

# THE JOURNAL of the Acoustical Society of America

Vol. 107, No. 6

June 2000

---

**SOUNDINGS SECTION**

<b>ACOUSTICAL NEWS—USA</b>	2937
USA Meetings Calendar	2937
<b>ACOUSTICAL NEWS—INTERNATIONAL</b>	2941
International Meetings Calendar	2941
<b>REPORTS OF RELATED MEETINGS</b>	2943
<b>REVIEWS OF ACOUSTICAL PATENTS</b>	2945

---

**GENERAL LINEAR ACOUSTICS [20]**

Experimental validations of the HELS method for reconstructing acoustic radiation from a complex vibrating structure	Nassif Rayess, Sean F. Wu	2955
Wave propagation in micro-heterogeneous porous media: A model based on an integro-differential wave equation	Andrzej Hanyga, Vladimir E. Rok	2965
Directive line source model: A new model for sound diffraction by half planes and wedges	Penelope Menounou, Ilene J. Busch-Vishniac, David T. Blackstock	2973
Limits of time-reversal focusing through multiple scattering: Long-range correlation	Arnaud Derode, Arnaud Tourin, Mathias Fink	2987
Analytical regularization based analysis of a spherical reflector symmetrically illuminated by an acoustic beam	Sergey S. Vinogradov, Elena D. Vinogradova, Alexander I. Nosich, Ayhan Altıntaş	2999
Effects of interaction between two bubble scatterers	George Kapodistrias, Peter H. Dahl	3006
Transformation of seismic velocity data to extract porosity and saturation values for rocks	James G. Berryman, Patricia A. Berge, Brian P. Bonner	3018
Analysis of multiscale scattering and poroelastic attenuation in a real sedimentary rock sequence	Chris L. Hackert, Jorge O. Parra	3028

**NONLINEAR ACOUSTICS [25]**

Model equation for strongly focused finite-amplitude sound beams	Tomoo Kamakura, Tsuneo Ishiwata, Kazuhisa Matsuda	3035
Propagation of acoustic pulses in material with hysteretic nonlinearity	Vitalyi Gusev	3047
Scattering of sonic booms by anisotropic turbulence in the atmosphere	Mark Kelly, Richard Raspet, Henry E. Bass	3059
The interaction of a laser-generated cavity with a solid boundary	S. J. Shaw, W. P. Schiffers, T. P. Gentry, D. C. Emmony	3065

(Continued)

## CONTENTS—Continued from preceding page

**AEROACOUSTICS, ATMOSPHERIC SOUND [28]**

Field measurements of sonic boom penetration into the ocean	R. A. Sohn, F. Vernon, J. A. Hildebrand, S. C. Webb	3073
Decay of large underwater bubble oscillations	B. Edward McDonald, Charles Holland	3084

**UNDERWATER SOUND [30]**

Regularized matched-mode processing for source localization	Nicole E. Collison, Stan E. Dosso	3089
Computed narrow-band time-reversing array retrofocusing in a dynamic shallow ocean	Michael R. Dungan, David R. Dowling	3101

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

Cell model calculations of dynamic drag parameters in packings of spheres	Olga Umnova, Keith Attenborough, Kai Ming Li	3113
Transient Lamb waves: Comparison between theory and experiment	M. Deschamps, O. Poncelet	3120
Nonlinear saturation of the thermoacoustic instability	S. Karpov, A. Prosperetti	3130
A thermoacoustic-Stirling heat engine: Detailed study	S. Backhaus, G. W. Swift	3148
Dynamics of gas bubbles in viscoelastic fluids. I. Linear viscoelasticity	John S. Allen, Ronald A. Roy	3167

**TRANSDUCTION [38]**

Diffractive acoustic elements for laser ultrasonics	M. Clark, S. D. Sharples, Michael G. Somekh	3179
---	---	------

**STRUCTURAL ACOUSTICS AND VIBRATION [40]**

Power flow in coupled bending and longitudinal waves in beams	Richard P. Szwerc, Courtney B. Burroughs, Stephen A. Hambric, Timothy E. McDevitt	3186
Influence of circumferential partial coating on the acoustic radiation from a fluid-loaded shell	J. M. Cuschieri, D. Feit	3196
Acoustic scattering from fluid-loaded stiffened cylindrical shell: Analysis using elasticity theory	André Baillard, Jean-Marc Conoir, Dominique Décultot, Gérard Maze, Aleksander Klauson, Jaan Metsaveer	3208

**NOISE: ITS EFFECTS AND CONTROL [50]**

Refined multiload method for measuring acoustical source characteristics of an intake or exhaust system	Seung-Ho Jang, Jeong-Guon Ih	3217
Control performance and robustness of an active noise control system for uncertain primary sound fields	Doo-Byung Yoon, Yang-Hann Kim	3226
A relaxed condition for “perfect” cancellation of broadband noise in 3D enclosures	Jing Yuan	3235
Community reaction to aircraft noise: Time-of-day penalty and tradeoff between levels of overflights	Henk M. E. Miedema, Henk Vos, Ronald G. de Jong	3245

**ARCHITECTURAL ACOUSTICS [55]**

Spatial-temporal correlation of a diffuse sound field	Boaz Rafaely	3254
---	--------------	------

**ACOUSTICAL MEASUREMENTS AND INSTRUMENTATION [58]**

Ultrasound field estimation method using a secondary source-array numerically constructed from a limited number of pressure measurements	Xiaobing Fan, Eduardo G. Moros, William L. Straube	3259
--	--	------

## CONTENTS—Continued from preceding page

**ACOUSTIC SIGNAL PROCESSING [60]**

<b>Auditory perception of objects by blind persons, using a bioacoustic high resolution air sonar</b>	Leslie Kay	3266
<b>Maximum <i>a posteriori</i> deconvolution of ultrasonic signals using multiple transducers</b>	Tomas Olofsson, Tadeusz Stepinski	3276
<b>Design of an optimal wave-vector filter for enhancing the resolution of reconstructed source field by near-field acoustical holography (NAH)</b>	Bong-Ki Kim, Jeong-Guon Ih	3289

**PHYSIOLOGICAL ACOUSTICS [64]**

<b>Group delays of distortion product otoacoustic emissions: Relating delays measured with <math>f_1</math>- and <math>f_2</math>-sweep paradigms</b>	Vera F. Prijs, Sandra Schneider, Ruurd Schoonhoven	3298
<b>The behavior of spontaneous otoacoustic emissions during and after postural changes</b>	Emile de Kleine, Hero P. Wit, Pim van Dijk, Paul Avan	3308
<b>Study of mechanical motions in the basal region of the chinchilla cochlea</b>	William S. Rhode, Alberto Recio	3317

**PSYCHOLOGICAL ACOUSTICS [66]**

<b>Efficient across-frequency integration: Evidence from psychometric functions</b>	Michelle L. Hicks, Søren Buus	3333
<b>Temporal dynamics of pitch strength in regular-interval noises: Effect of listening region and an auditory model</b>	L. Wiegrebe, H. S. Hirsch, R. D. Patterson, H. Fastl	3343
<b>Relative loudness perception of low and high frequency sounds in the open and occluded ear</b>	Gitte Keidser, Rickard Katsch, Harvey Dillon, Frances Grant	3351
<b>An effect of temporal asymmetry on loudness</b>	G. Christopher Stecker, Ervin R. Hafter	3358
<b>Perceptual classification based on the component structure of song in European starlings</b>	Timothy Q. Gentner, Stewart H. Hulse	3369
<b>Observer weighting of monaural level information in a pair of tone pulses</b>	Mark A. Stellmack, Neal F. Viemeister	3382
<b>Integration of monaural and binaural evidence of vowel formants</b>	Michael A. Akeroyd, A. Quentin Summerfield	3394
<b>Cross-correlation procedures for measuring noise and distortion in AGC hearing aids</b>	James M. Kates	3407

**SPEECH PRODUCTION [70]**

<b>Lexical boundary error analysis in hypokinetic and ataxic dysarthria</b>	Julie M. Liss, Stephanie M. Spitzer, John N. Caviness, Charles Adler, Brian W. Edwards	3415
<b>Effect of speaking rate and contrastive stress on formant dynamics and vowel perception</b>	Michel Pitermann	3425
<b>Acoustic effects of variation in vocal effort by men, women, and children</b>	Hartmut Traunmüller, Anders Eriksson	3438

**MUSIC AND MUSICAL INSTRUMENTS [75]**

<b>On operating deflection shapes of the violin body including in-plane motions</b>	Anna Runnemalm, Nils-Erik Molin, Erik Jansson	3452
<b>Vibration characteristics of pipe organ reed tongues and the effect of the shallot, resonator, and reed curvature</b>	G. R. Plitnik	3460

**BIOACOUSTICS [80]**

<b>Anisotropy of ultrasonic backscatter and attenuation from human calcaneus: Implications for relative roles of absorption and scattering in determining attenuation</b>	Keith A. Wear	3474
---	---------------	------

## CONTENTS—Continued from preceding page

<b>Cavitation nucleation agents for nonthermal ultrasound therapy</b>	Douglas L. Miller, Oliver D. Kripfgans, J. Brian Fowlkes, Paul L. Carson	3480
<b>Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985–1996</b>	G. Pavan, T. J. Hayward, J. F. Borsani, M. Priano, M. Manghi, C. Fossati, J. Gordon	3487
<b>Temporary shift in masked hearing thresholds of bottlenose dolphins, <i>Tursiops truncatus</i>, and white whales, <i>Delphinapterus leucas</i>, after exposure to intense tones</b>	Carolyn E. Schlundt, James J. Finneran, Donald A. Carder, Sam H. Ridgway	3496
<b>Calibration and comparison of the acoustic location methods used during the spring migration of the bowhead whale, <i>Balaena mysticetus</i>, off Pt. Barrow, Alaska, 1984–1993</b>	Christopher W. Clark, William T. Ellison	3509
<b>Recognizing transient low-frequency whale sounds by spectrogram correlation</b>	David K. Mellinger, Christopher W. Clark	3518
<b>Sound scattering and localized heat deposition of pulse-driven microbubbles</b>	Sascha Hilgenfeldt, Detlef Lohse, Michael Zomack	3530
<b>The impact of sound speed errors on medical ultrasound imaging</b>	M. E. Anderson, M. S. McKeag, G. E. Trahey	3540
<b>LETTERS TO THE EDITOR</b>		
<b>Displacement in a continuously twisted structurally chiral medium due to axial loading [20]</b>	Akhlesh Lakhtakia, Joseph A. Sherwin	3549
<b>An inexpensive passive acoustic system for recording and localizing wild animal sounds [30]</b>	Sean A. Hayes, David K. Mellinger, Donald A. Croll, Daniel P. Costa, J. Fabrizio Borsani	3552
<b>Sidelobe reduction of limited diffraction beams with Chebyshev aperture apodization [35]</b>	Shiping He, Jian-yu Lu	3556
<b>Real-time and off-line comparisons of standard array configurations containing three and four microphones [60]</b>	M. W. Hoffman, C. Pinkelman, X. F. Lu, Z. Li	3560
<b>Results of an independent analysis of the inverse beamformer for use in the advanced processor build process [60]</b>	G. Clifford Carter	3564
<b>Localization of amplitude-modulated high-frequency noise [66]</b>	Geoff Eberle, Ken I. McAnally, Russell L. Martin, Patrick Flanagan	3568
<b>Individual differences in speech adaptation to an artificial palate [70]</b>	Shari R. Baum, David H. McFarland	3572
<b>An analytical error invalidates the “depolarization” of the perceptual magnet effect [71]</b>	Frank H. Guenther	3576
<b>Reply to “An analytical error invalidates the ‘depolarization’ of the perceptual magnet effect” [J. Acoust. Soc. Am. 107, 3576–3577 (2000)] [71]</b>	Andrew J. Lotto	3578
<b>Frequency beamforming of dolphin whistles using a sparse three-element towed array [80]</b>	Aaron Thode, Thomas Norris, Jay Barlow	3581
<b>ACOUSTICS RESEARCH LETTERS ONLINE</b>		
<b>Design and construction of a solar-powdered, thermoacoustically driven, thermoacoustic refrigerator</b>	Jay A. Adeff, Thomas J. Hofler	L37
<b>INDEX TO VOLUME 107</b>		3585
<b>SUBJECT INDEX TO VOLUME 107</b>		3590
<b>AUTHOR INDEX TO VOLUME 107</b>		3615

# Design and construction of a solar-powered, thermoacoustically driven, thermoacoustic refrigerator

Jay A. Adeff, Thomas J. Hofler

*Department of Physics, Naval Postgraduate School, Monterey, CA 93943, USA  
adeff@nps.navy.mil tjhofler@nps.navy.mil*

**Abstract:** A thermoacoustically driven thermoacoustic refrigerator powered by solar thermal energy has been successfully built and tested. A 0.457 m diameter Fresnel lens focuses sunlight onto the hot end of a 0.0254 m diameter reticulated vitreous carbon prime mover stack, heating it to 475°C, thereby eliminating the need for the most troublesome component in a heat driven prime mover, the hot heat exchanger. The high intensity sound waves produced by the prime mover drive a thermoacoustic refrigerator to produce 2.5 watts of cooling power at a cold temperature of 5°C and a temperature span of 18°C.

© 2000 Acoustical Society of America

**PACS numbers:** 43.10.Ln, 43.35.Ud [HEB]

## Introduction

Recent research efforts in thermoacoustic refrigeration have moved away from the use of electrodynamic loudspeakers as a source of the high intensity sound waves required to power these devices. In particular, the authors have been successful in building a thermoacoustically driven thermoacoustic refrigerator (TADTAR), which is powered by electric heater cartridges delivering 600 watts of heat at 450°C directly to the hot heat exchanger of the prime mover section, while producing 90 watts of useful cooling power at a cold temperature of 0°C. With this device, the reliability of thermoacoustic refrigerators has been substantially increased because of a lack of moving parts associated with an electrodynamic loudspeaker. However, a new problem has been created as a result of the elevated temperatures at which the prime mover must operate.

A heat driven TADTAR requires that the prime mover's hot heat exchanger be capable of functioning at very high temperatures. In fact, the hotter this heat exchanger is capable of operating, the more efficiently the prime mover will function. The reader is referred to Swift's review and tutorial on thermoacoustic engines<sup>1</sup> for a complete description of the physics behind thermoacoustic engines. There is, however, a fundamental temperature limit to any material used for the heat exchanger. Copper, the most commonly used material because of its low thermal resistance, is capable only of withstanding temperatures up to 450°C before becoming too soft to maintain its structural integrity. High temperature alloys such as stainless steel require specialized fabrication techniques, which can be costly and time consuming and are not always readily available to researchers on a tight budget. Worse yet, these materials typically have very high thermal resistances relative to copper. Another possibility would be to apply electric heater wires directly to the hot end of the stack. However, the authors felt that the simplest solution to this problem was to replace the prime mover hot heat exchanger and any associated heater elements with direct solar heating.

### Direct solar heating of the prime mover

The idea to use sunlight concentrated and focused by an optical lens as a source of thermal energy, is not new. People have made use of just such a device for cooking food or heating water for many years. It is no great surprise, therefore, that a thin plastic Fresnel lens can be used to focus sunlight directly onto the hot end of a thermoacoustic prime mover or, better yet, directly onto the hot end of the prime mover stack. In the latter case, no heat exchanger is required to transfer the heat to the stack.

Reticulated vitreous carbon (RVC) was introduced by the authors as a stack material for thermoacoustic engines,<sup>2</sup> replacing previous materials such as plastic roll stacks in refrigerators<sup>3</sup> and wire mesh stacks in prime movers.<sup>4</sup> An RVC stack is not only easier and less labor intensive to produce, it also has the ability to withstand extremely high temperatures in non-oxidizing environments making it an ideal material for high temperature prime movers. In addition, RVC has a random fibrous structure that eliminates any optical path for light to penetrate directly through it, which is particularly attractive for direct solar heating.

### Experimental Apparatus

A low-powered prototype of a small temperature span, solar thermal powered TADTAR was designed, built, and tested at the Naval Postgraduate School in a period of only six months with the aid of a graphical numerical thermoacoustic engine simulation program developed by Purdy and Hofler.<sup>5</sup> The goal for this project was to build a compact device capable of producing five to ten watts of cooling power at 0°C and prove that a solar powered thermoacoustic refrigerator is not only possible but potentially very practical, simple, and reliable.

The project's objective was not necessarily to optimize the design of the refrigerator. Rather it was to arrive at a simple design that would have a high probability of working the first time out and produce a measurable amount of refrigeration. The resonator vessel measures just 0.318 m in length and has a stack diameter of 0.0254 m. A 0.457 m diameter Fresnel lens with a focal length of 0.61 m was used as the solar collector, delivering an estimated 100 watts of solar thermal power of heat energy to the prime mover. A 60-pore per linear inch (ppi) RVC stack was used for the prime mover, whereas a conventional plastic roll stack with a plate spacing of 0.38 mm was chosen for the refrigerator. To keep the device entirely solar powered, solar voltaic cells were used to power electric cooling fans impinging on external heat sinks to reject heat from the prime mover and refrigerator ambient heat exchangers. The final design, shown in Figure 1, is of a half-wave resonant tube with pressure anti-nodes at the closed ends.

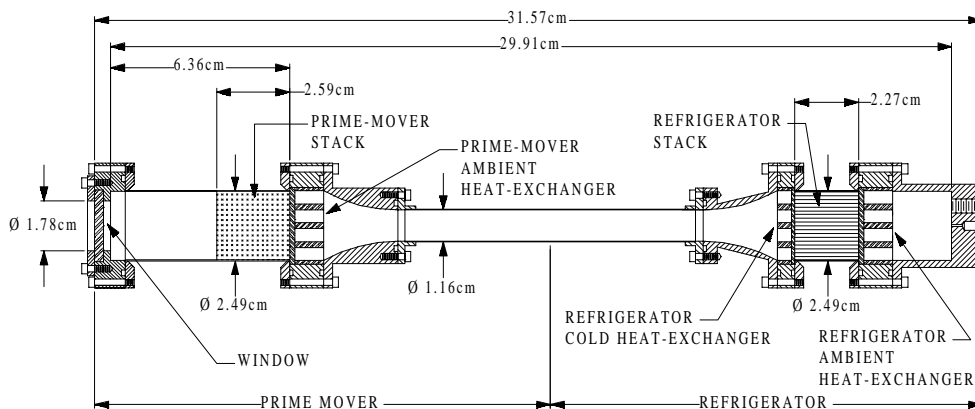


Fig. 1. An assembly section of the solar powered TADTAR

The solar TADTAR is designed to operate with an inert gas mixture of 18% argon and a balance of helium at a pressure of six bar absolute. The two ambient heat exchangers, the refrigerator cold heat exchanger, and their respective flanges were all made of copper, whereas the remainder of the resonator was made of stainless steel. Thin walled stainless steel tubes were used for the stack sections to limit heat conduction along the resonator wall, while still providing enough strength to withstand temperatures of up to 600°C without the danger of rupturing. A fused quartz window was located in the prime mover end plate to allow the focused sunlight to be projected directly onto the hot end of the prime mover stack. The diameter of this window was kept to a minimum by locating the focal point of the Fresnel lens close to the window. Three thermocouples were inserted directly into the heat exchangers, and a fourth was attached to the outside of the prime mover resonator wall at a location corresponding to the hot end of the prime mover stack. Finally, high temperature ceramic fiber insulation was used to cover the entire prime mover tube. Figure 2 shows the TADTAR ready to operate.

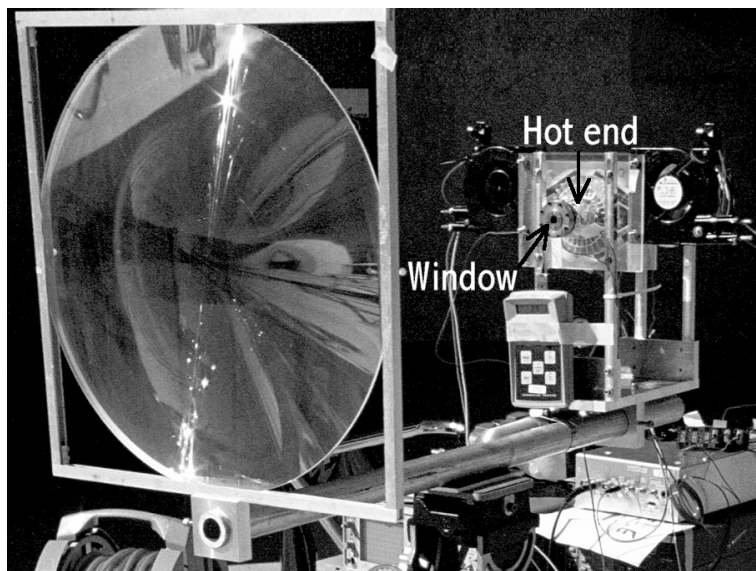


Fig. 2. The solar powered TADTAR ready to run.

### Measurements

The solar-powered TADTAR was run during the first two weeks of October when there was approximately 650 W/m<sup>2</sup> of solar heat flux available or 100 watts from the solar collector itself. A microphone placed at the refrigerator end of the resonator indicated that a peak acoustic pressure amplitude of 4.3% of the resonator mean pressure or 26.5 kPa was being generated by the prime mover. This was obtained with the hot and ambient ends of the prime mover at an equilibrium temperature of 480°C and 35°C, respectively. The procedure for quantifying the refrigerator's performance was to measure the cold heat exchanger's temperature as a function of time while the device was cooling down and again while it warmed back up to room temperature after being shut off. These measurements could then be used to estimate the cooling power and heat leak.<sup>6</sup>

The rate at which heat is pumped away from the cold end of the refrigerator,  $Q_{TOTAL}$ , can be determined from the following relation:

$$Q_{TOTAL} = \frac{dT_C}{dt} C_p, \quad (1)$$

where  $dT_c/dt$  is the time rate of change of the refrigerator's cold end temperature, and  $C_p$  is the total heat capacity of the cold end. The heat capacity is determined by multiplying the measured mass of the cold end by the heat capacity of the material from which it has been fabricated.

The time rate of change of the refrigerator's cold end temperature can be obtained by assuming that sections of the cool-down or warm-up data can be fitted with an exponential curve having the form

$$T_C(t) = T_0 + A_0 e^{-(t-t_0)/\tau}, \quad (2)$$

where  $T_C(t)$  is the cold temperature as a function of time,  $t_0$  is the time at which the experiment begins, and  $\tau$  is the cool-down time constant.  $T_0$  is a limiting temperature value that would be achieved in steady state equilibrium for very long time values. This limiting value is at or near room temperature for the warm-up data and is defined as  $T_{0,wu}$ , but it is the limiting coldest temperature observed for the cool-down data and is defined as  $T_{0,cd}$ .

The fact that the cool-down data can be fitted with an exponential curve is purely an empirical result and not derived from a theory for the thermoacoustic refrigeration mechanism. Taking the time derivative of Eq. (2) gives

$$\frac{dT_C}{dt} = \frac{-A_0}{\tau} e^{-(t-t_0)/\tau} = \frac{-1}{\tau} (\Delta T(t)), \quad (3)$$

where  $\Delta T(t) = T_c(t) - T_0$ . Substituting Eq. (3) into Eq. (1) gives

$$Q_{TOTAL} = \frac{-C_p}{\tau} \Delta T(t) = K \Delta T(t), \quad (4)$$

or

$$K = \frac{-C_p}{\tau}, \quad (5)$$

where  $K$  is defined as a total thermal conductance. The physical value of  $C_p$  for the refrigerator cold end is 70 Joules per degree centigrade.

For the warm-up data,  $K_{HL}$  is a conductance representing the total heat leak from the ambient end of the refrigerator stack to the cold end. For the cool-down data, we define  $K_{EC}$  as a thermal conductance, which corresponds to an "excess cooling" power given by the difference between the total cooling power of the refrigerator and the heat leak power from room temperature. The conductances are

$$K_{HL} = \frac{-C_p}{\tau_{wu}}, \quad (6)$$

and

$$K_{EC} = \frac{-C_p}{\tau_{cd}}, \quad (7)$$

where  $\tau_{wu}$  and  $\tau_{cd}$  are the warm-up and cool-down time constants, which are obtained by fitting an exponential curve of the form of Eq. (2) to the warm-up and cool-down data.

Figure 3 shows a plot of the refrigerator cold end temperature as a function of time, along with a curve fit to the data having the form of Eq. (2). Figure 4 shows a plot of the warm-up data taken immediately after this cool-down run, as well as a curve fit to the data.



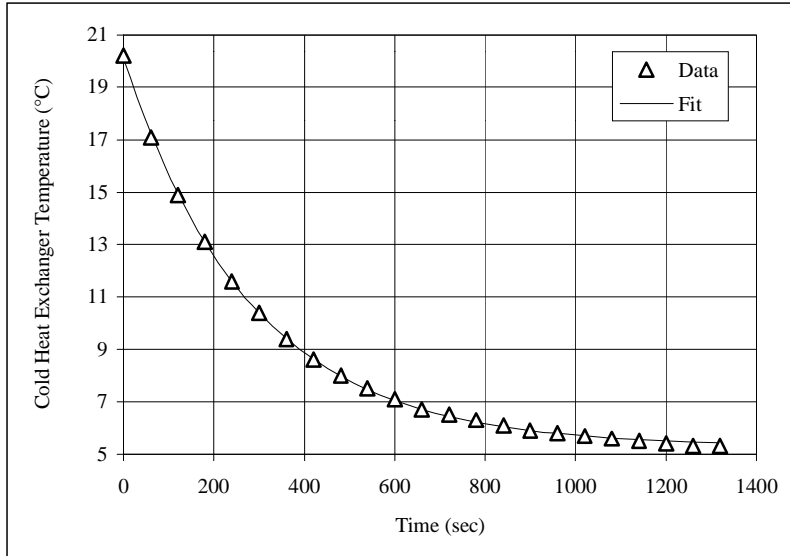


Fig. 3. A plot of the refrigerator cold heat exchanger temperature as a function of time showing the data and an exponential fit to the data.

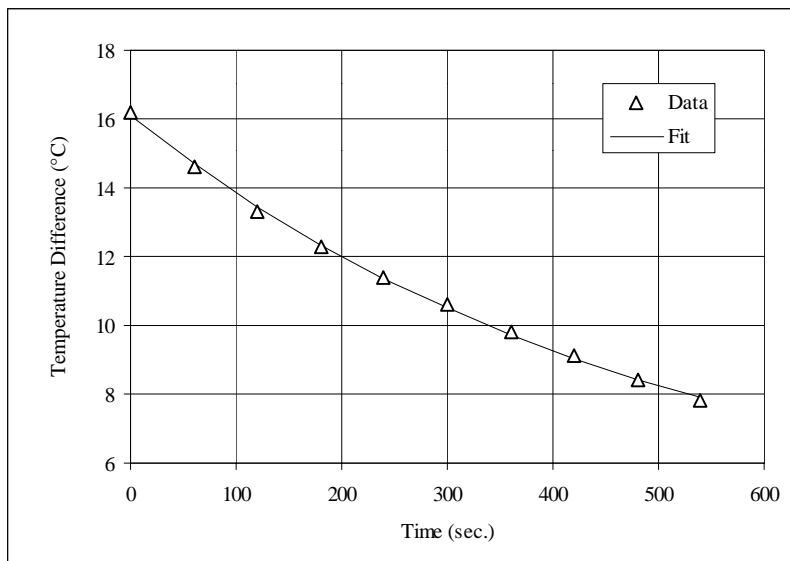


Fig. 4. A plot of the temperature difference between the refrigerator's ambient and cold heat exchangers as a function of time after the refrigerator had been shut down.

From the curve fits in Figures 3 and 4, the numbers obtained for the cool-down and warm-up time constants are:

$$\tau_{cd} = 281 \text{ s} \tag{8}$$

and

$$\tau_{wu} = 509 \text{ s}, \tag{9}$$

from which the total conductances are obtained:

$$K_{EC} = 0.249 \text{ W/}^\circ\text{C}, \quad (10)$$

$$K_{HL} = 0.138 \text{ W/}^\circ\text{C}. \quad (11)$$

The total cooling power is then

$$Q_{TC} = Q_{EC} + Q_{HL} = K_{EC} \Delta T_{EC} + K_{HL} \Delta T_{HL}. \quad (12)$$

At the largest span observed by the refrigerator of  $\Delta T_{HL} = 17.7^\circ\text{C}$  and  $\Delta T_{EC} = 0$ , the total cooling power was therefore 2.5 watts. Figure 5 shows a plot of the cooling power as a function of the cold heat exchanger temperature.

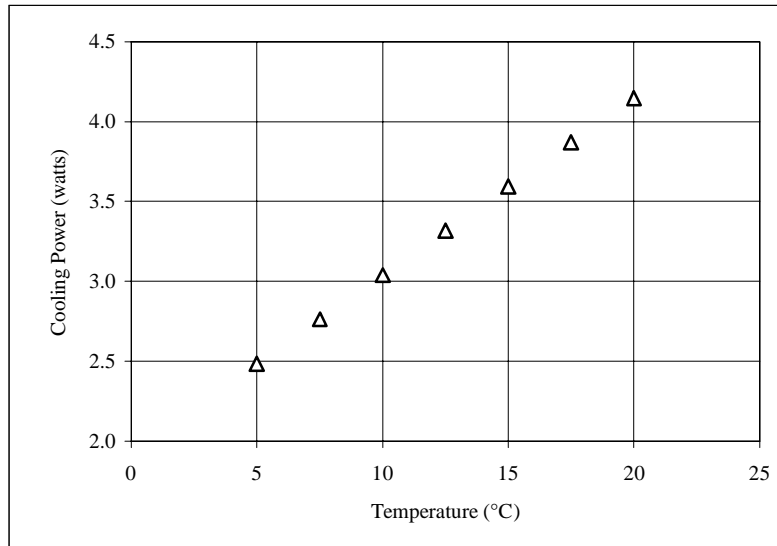


Fig. 5. A plot of the refrigerator total cooling power as a function of the cold heat exchanger temperature.

## Conclusions

Based on these measurements, we believe that solar-powered thermoacoustic refrigeration is a viable technology worthy of further investigation and significant improvements can be made in the cooling power available. This could be accomplished by using a larger solar collector to increase the amount of solar power available and improving the design of the heat exchangers and insulation. Such a device could be valuable in geographical areas where sunlight is abundant year round but electrical power is not. Ice could be manufactured during daylight hours and used to store perishables within an insulated icebox.

## References

- <sup>1</sup>G.W. Swift, "Thermoacoustic engines," *J. Acoust. Soc. Am.* **84**, 1145 (1988).
- <sup>2</sup>J.A. Adeff, T.J. Hofler, A.A. Atchley, W.C. Moss, "Measurements with reticulated vitreous carbon stacks in thermoacoustic prime movers and refrigerators," *J. Acoust. Soc. Am.* **104**(1), July 1998.
- <sup>3</sup>J.A. Adeff, "Measurement of the space thermoacoustic refrigerator performance," Master's thesis, Naval Postgraduate School, Monterey, California, 1990.
- <sup>4</sup>M.S. Reed and T.J. Hofler, "Measurements with wire mesh stacks in thermoacoustic prime movers," *J. Acoust. Soc. Am.* **99**, 2559 (1996).
- <sup>5</sup>Purdy, E.W., "Development of a graphical numerical simulation for thermoacoustic research," Master's thesis, Naval Postgraduate School, Monterey, California, 1998.
- <sup>6</sup>T.J. Berhow, "Construction and performance measurement of a portable thermoacoustic refrigerator demonstration apparatus," Master's thesis, Naval Postgraduate School, Monterey, California, 1994.

# 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.

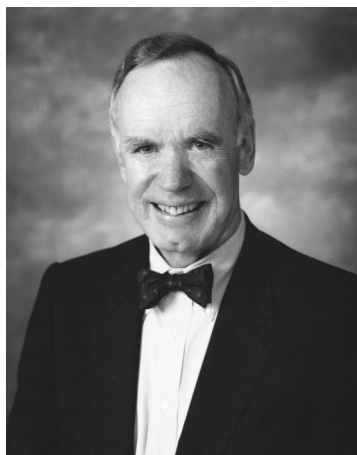
## ACOUSTICAL NEWS—USA

**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.*

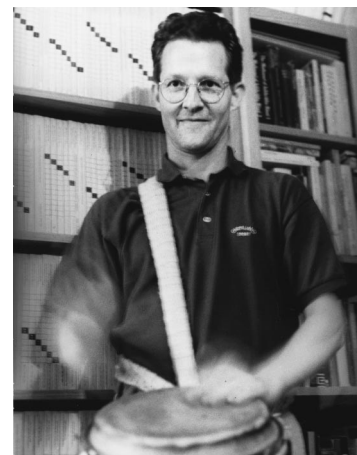
### New Fellows of the Acoustical Society of America



**Charles R. Greene**—For contributions to the understanding of the effects of noise on Arctic marine mammals.



**Avraham Hirschberg**—For contributions to understanding flow dynamics in musical instruments and in the human voice.



**Stephen E. McAdams**—For contributions to the perception and cognition of music.

### Pranab Saha receives award from Michigan Society of Professional Engineers

ASA member Pranab Saha was selected as the “2000 Outstanding Engineer in Private Practice” by the Michigan Society of Professional Engineers and the American Consulting Engineers Council of Michigan. Dr. Saha received his Ph.D. in Mechanical Engineering with a specialty in acoustics from Georgia Institute of Technology. He is a Principal Consultant and Co-owner of Kolano and Saha Engineers, Inc., an independent, professional engineering company located in Waterford, Michigan.

### 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**  
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, Mu-

27–28 July

2–4 August

21–23 September

22–25 October

4–8 December

sic Dept., Univ. of Central Florida, P.O. Box 161354, Orlando, FL 32816-1354; Tel.: 407-823-5116; E-mail: kkoons@pegasus.cc.ucf.edu].

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/].

2000 Ultrasonic Transducer Engineering Conference, University Park, PA [K. Kirk Shung, NIH Resource on Medical Ultrasonic Transducer Technology, Pennsylvania State Univ., 231 Hallowell Bldg., University Park, PA 16802; Tel.: 814-865-1407; Fax: 814-863-0490; E-mail: kksbio@engr.psu.edu].

Eighth Annual Conference on the Management of the Tinnitus Patient, Iowa City, IA [Richard Tyler; Tel.: 391-356-2471; E-mail: tyler@uiowa.edu; WWW: www.medicine.uiowa.edu/otolaryngology/news/news].

IEEE Ultrasonics Symposium, San Juan, Puerto Rico [R. Almar, 896 Buttonwood Ln., Altamonte Springs, FL 32714; Fax: 407-290-5181; WWW: www.uffcsymp2000.org].

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].

4–8 June

141st Meeting of the Acoustical Society of America, Chicago, IL [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].

30 April–3 May

**2001**

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.

7–10 October

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].

# ACOUSTICAL NEWS—INTERNATIONAL

Walter G. Mayer

Physics Department, Georgetown University, Washington, DC 20057

## Papers published in JASJ(E)

A listing of Invited Papers and Regular Papers appearing in the latest issue of the English language version of the *Journal of the Acoustical Society of Japan*, JASJ(E), was published for the first time in the January 1995 issue of the Journal. This listing is continued below.

The March issue of JASJ(E), Vol. 21, No. 2 (2000) contains the following contributions:

T. S. Hsu and K. A. Poornima, "Temperature prediction of the voice coil of a moving coil loudspeaker by computer simulation"

K. Okanoya, "Perception of missing fundamentals in zebra finches and Bengalese finches"

A. Preis, M. Ishibashi, and H. Tachibana, "Psychoacoustic studies on assessment of floor impact sounds"

K. Shinoda and T. Watanabe, "MDL-based context-dependent subword modeling for speech recognition"

T. Otsuru and R. Tomiku, "Basic characteristics and accuracy of acoustic element using spline function in finite element sound field analysis"

K. Oku, A. Yarai, and T. Nakanishi, "A new tuning method for glass harp based on a vibration analysis that uses a finite element method"

## International Meetings Calendar

Below are announcements of meetings to be held abroad. Entries preceded by an \* are new or updated listings with full contact addresses given in parentheses. *Month/year* listings following other entries refer to meeting announcements, with full contact addresses, which were published in previous issues of the *Journal*.

### June 2000

5-9 **International Conference on Acoustics, Speech and Signal Processing (ICASSP-2000)**, Istanbul. (Fax: +1 410 455 3969; Web: [icassp2000.sdsu.edu](http://icassp2000.sdsu.edu)) 6/99

6-9 **5th International Symposium on Transport Noise and Vibration**, St. Petersburg. (Fax: +7 812 127 9323; E-mail: [noise@mail.rcom.ru](mailto:noise@mail.rcom.ru)) 6/99

14-17 **IUTAM Symposium on Mechanical Waves for Composite Structures Characterization**, Chania. (Fax: +30 821 37438; Web: [www.tuc.gr/iutam](http://www.tuc.gr/iutam)) 10/99

### July 2000

4-7 **7th International Congress on Sound and Vibration**, Garmisch-Partenkirchen. (Fax: +49 531 295 2320; Web: [www.iiav.org/icsv7.html](http://www.iiav.org/icsv7.html)) 12/98

9-13 **19th International Congress on Education of the Deaf & 7th Asia-Pacific Congress on Deafness**, Sydney. (Fax: +61 2 9262 3135; Web: [www.iced2000.com](http://www.iced2000.com)) 4/00

10-13 **5th European Conference on Underwater Acoustics**, Lyon. (Fax: +33 4 72 44 80 74; Web: [www.ecua2000.cpe.fr](http://www.ecua2000.cpe.fr)) 12/99

10-14 **5th International Conference on Mathematical and Numerical Aspects of Wave Propagation**, Santiago de Compostela. (Web: [www.usc.es/waves2000](http://www.usc.es/waves2000)) 4/00

### August 2000

28-30 **INTER-NOISE 2000**, Nice. (Fax: +33 1 47 88 90 60; \*New Web: [internoise2000.loa.espci.fr](http://internoise2000.loa.espci.fr)) 6/99

31-2 **International Conference on Noise & Vibration Pre-Design and Characterization Using Energy Methods (NOVEM)**, Lyon. (Fax: +33 4 72 43 87 12; Web: [www.insa-lyon.fr/laboratories/lva.html](http://www.insa-lyon.fr/laboratories/lva.html)) 6/99

### September 2000

3-6

**5th French Congress on Acoustics—Joint Meeting of the Swiss and French Acoustical Societies**, Lausanne. (Fax: +41 216 93 26 73) 4/99

13-15

**International Conference on Noise and Vibration Engineering (ISMA 25)**, Leuven. (Fax: +32 16 32 24 82; E-mail: [lieve.notre@mech.kuleuven.ac.be](mailto:lieve.notre@mech.kuleuven.ac.be)) 12/99

17-21

**Acoustical Society of Lithuania 1st International Conference**, Vilnius. (Fax: +370 2 223 451; E-mail: [daumantas.ciblys@ff.vu.lt](mailto:daumantas.ciblys@ff.vu.lt)) 8/99

18-22

**47th Seminar on Acoustics (OSA2000)**, Zalew Solinski, Poland. (E-mail: [osa@atena.univ.rzeszow.pl](mailto:osa@atena.univ.rzeszow.pl)) 4/00

29-1

**\*Tone Wood Forum**, Quesnel, BC, Canada. (J. Griffin, Island Mountain Arts, P.O. Box 65, Wells, BC V0K 2R0, Canada; Fax: +1 250 994 3433; Web: [www.imarts.com](http://www.imarts.com))

### October 2000

3-5

**WESTPRAC VII**, Kumamoto. (Web: [cogni.eecs.kumamoto-u.ac.jp/others/westprac7](http://cogni.eecs.kumamoto-u.ac.jp/others/westprac7)) 6/98

3-6

**EUROMECH Colloquium on Elastic Waves in NDT**, Prague. (Fax: +420 2 858 4695; E-mail: [ok@bivoj.it.cas.cz](mailto:ok@bivoj.it.cas.cz)) 10/99

12-14

**International Conference on Newborn Hearing Screening**, Milan. (Fax: +39 2 23993367/60; Web: [www.biomed.polimi.it/nh2000](http://www.biomed.polimi.it/nh2000)) 2/00

16-18

**2nd Iberoamerican Congress on Acoustics, 31st National Meeting of the Spanish Acoustical Society, and EAA Symposium**, Madrid. (Fax: +34 91 411 7651; E-mail: [ssantiago@fresno.csic.es](mailto:ssantiago@fresno.csic.es)) 12/98

16-20

**6th International Conference on Spoken Language Processing**, Beijing. (Fax: +86 10 6256 9079; Web: [www.icslp2000.org](http://www.icslp2000.org)) 10/98

### November 2000

24-27

**\*21st Tonmeistertagung (VDT International Audio Convention)**, Hannover, Germany. (Convention Office VDT, Am Zaarshäuschen 9, 51427 Bergisch-Gladbach, Germany; Fax: +49 2204 21584; Web: [www.tonmeister.de](http://www.tonmeister.de))

### August 2001

28-30

**INTER-NOISE 2001**, The Hague. (Web: [internoise2001.tudelft.nl](http://internoise2001.tudelft.nl)) 6/99

### September 2001

2-7

**17th International Congress on Acoustics (ICA)**, Rome. (Fax: +39 6 4424 0183; Web: [www.uniroma1.it/energ/ica.html](http://www.uniroma1.it/energ/ica.html)) 10/98

10-13

**International Symposium on Musical Acoustics (ISMA 2001)**, Perugia. (Fax: +39 75 577 2255; E-mail: [perusia@classico.it](mailto:perusia@classico.it)) 10/99

### October 2001

17-19

**32nd Meeting of the Spanish Acoustical Society**, La Rioja. (Fax: +34 91 411 76 51; Web: [www.ia.csic.es/sea/index.html](http://www.ia.csic.es/sea/index.html)) 10/99

### September 2002

16-21

**Forum Acusticum 2002 (Joint EAA-SEA-ASJ Meeting)**, Sevilla. (Fax: +34 91 411 7651; Web: [www.cica.es/aliens/forum2002](http://www.cica.es/aliens/forum2002)) 2/00

# REVIEWS OF ACOUSTICAL PATENTS

**Lloyd Rice**

11222 Flatiron Drive, Lafayette, Colorado 80026

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

## Reviewers for this issue:

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

RONALD B. COLEMAN, *BBN Technologies, 70 Fawcett Street, Cambridge, Massachusetts 02138*

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

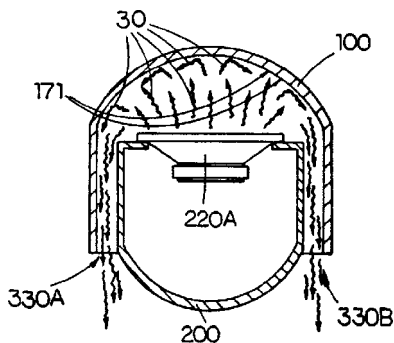
**5,933,509**

### 43.38.Ja BAND PASS SPEAKER

George Wu, Simi Valley, California

3 August 1999 (Class 381/345); filed 18 April 1997

In a sealed/vented bandpass system, the loudspeaker is buried inside the enclosure and a removable panel must be provided to install the speaker. This invention configures the design as a sealed box nested in a shell. The space between the two defines the front chamber and the vent. The assembly is held together by two bolts and no airtight gaskets are required. A simple,



clever idea. However, the patent document is not content with simplicity. It would have us believe that, "The present invention is a breakthrough in the new generation of tight and focused bass...driving all the sound waves through and out a sound wave focusing pocket." Oh well, the geometry is still clever.—GLA

**5,940,347**

### 43.38.Ja DIRECTED STICK RADIATOR

Hans-Joachim Raida, Koelm, Germany and Oscar Friedrich Bschorr, Munich, Germany

17 August 1999 (Class 367/138); filed in Germany 26 November 1996

An elongated mechanical delay line is driven at one end by an electrodynamic actuator and terminated at the other end by passive or active means. In between, a multiplicity of "transformer elements" convert mechanical energy into acoustical energy. By adjusting propagation velocity and the properties of transformer elements, a variety of directional characteristics can be achieved. The patent includes numerous charts, diagrams, and impressive math equations.

So far, so good. However, the inventors assume that such a device not only can be smaller and more efficient than an array of conventional loudspeakers, but can have virtually unlimited bandwidth extending from infrasonic to ultrasonic frequencies. There is no hint that a prototype has actually been built or tested, but every conceivable variation is covered in at least one of the 63 patent claims.—GLA

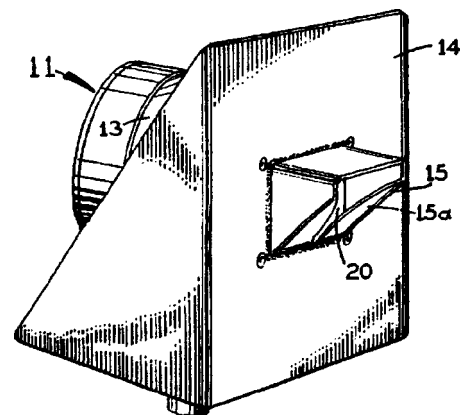
**5,943,431**

### 43.38.Ja LOUDSPEAKER WITH TAPERED SLOT COUPLER AND SOUND REPRODUCTION SYSTEM

Alan Weiss and Thomas Brown, both of North Lauderdale, Florida

24 August 1999 (Class 381/305); filed 6 March 1997

Conventional loudspeaker 11 is mounted to baffle board 14. Sound is squirted through a small aperture and emanates from a miniaturized Karlson tapered slot coupler 15, which looks somewhat like the pouring spout on a box of salt. The patent asserts that this little gadget has the miraculous



property of maintaining uniformly wide horizontal coverage and well-defined vertical coverage regardless of frequency. The patent was issued so it must be true.—GLA

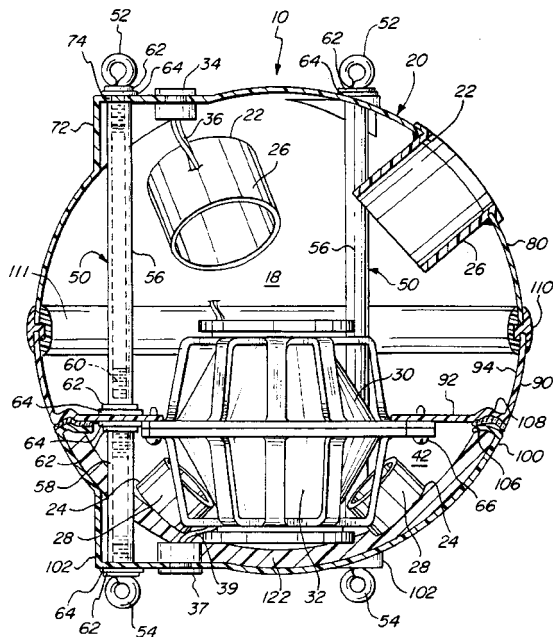
**5,952,620**

### 43.38.Ja OMNI-DIRECTIONAL SUB-BASS LOUDSPEAKER

Peter A. Hamilton and Mikhail Malkovitch, assignors to Sonic Systems, Incorporated

14 September 1999 (Class 181/153); filed 22 October 1997

By configuring the Bose two-chamber bandpass design (United States Patent 4,549,631) as a generally spherical enclosure, the inventors have



managed to achieve nondirectional propagation of very low frequencies. What won't they think of next?—GLA

**5,930,373**

**43.38.Lc METHOD AND SYSTEM FOR ENHANCING QUALITY OF SOUND SIGNAL**

**Meir Shashoua and Daniel Glotter, assignors to K. S. Waves Limited**  
**27 July 1999 (Class 381/98); filed 4 April 1997**

Stimulated by a suitable series of harmonics, the human hearing mechanism will perceive a fundamental frequency even if the actual fundamental is absent. The phenomenon has been used to generate "synthetic bass" in pipe organs and later in radio receivers and phonographs. This patent brings the concept into the computer age with sophisticated psycho-acoustic algorithms. The invention is interesting and the patent is clearly written.—GLA

**5,943,427**

**43.38.Vk METHOD AND APPARATUS FOR THREE DIMENSIONAL AUDIO SPATIALIZATION**

**Dana C. Massie et al., assignors to Creative Technology Limited**  
**24 August 1999 (Class 381/17); filed 21 April 1995**

The patent describes a sophisticated method for generating virtual sound sources (including moving sources) from two speakers. "Three-dimensional audio spatialization capabilities can be provided to video games, multimedia computer systems, virtual reality, cinema sound systems, home theater, and home digital audio systems."—GLA

**5,949,894**

**43.38.Vk ADAPTIVE AUDIO SYSTEMS AND SOUND REPRODUCTION SYSTEMS**

**Philip Arthur Nelson et al., assignors to Adaptive Audio Limited**  
**7 September 1999 (Class 381/300); filed 18 March 1997**

One general class of playback equalization methods uses adaptive digital filters to simultaneously correct time and frequency domain irregularities. The result can be startlingly effective so long as nothing changes, including temperature, humidity, and the exact positions of loudspeakers and listener. Numerous previous patents address the problem of opening up the listening area while retaining the benefits of equalization. The invention

at hand is a fairly complex method of equalizing a practical multi-channel playback system in which loudspeakers are not located symmetrically in relation to the preferred listening location.—GLA

**5,970,152**

**43.38.Vk AUDIO ENHANCEMENT SYSTEM FOR USE IN A SURROUND SOUND ENVIRONMENT**

**Arnold I. Klayman, assignor to SRS Labs, Incorporated**  
**19 October 1999 (Class 381/1); filed 30 April 1996**

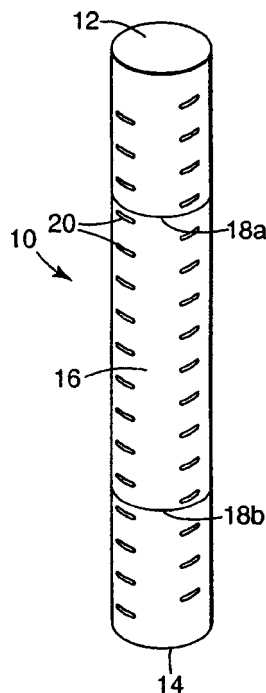
The inventor holds several previous patents relating to stereo sound enhancement. This latest invention is intended to improve 5.1 multi-channel sound reproduction by using dynamic electronic processing to create a more diffuse surround sound field and smoother panning between the five full-range loudspeakers.—GLA

**5,945,642**

**43.38.Zc ACOUSTIC HORN**

**Satinder K. Nayar et al., assignors to Minnesota Mining and Manufacturing Company**  
**31 August 1999 (Class 181/184); filed 13 March 1998**

Funny, it doesn't look like a horn. Most readers will assume that an acoustic horn operates in air. In this case, however, the horn is a machine tool made of solid aluminum or titanium. It is used to couple mechanical energy at ultrasonic frequencies. The invention is an improved design incor-



porating cutouts 20 distributed along its length, thus establishing a natural resonance frequency lower than the nominal halfwave frequency of the device.—GLA

**5,899,443**

**43.40.Vn PASSIVE-ACTIVE VIBRATION ISOLATION**

**Jen-Houne Hannsen Su, assignor to the United States of America as represented by the Secretary of the Navy**  
**4 May 1999 (Class 267/140.14); filed 22 October 1996**

An activated air mount is described. The foundation plate of the passive air mount is equipped with vibration sensors and inertial actuators. The nominally collocated sensors and actuators are used in a feedback control

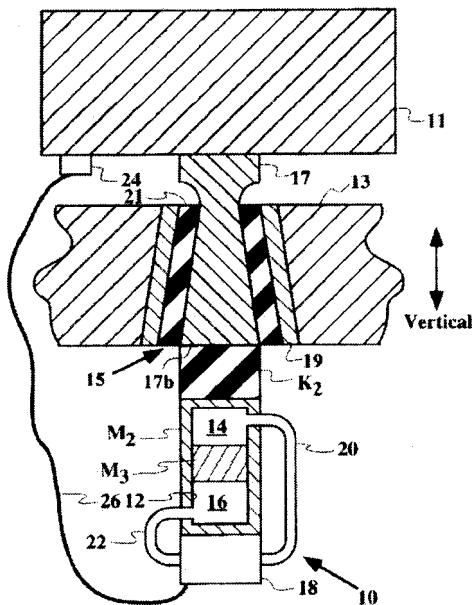
system to reduce transmissibility of the mount. A similar approach is discussed for use with elastomeric mounts, and for controlling multiple response directions of the foundation plate.—RBC

5,906,254

**43.40.Vn ACTIVE SYSTEMS AND DEVICES INCLUDING ACTIVE VIBRATION ABSORBERS (AVAS)**

Warren E. Schmidt *et al.*, assignors to Lord Corporation  
25 May 1999 (Class 188/378); filed 11 October 1995

A flexibly-mounted active vibration absorber (AVA) is described that has a secondary controllable mass  $M_3$ . If the passive resonance of the mass  $M_2$  of the tuned absorber is insufficient to control the input vibration, a secondary mass  $M_3$  is vibrated at a frequency, amplitude, and phase suffi-



cient to make up the deficiency in mass  $M_2$ . Many embodiments are described that depict variations on the first spring used for flexibly mounting the AVA, the actuators used to drive the AVA, and features of the attachment to a vibrating member.—RBC

5,984,062

**43.40.Vn METHOD FOR CONTROLLING AN ACTIVE TRUSS ELEMENT FOR VIBRATION SUPPRESSION**

James E. Bobrow and Faryar Jabbari, both of Irvine, California  
16 November 1999 (Class 188/378); filed 24 February 1995

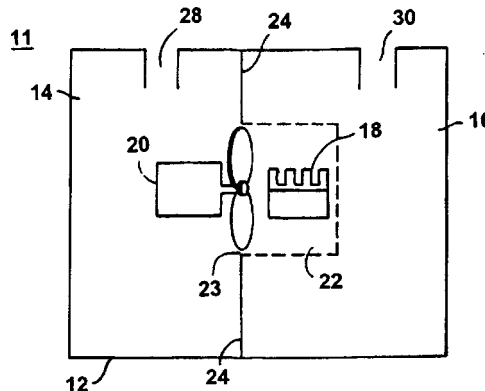
An active truss element and control law are described to dissipate energy during the motion of a structure, such as a building excited by seismic excitation. A decentralized switching control law is used along with a compressible fluid in the truss element. The truss element retains its maximum stiffness, but has a resettable nominal unstressed length. Energy is absorbed in the working fluid of the truss element through heat transfer to the environment when the nominal length is reset. The control law is claimed to be insensitive to changes in structural parameters such as mass, stiffness, and damping.—RBC

5,792,999

**43.50.Gf NOISE ATTENUATING IN PORTED ENCLOSURE**

Finn Arnold and Stephen R. O'Dea, assignors to Bose Corporation  
11 August 1998 (Class 181/141); filed 23 January 1997

Yes, a Helmholtz resonator can be used to attenuate the noise of fan 20 without blocking the flow of cooling air passing over electronic circuitry 18.



It can also do double duty by being combined with the bandpass speaker system from which it was derived.—GLA

5,917,919

**43.50.Ki METHOD AND APPARATUS FOR MULTI-CHANNEL ACTIVE CONTROL OF NOISE OR VIBRATION OR OF MULTI-CHANNEL SEPARATION OF A SIGNAL FROM A NOISY ENVIRONMENT**

Felix Rosenthal, Annandale, Virginia  
29 June 1999 (Class 381/71.11); filed 3 December 1996

A frequency-domain algorithm for multi-channel feedforward control or signal separation is described. The approach calculates the expected values of the reference-to-reference and reference-to-residual cross-spectral-density (CSD) matrices. These matrices, together with an estimate of the plant transfer matrix, are used to solve for the control-filter coefficients. In order to account for possible ill-conditioning of the reference-to-reference CSD matrix, singular-valued-decomposition and thresholds are used to improve the conditioning of this matrix prior to solving the Wiener-filter equation for the control filter coefficients.—RBC

5,978,489

**43.50.Ki MULTI-ACTUATOR SYSTEM FOR ACTIVE SOUND AND VIBRATION CANCELLATION**

Eric Andrew Wan, assignor to Oregon Graduate Institute of Science and Technology  
2 November 1999 (Class 381/71.11); filed 4 May 1998

An alternative algorithm to the Widrow Filtered-X LMS algorithm for feedforward control is presented. The algorithm filters the error signals, as opposed to filtering the reference signals, to obtain the drive signals for the actuators. For multiple-input-multiple-output systems, the algorithm is claimed to provide a significant reduction in computational complexity. It is similar, if not the same, as the control algorithms described by Strothers and Elliott in United Kingdom Patent Applications 9222102.7 and 9405953.2.—RBC



5,987,144

#### 43.50.Ki PERSONAL ACTIVE NOISE CANCELLATION METHOD AND DEVICE HAVING INVARIANT IMPULSE RESPONSE

Christian Carme and Alain Roure, assignors to Technofirst; Centre National de la Recherche Scientifique  
16 November 1999 (Class 381/72); filed in France 4 April 1995

An active control system is described to reduce the local sound field around the head of a passenger sitting on an automobile or aircraft seat. The system uses speakers and residual microphones located within a cavity in the seat near the passenger's head. Remote reference sensors are filtered by a digital controller to generate drive signals to the speakers. A LMS-based algorithm is used to minimize the mean-square response of the residual microphones. The system supports plant identification as well as control functionality. The authors claim that the mechanical design of the seat cavity holding the speakers and microphones is such that the plant transfer functions are invariant relative to passenger head location.—RBC

5,991,663

#### 43.64.Me MULTIPLE PULSE STIMULATION

Laurence Irlicht and Graeme Clark, assignors to The University of Melbourne  
23 November 1999 (Class 607/57); filed 17 October 1995

A cochlear implant stimulation strategy seeks to change the time domain response of a patient's neural system for acoustic stimuli to an approximation of that of a normal-hearing person to the same stimuli. A table look-up is used after analysis of the electrical signals produced by the acoustic signals to identify portions of the signal corresponding to one of a group of predefined features. The resultant are stimuli to be presented via the electrode array.—DAP

5,991,417

#### 43.66.Ts PROCESS FOR CONTROLLING A PROGRAMMABLE OR PROGRAM-CONTROLLED HEARING AID FOR ITS *IN-SITU* FITTING ADJUSTMENT

Jan Topholm, assignor to Topholm & Westerman ApS  
23 November 1999 (Class 381/60); filed 2 May 1995

To prevent acoustic feedback oscillation in fitting persons with more severe hearing losses, an automatic adjustment is made on the maximum gain in at least one of several frequency bands via continuous monitoring *in situ* of ear canal SPL. If oscillation is still detected at the lowest gain levels, or if too much background noise is detected, the fitting processes stops automatically.—DAP

5,991,419

#### 43.66.Ts BILATERAL SIGNAL PROCESSING PROSTHESIS

Richard Brander, assignor to Beltone Electronics Corporation  
23 November 1999 (Class 381/312); filed 29 April 1997

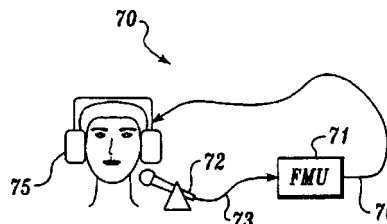
A binaural hearing aid fitting is described in which a miniature, low-current rf wireless link transceiver operating from a 1.1-V power source couples processed signals from one side of the head to the other. Algorithms that utilize differences between signals from both sides of the head enhance directionality, cancel noise from the sides of the wearer with adaptive or fixed processing, and spread out signal source locations perceptually. The system may either be asymmetric using one signal processor for the two outputs, or may use two signal processors for the two outputs.—DAP

5,940,798

#### 43.70.Dn FEEDBACK MODIFICATION FOR REDUCING STUTTERING

John F. Houde, assignor to Scientific Learning Corporation  
17 August 1999 (Class 704/271); filed 31 December 1997

It has long been known that auditory feedback of a stutterer's voice with a delay on the order of 50 ms will significantly reduce the stuttering. However, the improvements disappear when the feedback is discontinued. This system, based on a newer model of human speech control mechanisms,



adds a small amplitude perturbation to the feedback. This "tricks" the vocal tract controls into reducing an internal gain setting. The beneficial effects are said to persist after the auditory feedback is stopped.—DLR

5,946,649

#### 43.70.Dn ESOPHAGEAL SPEECH INJECTION NOISE DETECTION AND REJECTION

Hector Raul Javkin *et al.*, assignors to Technology Research Association of Medical Welfare Apparatus  
31 August 1999 (Class 704/203); filed 16 April 1997

United States Patent 5,890,111, reviewed previously, described a method for reducing disturbing, artificial noises generated by an esophageal speech device. This patent presents an alternative method for achieving the same result. A FFT analysis is done of the esophageal speech. A few rules testing timing and spectral features detect normal speech, injection noise, or silence.—DLR

5,930,747

#### 43.72.Ar PITCH EXTRACTION METHOD AND DEVICE UTILIZING AUTOCORRELATION OF A PLURALITY OF FREQUENCY BANDS

Kazuyuki Iijima *et al.*, assignors to Sony Corporation  
27 July 1999 (Class 704/207); filed in Japan 1 February 1996

This device is similar to the traditional autocorrelation pitch detector, except that there are two of them. One channel follows a high-pass filter, the other, a low-pass filter. Pitch evaluation steps following each autocorrelation determine the best pitch estimate based on the known error characteristics of each channel.—DLR

5,930,749

#### 43.72.Ar MONITORING, IDENTIFICATION, AND SELECTION OF AUDIO SIGNAL POLES WITH CHARACTERISTIC BEHAVIORS, FOR SEPARATION AND SYNTHESIS OF SIGNAL CONTRIBUTIONS

Stephane Herman Maes, assignor to International Business Machines Corporation  
27 July 1999 (Class 704/228); filed 2 February 1996

This audio signal classifier computes the transfer function poles from a linear prediction analysis of each frame. Based on a characterization of the pole movements, the frame is then analyzed into different audio compo-

nents, such as speech, music, a particular background signal, etc. Those poles contributing to the target signal are extracted, allowing the target signal to be cleaned up, substantially reducing the unwanted signals.—DLR

5,933,805

**43.72.Ar RETAINING PROSODY DURING SPEECH ANALYSIS FOR LATER PLAYBACK**

Dale Boss *et al.*, assignors to Intel Corporation  
3 August 1999 (Class 704/249); filed 13 December 1996

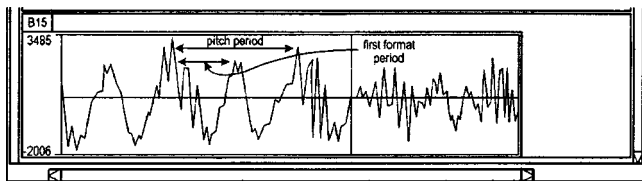
This is a version of what was known in the 1970s as a phoneme vocoder. A segment (phoneme) classification is performed using recognition technology. There is a brief mention, without elaboration, that recognized words, phrases, or sentences might be useful for error correction at the phonetic level. Each segment is transmitted with pitch, amplitude, and duration information from a prosodic analysis. At the receiver, standardized baseline values, such as average pitch or amplitudes, may be inserted. These are referred to as "voice fonts."—DLR

5,937,374

**43.72.Ar SYSTEM AND METHOD FOR IMPROVED PITCH ESTIMATION WHICH PERFORMS FIRST FORMANT ENERGY REMOVAL FOR A FRAME USING COEFFICIENTS FROM A PRIOR FRAME**

John G. Bartkowiak and Mark A. Ireton, assignors to Advanced Micro Devices, Incorporated  
10 August 1999 (Class 704/209); filed 15 May 1996

This linear prediction (LP) vocoder is mostly of traditional design, including the use of a switched pulse or noise excitation source. The novel element is a look-ahead pitch analysis. From the LP coefficients of the



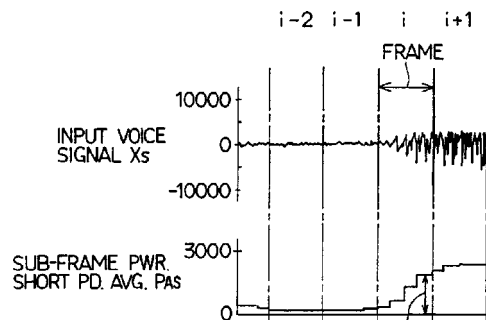
previous frame, a two-pole spectral model is extracted. This model is used to prefilter the next frame, improving the accuracy of the pitch estimate on that frame prior to the normal LP analysis.—DLR

5,937,375

**43.72.Ar VOICE-PRESENCE/ABSENCE DISCRIMINATOR HAVING HIGHLY RELIABLE LEAD PORTION DETECTION**

Kazuo Nakamura, assignor to Denso Corporation  
10 August 1999 (Class 704/215); filed in Japan 30 November 1995

This voice presence detector uses an improved method of detecting the beginning of an utterance. Each frame is divided into four subframes and a speech energy computation is performed in each subframe. This avoids a



problem of failure to detect a voice start when it occurs very late in a frame.—DLR

5,946,650

**43.72.Ar EFFICIENT PITCH ESTIMATION METHOD**

Ma Wei, assignor to Tritech Microelectronics, Limited  
31 August 1999 (Class 704/207); filed 19 June 1997

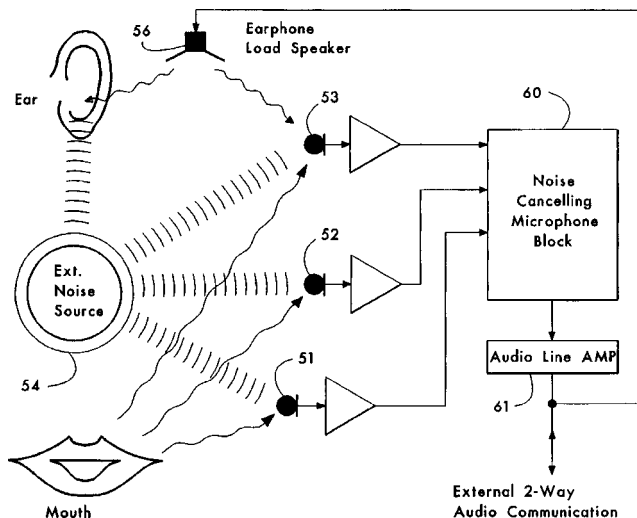
This speech pitch tracker begins simply enough with center clipping and low-pass filtering. These are followed by a series of error checking procedures, including auto- and cross-correlations in adjacent, variable-length windows, finding the minima of functions of those correlations and a final test for pitch doubling.—DLR

5,937,070

**43.72.Dv NOISE CANCELLING SYSTEMS**

Chris Todter, San Diego, California, *et al.*  
10 August 1999 (Class 381/71.6); filed in Australia 14 September 1990

This active noise cancelling system uses one or more additional microphones positioned so as to pick up the output of a loudspeaker in the communications loop. The loudspeaker cone resonance and other transfer

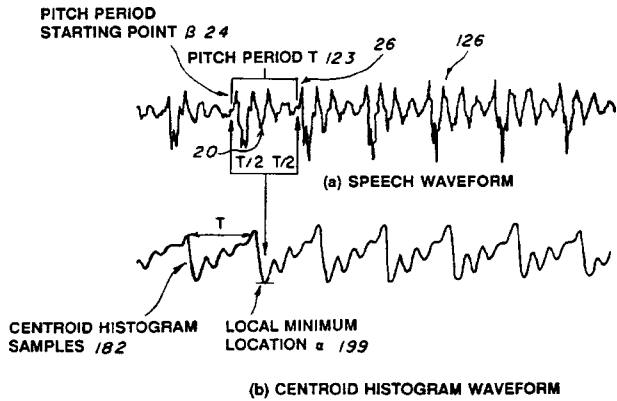


function characteristics are determined when the system is not in use for speech. An on-line, adaptive noise compensation computation is also described.—DLR

**43.72.Ew METHOD AND APPARATUS FOR GENERATING MODIFIED SPEECH FROM PITCH-SYNCHRONOUS SEGMENTED SPEECH WAVEFORMS**

George S. Kang and Lawrence J. Fransen, assignors to the United States of America as represented by the Secretary of the Navy 3 August 1999 (Class 704/278); filed 7 November 1995

The patent describes a method of identifying the pitch period of a speech signal for purposes of pitch, rate, or other alterations of the signal. A running histogram is computed from a low-pass filtered version of the input



signal. A local minimum of the histogram indicates a suitable cutting point according to a centroid definition.—DLR

**43.72.Fx SYSTEM FOR VOICE VERIFICATION OF TELEPHONE TRANSACTIONS**

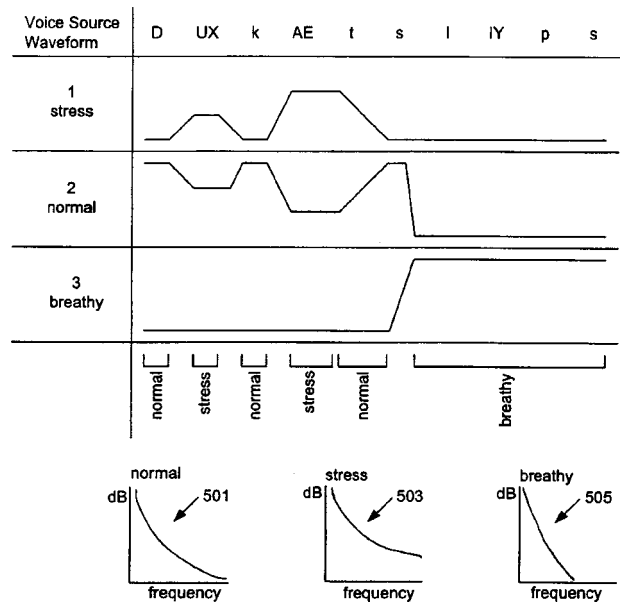
William Yee-Ming Huang *et al.*, assignors to ITT Defense, Incorporated 10 August 1999 (Class 704/247); filed 10 April 1996

This speaker verification system for telephone use is based on known speech recognition technology with improvements to speed up operation. An applicant is prompted to speak one of about 24 predetermined phrases. Speech recognition methods, including Viterbi decoding, are used to segment the input speech and to attach phonetic labels. Only segments with matching phonetic labels are compared in the speaker verification steps.—DLR

**43.72.Ja UTILIZATION OF A RECORDED SOUND SAMPLE AS A VOICE SOURCE IN A SPEECH SYNTHESIZER**

Mark L. Cecys, assignor to Apple Computer, Incorporated 27 July 1999 (Class 704/260); filed 11 March 1994

The patent makes the assertion that the complicating factors in the design of a voicing source for a formant synthesizer have been identified and mastered such that segments of real speech or various nonspeech sounds can be appropriately adapted for use as a high-quality voicing source. If true, this would provide new degrees of flexibility and output speech quality



from synthesizers, such as used in typical text-to-speech systems. The discussion seems to concentrate primarily on the spectral rolloff of the source. In addition to adjusting the rolloff, the device can switch or fade rapidly between sources.—DLR

**43.72.Ja CAR NAVIGATION SYSTEM**

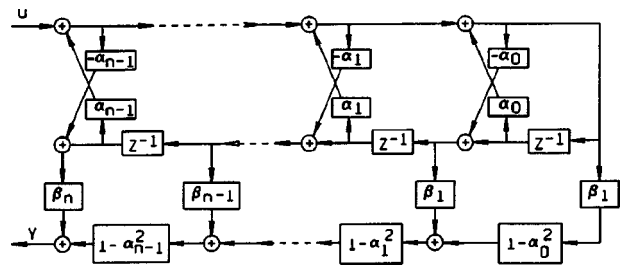
Masahiro Saiki, assignor to Matsushita Electric Industrial Company, Limited 10 August 1999 (Class 701/211); filed in Japan 7 September 1995

This automobile navigation system includes an elaborate audio system used for playback of the synthesized navigation data. Various voices are used to indicate different functions. Synthesized spatial patterns and amplitude variations for the voice source can be used to indicate the direction of referenced objects, oncoming traffic, etc.—DLR

**43.72.Ja METHOD AND APPARATUS FOR SPEECH ANALYSIS AND SYNTHESIS USING LATTICE LADDER NOTCH FILTERS**

Christopher I. Byrnes and Anders Lindquist, assignors to Washington University 17 August 1999 (Class 704/219); filed 9 May 1997

A speech analyzer is presented which recovers both poles and zeroes in the spectrum model. As this reviewer understands the presentation in thick patentese, the method consists of performing a standard all-pole analysis, inserting zeroes according to one or more simple procedures, and then



running an iterative spectral matching loop until a satisfactory pole-zero approximation has been reached. The final speech spectral filter is implemented in an augmented lattice configuration as shown in the figure.—DLR

5,940,797

**43.72.Ja SPEECH SYNTHESIS METHOD UTILIZING AUXILIARY INFORMATION, MEDIUM RECORDED THEREON THE METHOD AND APPARATUS UTILIZING THE METHOD**

Masanobu Abe, assignor to Nippon Telegraph and Telephone Corporation  
17 August 1999 (Class 704/260); filed in Japan 24 September 1996

The patent describes in some detail the prosody controls for a text-to-speech synthesizer, while only briefly mentioning a clearly inadequate phonetic segment synthesis strategy. Prosodic information is combined from a word dictionary lookup, extracts from recordings of human speech, and a fairly elaborate syntactic analysis. The primary application is the adaptation of prerecorded synthesis materials when changes are required but the original speaker is no longer available.—DLR

5,946,651

**43.72.Ja SPEECH SYNTHESIZER EMPLOYING POST-PROCESSING FOR ENHANCING THE QUALITY OF THE SYNTHESIZED SPEECH**

Kari Jarvinen and Tero Honkanen, assignors to Nokia Mobile Phones  
31 August 1999 (Class 704/223); filed 13 June 1996

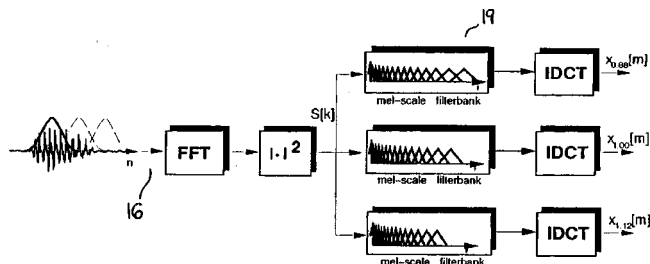
Disclosed is a method for improving the perceptual quality of the speech produced by a code-excited linear prediction vocoder receiver. The method is particularly suitable for a vocoder having a split excitation codebook, with a fixed part and an adaptive part. A correction or enhancement signal is derived from the adaptive codebook and is added to the overall original codebook output, either before or after spectral information is added by the short-term LPC filter. Said to have the greatest effect during voiced intervals, the enhancement seems to operate primarily by cleaning up the pitch periodicity.—DLR

5,930,753

**43.72.Ne COMBINED FREQUENCY WARPING AND SPECTRAL SHAPING IN HMM BASED SPEECH RECOGNITION**

Alexandros Potamianos and Richard Cameron Rose, assignors to AT&T Corporation  
27 July 1999 (Class 704/256); filed 20 March 1997

One of the primary sources of difference in an individual's speech as compared to an ensemble of speakers is a shift in the frequencies of corresponding formant patterns. However, accounting for the frequency shifts alone is not sufficient. This patent covers a method of combining frequency



shifts with linear spectral amplitude changes to improve recognition across speakers. Frequency shifting is done in the cepstral coefficient space, while spectral warping is done by adapting the HMM mixture Gaussian parameters.—DLR

5,940,794

**43.72.Ne BOUNDARY ESTIMATION METHOD OF SPEECH RECOGNITION AND SPEECH RECOGNITION APPARATUS**

Yoshiharu Abe, assignor to Mitsubishi Denki Kabushiki Kaisha  
17 August 1999 (Class 704/253); filed in Japan 2 October 1992

What this patent refers to as a boundary estimation problem is what is more commonly called phonetic segmentation. Utterance boundary detection is not the issue. What is done is a preliminary mel-scaled cepstral analysis of each speech frame. Based on stability of the coefficients across a small number of nearby frames, the frame is assigned a probability of representing a phonetic transition. The resulting transition probability assignments are then used to improve the performance of the following segment classifier.—DLR

5,937,384

**43.72.Ne METHOD AND SYSTEM FOR SPEECH RECOGNITION USING CONTINUOUS DENSITY HIDDEN MARKOV MODELS**

Xuedong D. Huang and Milind V. Mahajan, assignors to Microsoft Corporation  
10 August 1999 (Class 704/256); filed 1 May 1996

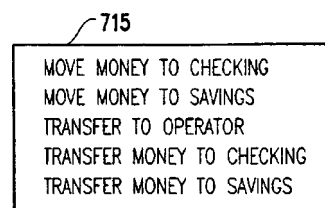
The patent describes a method for improving the match between an input speech signal and a hidden Markov model (HMM) sequence in a speech recognizer, especially in the case where sparse training data has limited the definition of some of the HMM density functions. The available training data is used to train context-independent HMMs for each phonetic unit as well as the set of available context-dependent HMMs. An estimate is then generated to indicate the probability that an undefined context would be approximated by the available training information.—DLR

5,937,385

**43.72.Ne METHOD AND APPARATUS FOR CREATING SPEECH RECOGNITION GRAMMARS CONSTRAINED BY COUNTER EXAMPLES**

Wlodek Wlodzimierz Zadrozny and Nandakishore Kambhatla, assignors to International Business Machines Corporation  
10 August 1999 (Class 704/257); filed 20 October 1997

Most speech recognition patents deal with problems in the phonetic matching area. This one concerns a less commonly seen area, that of grammar management, or more specifically, automatic grammar management. Based on a set of sample phrases, a trial grammar is proposed to cover the set of utterances allowable at a certain point in the dialogue. The full set of



possible phrases is then generated using that grammar and a machine or human judge identifies the nonapplicable counter examples. These are fed back into a grammar revision step, which produces an updated grammar.—DLR

5,937,383

**43.72.Ne APPARATUS AND METHODS FOR SPEECH RECOGNITION INCLUDING INDIVIDUAL OR SPEAKER CLASS DEPENDENT DECODING HISTORY CACHES FOR FAST WORD ACCEPTANCE OR REJECTION**

**Abraham Poovakunnel Ittycheriah and Stephane Herman Maes, assignors to International Business Machines Corporation**  
10 August 1999 (Class 704/255); filed 2 February 1996

This strategy for rapid keyword decoding is intended to speed up processing in a speech recognizer, especially in a large, multi-user system where one server handles voice inputs from many users. A small cache of recently recognized words is checked first, before resorting to checking against the full vocabulary. The score of the new input against the cache items is compared to the original scores when those cached items were first recognized.—DLR

5,943,647

**43.72.Ne SPEECH RECOGNITION BASED ON HMMS**

**Jari Ranta, assignor to Tecnomen Oy**  
24 August 1999 (Class 704/251); filed in Finland 30 May 1994

This speech recognizer uses dual-model hidden Markov models which have one model in each state for the usual first-order features and a second model in each state for the time derivatives of each of the features.—DLR

5,946,653

**43.72.Ne SPEAKER INDEPENDENT SPEECH RECOGNITION SYSTEM AND METHOD**

**William Michael Campbell *et al.*, assignors to Motorola, Incorporated**  
31 August 1999 (Class 704/243); filed 1 October 1997

A matrix processing method of speaker-independent speech recognition is presented. A set of cepstral coefficients and their derivatives are expanded by a high-order polynomial expansion. For example, a fourth-order expansion of 12 cepstral and 12 derivative coefficients results in a vector of 20 475 elements for each speech frame. These vectors are operated on by various summation, product, and normalizing steps, resulting in a final dot product which indicates the recognized speech phrase.—DLR

5,930,748

**43.72.Pf SPEAKER IDENTIFICATION SYSTEM AND METHOD**

**John Eric Kleider and Khaled Assaleh, assignors to Motorola, Incorporated**  
27 July 1999 (Class 704/219); filed 11 July 1997

The patent describes a method of adapting a speaker identification model called Row Action Projection. In this method, a cepstral vector is expanded by all product terms to some polynomial order. For example, a second-order expansion of 12 cepstral coefficients would have 91 terms. The expansions of all frames for a large number of speakers are processed to compute error terms by which initially randomized 91-element speaker model vectors are updated. Interestingly, the entire patent text contains no math formulas. Instead, all vector and matrix operations are described in fairly clearly written text.—DLR

# Experimental validations of the HELS method for reconstructing acoustic radiation from a complex vibrating structure

Nassif Rayess and Sean F. Wu

*Department of Mechanical Engineering, Wayne State University, Detroit, Michigan 48202*

(Received 31 July 1999; revised 7 August 1999; accepted 9 February 2000)

This paper presents experimental validations of the Helmholtz Equation Least Squares (HELs) method [Wang and Wu, *J. Acoust. Soc. Am.* **102**, 2020–2032 (1997); Wu and Wang, U.S. Patent Number 5712805 (1998); Wu, *J. Acoust. Soc. Am.* **107**, 2511–2522 (2000)] on reconstruction of the radiated acoustic pressures from a complex vibrating structure. The structure under consideration has geometry and dimensions similar to those of a real passenger vehicle front end. To simulate noise radiation from a vehicle, a high fidelity loudspeaker installed inside the structure at the location of the engine is employed to generate both random and harmonic acoustic excitations. The radiated acoustic pressures are measured over a finite planar surface above the structure by a microphone. The measured data are taken as input to the HELs formulation to reconstruct the acoustic pressures on the top surface of the structure as well as in the field. The reconstructed acoustic pressures are then compared with measured ones at the same locations. Also shown are comparisons of the reconstructed and measured acoustic pressure spectra at various locations on the surface. Results show that satisfactory reconstruction can be obtained on the top surface of the structure subject to both random and harmonic excitations. Moreover, the more measurements and the closer their distances to the source surface, the more accurate the reconstruction. The efficiency of the HELs method may decrease with increasing of the excitation frequency. This high frequency difficulty is inherent in all expansion theories. © 2000 Acoustical Society of America. [S0001-4966(00)03005-8]

PACS numbers: 43.20.Tb, 43.20.Rz, 43.40.Rj, 43.50.Yw [CBB]

## INTRODUCTION

Since its first publication, the Helmholtz Equation Least Squares (HELs) method<sup>1,2</sup> has been used to reconstruct the radiated acoustic pressure fields from complex vibrating structures in both exterior<sup>3,4</sup> and interior<sup>5</sup> regions. In these studies, general three-dimensional structures are considered that contain sharp edges and corners and abrupt changes on surface contours. Acoustic radiation is simulated by applying harmonic forces to arbitrarily selected surfaces on the structures under consideration and determined by using the finite element method (FEM) and boundary element method (BEM). The radiated acoustic pressures are collected in the field and taken as input to the HELs formulation, and the reconstructed acoustic pressures on the source surface and in the field are then compared with the benchmark values.

These investigations<sup>1–5</sup> show that the HELs method enables one to reconstruct the radiated acoustic pressure fields everywhere with a few expansion functions, and is relatively insensitive to measurement locations. Reconstruction can be done by taking field pressures over either conformal or planar surfaces that enclose the vibrating structure. Results indicate that the accuracy of reconstruction is higher for a conformal measurement surface than a planar one.<sup>6</sup> This is because a planar surface often extends beyond the near-field region and makes the input data less accurate. However, measurements over a conformal surface require a properly programmed robot arm which may not be easily realized in practice, while those over planar surfaces can be facilitated

by a simple  $x$ - $y$  translator. Hence a preferable way is to take measurements over small planar surfaces that enclose the source at the closest possible distances.

The present paper demonstrates experimental validations of the HELs method on a complex structure with dimensions and geometry similar to those of a real vehicle front end. To simulate noise radiation from a passenger vehicle, white noise is generated by a high fidelity loudspeaker installed inside the structure at the same location as the engine. The radiated acoustic pressures are measured by a condenser microphone sweeping over a planar surface at certain distance away from the structure. These measured signals are taken as input to the HELs formulation to reconstruct the acoustic pressure distribution on the structure surface, and the values thus obtained are compared with those measured at the same locations.

The ultimate goal of the present investigation is to develop an effective tool for the design engineers to diagnose noise sources and assess noise performance of a complex vibrating structure such as a passenger vehicle, aircraft cabin, machine tool, etc. Such a diagnosis is important for it may not only yield insight into better modifications of structural designs to lower noise radiation, but also provide a means to visualize acoustic radiation paths, thus enabling the engineers to tackle noise problems more cost-effectively.

B&K<sup>®</sup> has developed the State Transformation of Sound Field (STSF) method,<sup>7</sup> which combines the Near-field Acoustic Holography (NAH)<sup>8–10</sup> for reconstructing the acoustic near field and BEM-based Helmholtz integral for-

mulation for the acoustic far field. This method has also been adopted in the LMS<sup>®</sup> Acoustic Holography software<sup>11</sup> to produce images of noise source locations. However, STSF has a major defect in that it assumes the acoustic pressure field radiated from any source to be equivalent to that from a finite planar surface. This assumption is mathematically indefensible, which renders STSF suitable for a planar surface within an extremely close distance to the source. Even so, visualization of acoustic pressure is limited to the size of the measurement plane on one side. For a nonplanar source, this assumption is wrong and the resulting images may be false, regardless of projecting backward or forward. Moreover, reconstruction of the acoustic pressure field is done on a spatial sampling basis. Consequently, its applicability may be limited and become increasingly inefficient at higher frequencies.

Theoretically, one may replace the acoustic pressure field radiated from an arbitrary source by that from a control surface which completely encloses the source. If the acoustic pressure distribution on this enclosure is specified, say, through microphone measurements or numerical calculations, then the radiated acoustic pressures anywhere, including the source surface can be reconstructed by using a BEM-based Helmholtz integral formulation. Examples of such have been shown in both exterior<sup>12–14</sup> and interior<sup>15,16</sup> regions.

However, this integral approach has many shortcomings as pointed out in Ref 1, the major one being due to the fact that BEM can only depict acoustic pressures at the certain nodes discretized on the surface. To avoid aliasing and acquire certain resolution in reconstruction, one must have a minimum number of nodes per wavelength on the surface and take the same number of measurements in the field. Then by solving a matrix equation the pressure amplitudes at the nodes on the surface can be determined. Since the number of discrete nodes increases with the excitation frequency and complexity of the source geometry, this technique may become impractical for a finite-sized object, for example, an engine or a vehicle front end in the low-to-mid frequency regime. This is because the corresponding number of measurements needed to reconstruct the acoustic pressure field would be too large to make this process cost-effective.

Unlike STSF and BEM-based techniques, the HELS method employs an expansion of spheroidal functions to reconstruct the acoustic pressure field. The number of expansion terms determines the number of measurements, which is small when the source geometry is spherical or its aspect ratio is close to unity,  $x:y:z \rightarrow 1:1:1$ . For example, for sources radiating primarily combinations of monopole and dipole sounds, reconstruction can be accomplished by using four expansion terms or taking four measurements in the field, regardless of the excitation frequency. For general sound radiation, the number of expansion terms may be large. However, the high-order terms represent the small-scaled effects, which do not contribute to the far field and which may have been lost in the measurements and cannot be recovered anyway. Therefore these high-order terms can be neglected and the total number of measurements remains relatively small.

Theoretically, the HELS formulation is exact for a spheroidal surface. However, in engineering applications most vibrating structures are of arbitrary shapes. Section IV shows that for a nonspheroidal but smooth surface, satisfactory reconstruction of the radiated acoustic pressure fields can be obtained by using the HELS method. This is not surprising because one can always approximate the acoustic pressure field radiated from a nonspheroidal surface by using an expansion of spheroidal functions.<sup>17</sup> For example, one can use ellipsoidal functions to describe the acoustic pressure fields generated by chunky objects of aspect ratios  $x:y:z \rightarrow 1:1:1$ , elongated surfaces of  $x:y:z \rightarrow 1:1:10$ , flat surfaces of  $x:y:z \rightarrow 1:10:10$ , or both elongated and flat surfaces of aspect ratios  $x:y:z \rightarrow 1:10:100$ . Among the coordinate systems, the spherical coordinates are often selected because the corresponding wave functions are readily available in the mathematical libraries such as IMSL subroutines. On the other hand, analytic expressions for spheroidal functions in oblate, prolate, and elliptic coordinates are not available and the corresponding numerical computations can be extremely time-consuming. In this paper, we use spherical coordinates to reconstruct acoustic pressures radiated from a nonspherical surface.

The main objective of the present paper is to validate the HELS method experimentally and to examine its effectiveness and robustness under a simulated working environment. In engineering applications, one may often encounter practical difficulties in carrying out noise diagnostics. For example, in some cases it may not be possible to take measurements at very close distances and the near-field information may be lost in the input data, while reconstruction must be done in the region where the near-field effect is predominant. Sometimes, it may not even be possible to take a sufficient number of measurements. Needless to say, these restrictions will affect the accuracy of reconstruction. Such adverse scenarios are embedded in the present investigation, and the results thus obtained are demonstrated and discussed.

## I. THEORY

The HELS method assumes that the radiated acoustic pressure from a vibrating object in an unbounded fluid medium is expressible in terms of an expansion of certain basis functions,<sup>3,18</sup>

$$\hat{p}(\mathbf{x}, \omega) = \rho c \sum_{j=1}^J C_j \Psi_j(\mathbf{x}, \omega), \quad (1)$$

where  $\hat{p}(\mathbf{x}, \omega)$  is the complex amplitude of the acoustic pressure at any field point  $\mathbf{x}$  and angular frequency  $\omega$ ,  $\rho$  and  $c$  are the density and speed of sound of the fluid medium, respectively,  $\Psi_j$  are the basis functions obtained by the Gram–Schmidt orthonormalization on the source surface  $\partial B$  with respect to the particular solutions  $\psi_i$  to the Helmholtz equation,

$$\langle \psi_i(\mathbf{x}_B, \omega), \Psi_j(\mathbf{x}_B, \omega) \rangle = \int_{\partial B} \psi_i(\mathbf{x}_B, \omega) \Psi_j(\mathbf{x}_B, \omega) dS, \quad (2)$$

$$\mathbf{x}_B \in \partial B,$$

TABLE I. Relationship among indices  $i, j, n$ , and  $m$ .

$i, j$	$n$	$m$	$\cos(m\phi), \sin(m\phi)$
1	0	0	1
2	1	0	1
3	1	-1	$\sin \phi$
4	1	1	$\cos \phi$
5	2	0	1
6	2	-1	$\sin \phi$
7	2	1	$\cos \phi$
8	2	-2	$\sin(2\phi)$
9	2	2	$\cos(2\phi)$

where  $\psi_i$  can be expressed in any spheroidal coordinates.<sup>19</sup> In particular, one can write  $\psi_i$  in spherical coordinates as<sup>20</sup>

$$\psi_i(\mathbf{x}, \omega) = \psi_{n,m}(r, \theta, \phi) = h_n(kr)P_{n,m}(\cos \theta) \times \begin{cases} \cos(m\phi) \\ \sin(m\phi) \end{cases}, \quad (3)$$

where  $h_n(kr)$  and  $P_{n,m}(\cos \theta)$  denote the spherical Hankel functions and Legendre functions, respectively, and  $k = \omega/c$  is the acoustic wave number. Table I shows the relationships among the indices ( $i, j, m, n$ ) for the first nine expansion functions.

The coefficients  $C_j$  in Eq. (1) are determined by requiring the assumed-form solution to satisfy the boundary condition  $\hat{p}(\mathbf{x}_m, \omega)$  at the measurement point  $\mathbf{x}_m$

$$\rho c \sum_{j=0}^J C_j \Psi_{m,j}(\mathbf{x}_m, \omega) = \hat{p}(\mathbf{x}_m, \omega). \quad (4)$$

If a  $J$ -term expansion is used in Eq. (4), then  $M$  measurements must be taken ( $M \geq J$ ). To enhance the accuracy of reconstruction, the least-squares method is utilized to eliminate the first-order errors incurred during this process. Accordingly, we obtain the following equation<sup>3</sup>

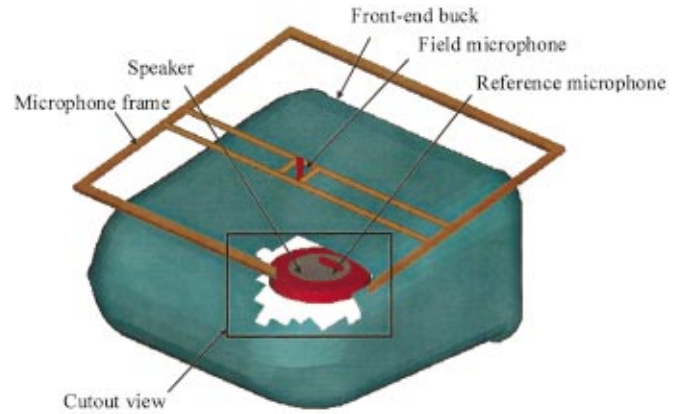


FIG. 2. Schematic of the test setup.

$$[T]_{J \times J} \{C\}_{J \times 1} = \{D\}_{J \times 1}, \quad (5)$$

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

$$T_{i,j} = \rho c \sum_{m=1}^M \Psi_{m,i}(\mathbf{x}_m, \omega) \Psi_{m,j}(\mathbf{x}_m, \omega), \quad (6a)$$

$$D_i = \sum_{m=1}^M \hat{p}(\mathbf{x}_m, \omega) \Psi_{m,i}(\mathbf{x}_m, \omega). \quad (6b)$$

## II. ALGORITHM

Figure 1 shows a flow chart of the computer model thus developed. The input data include the geometry and dimension of a structure and measured acoustic pressure signals. The former is needed to generate the expansion functions via an orthonormalization process, while the latter is used to

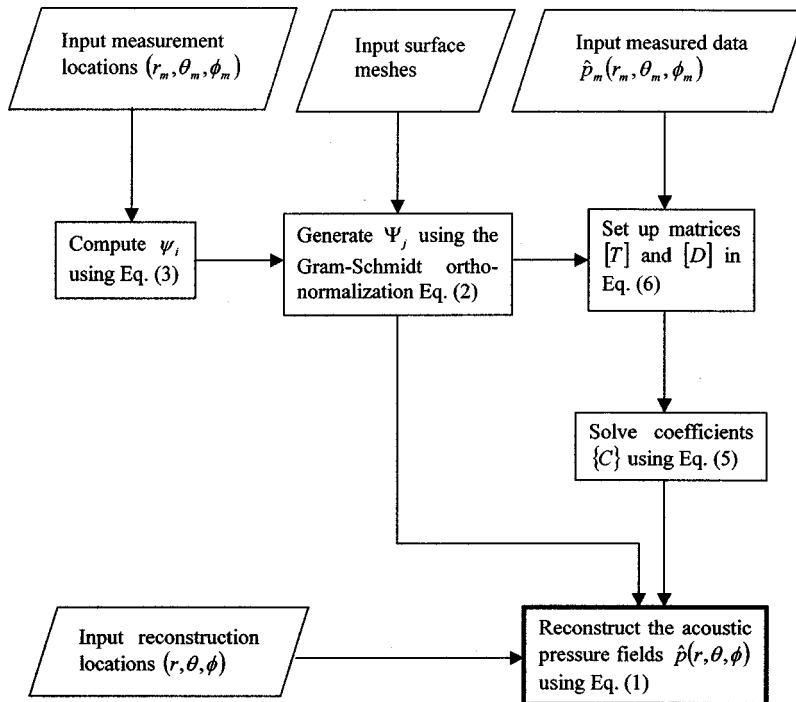


FIG. 1. Flow chart of the computer program for the HELS method.



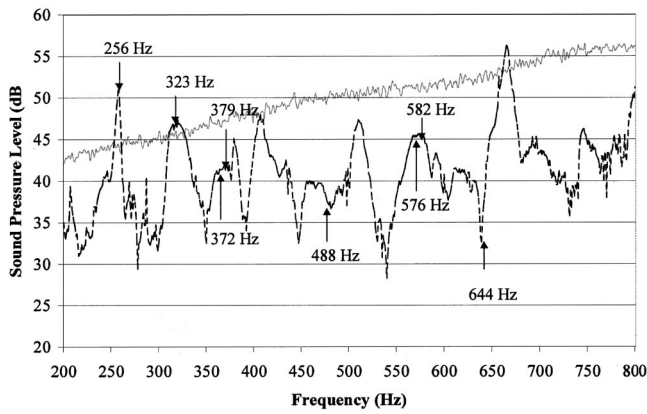


FIG. 3. Effect of the vehicle buck on the resulting acoustic pressure field. Solid line: Without buck; Dashed line: With buck.

determine the associated coefficients. Once these expansion functions and coefficients are specified, the acoustic pressure anywhere can be approximated by Eq. (1).

The acoustic pressure signals have two components: amplitudes and phases. The former can be obtained by taking auto-correlation of the acoustic pressure signals measured at any field point, while the latter can be determined by taking

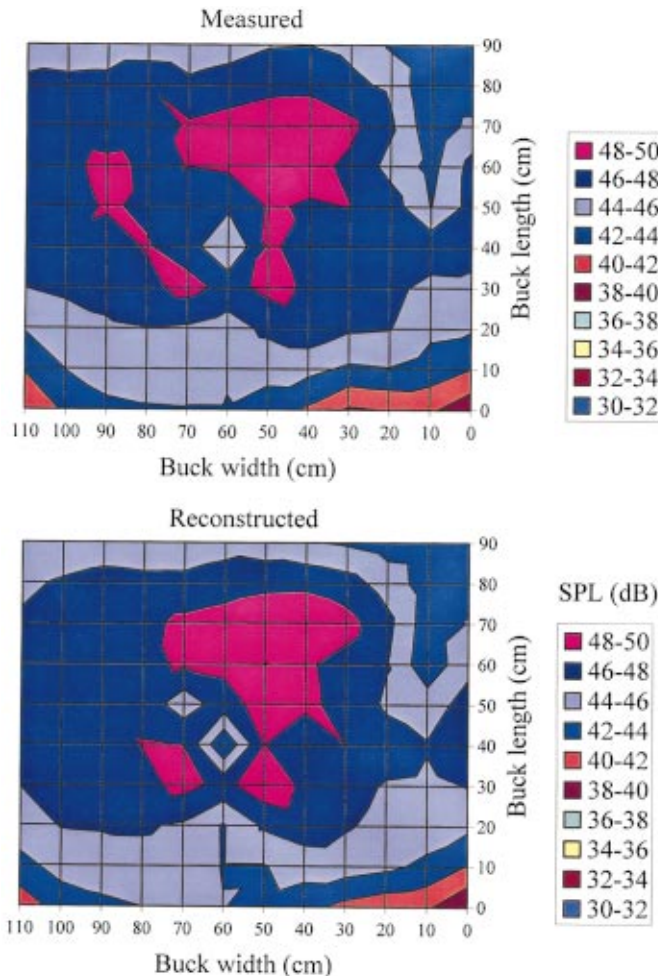


FIG. 4. Comparison of acoustic pressure distributions on the buck top surface subject to random excitation at 256 Hz.

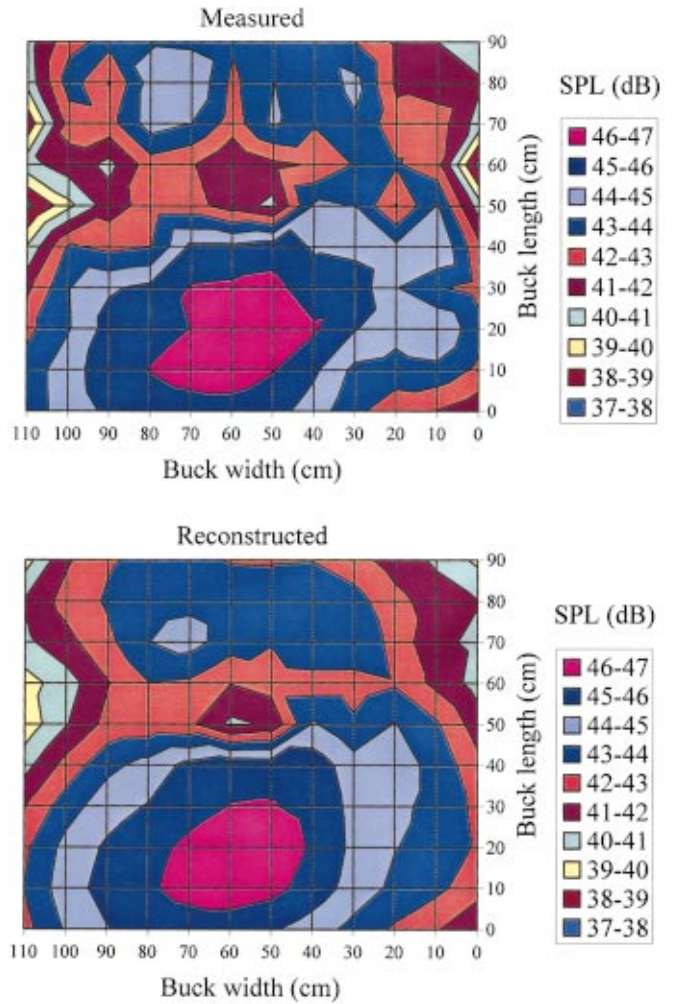


FIG. 5. Comparison of acoustic pressure distributions on the buck top surface subject to random excitation at 323 Hz.

cross-correlation of the signals measured at the source and field points, respectively.

Note that the HELS method imposes no restrictions on the measurement locations, so long as they do not overlap each other. For example, the radiated acoustic pressures can be measured over either conformal or planar surfaces. Test results have shown that measurements over a conformal surface always yield more accurate reconstruction than those over a planar surface.<sup>6</sup> This is because a conformal surface ensures that all measurements are taken in the same region, thus guaranteeing a consistent accuracy in the input data. On the other hand, a planar surface allows measurements to extend beyond the near field, thus making the measured signals less consistent and prone to errors. However, measurements over a conformal surface require a properly programmed robot arm that may not be easily realized in engineering applications, while those over a planar surface can be facilitated by a simple  $x$ - $y$  translator. Therefore, a preferred way is to measure acoustic pressures over a control surface that consists of several small planar surfaces enclosing completely the source at the closest possible distances.

Since the expansion functions  $\Psi_j$  are frequency dependent, the associated coefficients  $C_j$  must be solved repeatedly for each frequency. This process is accomplished by a

do-loop in the computer program that covers the entire frequency range of interest. Because of the simplicity of the mathematics involved, the numerical computations are very fast. For example, computations for solving the coefficients  $C_j$  associated with a  $J$ -term expansion ( $J < 100$ ) over an 800-line spectrum can be completed on an IBM Computer with a Pentium processor III 500 MHz within a few minutes.

### III. TEST SETUP

In this paper, noise radiation from an engine installed inside a vehicle front end was simulated. Since a real vehicle front end was not available, a replica was built. Figure 2 depicts such a structure modeled from a real, full-size vehicle front end, also known as a front-end buck. This buck had similar shape and sizes as a real one, but with the windshield, tires, and rear-view mirrors removed in order to fit inside a 12 in. by 12 in. by 6.5 in. fully anechoic chamber in the Acoustics, Vibration, and Noise Control (AVNC) Laboratory at Wayne State University.

The replicated buck was made of a skeleton of 6-mm steel rods machined and welded together in the shape of a vehicle front end. Attached to this skeleton was a layer of thick wire mesh, followed by a layer of papier-mâché, and finally by another layer of joint compound to yield similar hardness and rigidity as those of sheet metal in a passenger vehicle. Care was taken to provide ample space inside the buck and to leave the undercarriage area unobstructed. This buck was then screwed on wood posts in the anechoic chamber at a height consistent with that of a real vehicle. A high-fidelity loudspeaker (Electro-Voice, Inc.) was installed inside the buck at the same location as that of an engine. The speaker was connected to the B&K Multi-channel Signal Analyzer Type 3550. Signals were generated by the HP 8904A Multifunction Synthesizer DC-600 kHz and amplified by the B&K Power Amplifier Type 2712. Both random and harmonic signals were generated and the radiated acoustic pressures were measured by the B&K Prepolarized Free-field Microphone Type 4188.

Figure 2 also shows the schematic of the test setup. Here measurements were taken over two planar surfaces above the ‘hood’: one at 12 cm (called measurement surface) and the other at 1 cm (called reconstruction surface). The measurement surface was 110 cm long and 90 cm wide and contained 120 measurements with an interval of 10 cm between the neighboring points. The reconstruction surface had the same dimensions as those of the buck top surface on which 120 measurements were made for the validation purpose.

Note that while the HELS method could produce images of acoustic pressures on the entire buck surface under the present test setup, the accuracy of reconstruction on other surfaces may not be as accurate as that on the top surface.<sup>3,5</sup> This is because the input data based on a finite planar surface are incomplete, thus they should not be used to reconstruct the whole acoustic pressure field. To obtain a satisfactory reconstruction of the acoustic pressures over the entire surface, measurements must be taken over a control surface that completely encloses the source.<sup>6</sup> Therefore, we choose to focus our validation on the top surface of the buck only.

As pointed out above, taking more measurements at closer distances around the source can improve the accuracy of reconstruction. In this investigation, measurements were taken at certain distances away from the source and measurement area barely covered the top surface of the buck. Consequently, some near-field information may be lost in the measurements. These data were used to reconstruct the acoustic pressures on the buck surface, which was dominated by the near-field effects. Therefore it would be unrealistic to expect a perfect reconstruction because the problem was ill-posed. Such an ill-posedness difficulty exists in most inverse problems.

The presence of the buck structure further complicated the acoustic pressure field generated by the speaker. Figure 3 illustrates the comparison of the noise spectra with and without the structure. Results show that the buck structure acted like a series of band-pass filters. Without the buck, the spectrum was essentially a straight line, indicating a white noise as expected. With the buck, however, the sound pressure levels were greatly reduced except at certain frequency bands that coincided with the structural resonance frequencies, at which the transmission losses were minimum. Therefore, the presence of the buck and its interaction with sound significantly changed acoustic radiation from the loudspeaker, which in turn made reconstruction more difficult.

Because of a lack of a microphone array and the corresponding multi-channel signal analyzer, data acquisition was done by using two microphones, one fixed in the proximity to the speaker and the other in the field. Accordingly, measurements were taken one point at a time and the data were recorded in a desktop computer. At the end of each recording cycle, the field microphone was moved to a new location manually. This required entering the anechoic chamber, sliding the microphone along the rail, repositioning the rail, and exiting and closing the chamber door. These actions would inevitably disturb the setup and make the input data susceptible to errors, no matter how careful one might be. Pains-taking efforts were made to maintain repeatability of the measurements. Even so, there were random fluctuations in the signal levels that would make the input inconsistent over a long period of time. To circumvent this difficulty, transfer functions rather than the acoustic pressures were measured

$$G(\mathbf{x}_m | \mathbf{x}_{\text{ref}}; \omega) = \frac{\hat{p}_{\text{field},m}(\mathbf{x}_m, \omega)}{\hat{p}_{\text{ref}}(\mathbf{x}_{\text{ref}}, \omega)}, \quad (7)$$

where  $G(\mathbf{x}_m | \mathbf{x}_{\text{ref}}; \omega)$  represents the transfer function between the acoustic pressure at any measurement location  $\hat{p}_{\text{field},m}(\mathbf{x}_m, \omega)$  and that at the reference  $\hat{p}_{\text{ref}}(\mathbf{x}_{\text{ref}}, \omega)$ . Once  $G(\mathbf{x}_m | \mathbf{x}_{\text{ref}}; \omega)$  were collected from all measurement points, the acoustic pressures were obtained by multiplying the transfer functions by the acoustic pressure at the reference, which in effect was equivalent to using a microphone array to measure all field acoustic pressures simultaneously.

### IV. RESULTS AND DISCUSSIONS

In this investigation, two sets of data were collected. In the first set the buck was excited by random signals. The resulting acoustic pressures collected on the measurement

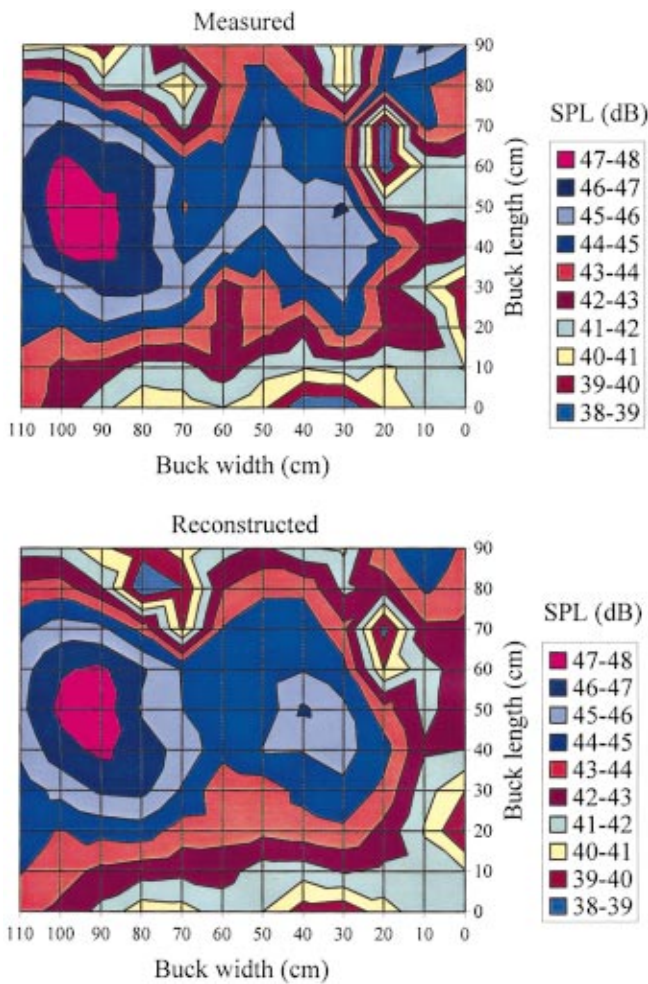


FIG. 6. Comparison of acoustic pressure distributions on the buck top surface subject to random excitation at 379 Hz.

surface were used to calculate acoustic pressures on the reconstruction surface. These reconstructed acoustic pressures were then compared with the measured values at the same locations. In the second set the buck was excited at 372, 488, 576, and 644 Hz simultaneously, which corresponded to a small coupled resonance, nonresonance, large coupled resonance, and anti-resonance frequencies, respectively (see Fig. 3). These experiments allowed us to test the effectiveness of the HELS method in reconstructing the acoustic pressure fields under both random and harmonic excitations.

Because of the errors involved in the measured data due either to the loss of the near-field effect or to measurement uncertainties, the matrix equation (5) was ill-posed. Under this condition, increasing the number of expansion terms may not necessarily enhance the accuracy of reconstruction. In fact, it may even worsen the accuracy since any addition of expansion terms may introduce new errors to Eq. (5). Results indicated that the accuracy of reconstruction on the source surface increased with the number of expansion terms up to a certain value, then decreased monotonically thereafter. On the other hand, the accuracy of reconstruction on the measurement surface increased monotonically with the number of expansion terms. This is because the errors on the measurement surface have been minimized by the least-squares method. Hence there exists an optimal number of

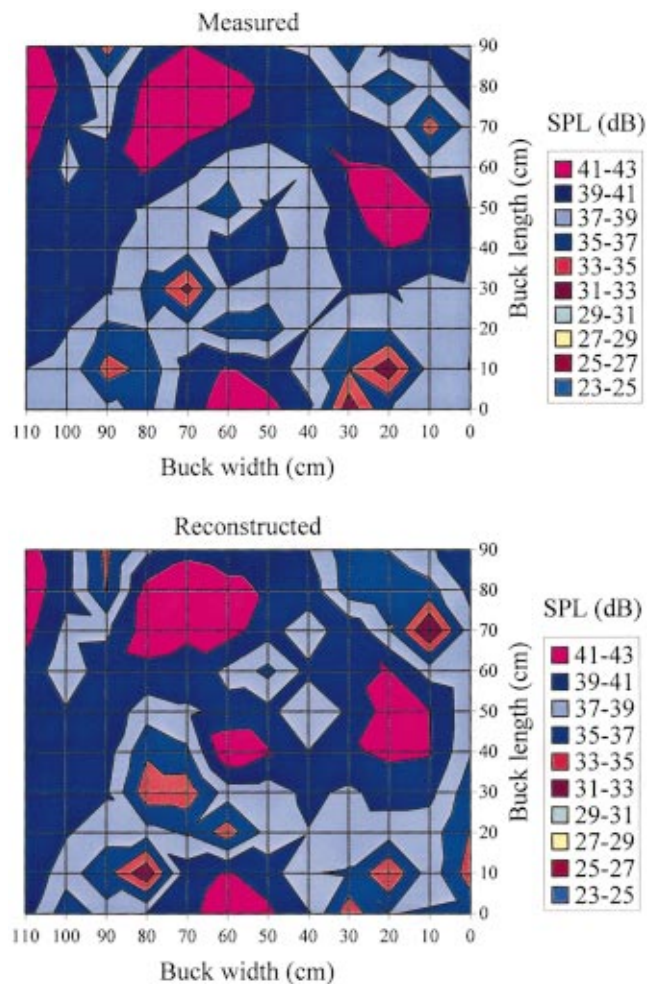


FIG. 7. Comparison of acoustic pressure distributions on the buck top surface subject to random excitation at 586 Hz.

expansion functions which may yield the best reconstruction with a minimal number of expansion.

Since the exact values of the acoustic pressure distribution on the source surface are unknown a priori, there is no way of determining the optimal expansion numbers analytically. However, for the HELS method to become a viable noise diagnostic tool, it is imperative to develop certain pro-

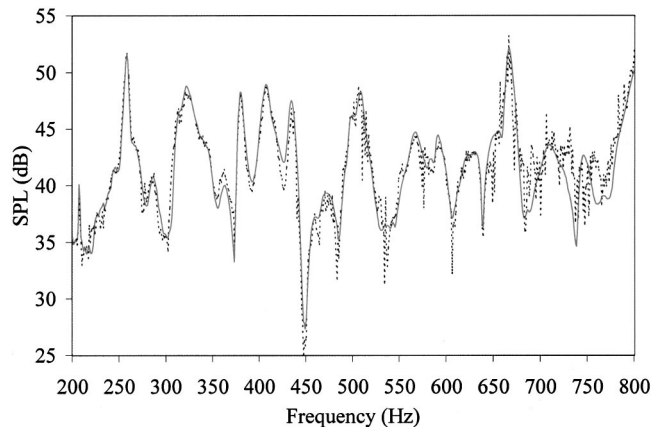


FIG. 8. Comparison of acoustic pressure spectra at a point halfway between the center and the right edge on the buck top surface. Solid line: Reconstructed; Dashed line: Measured.

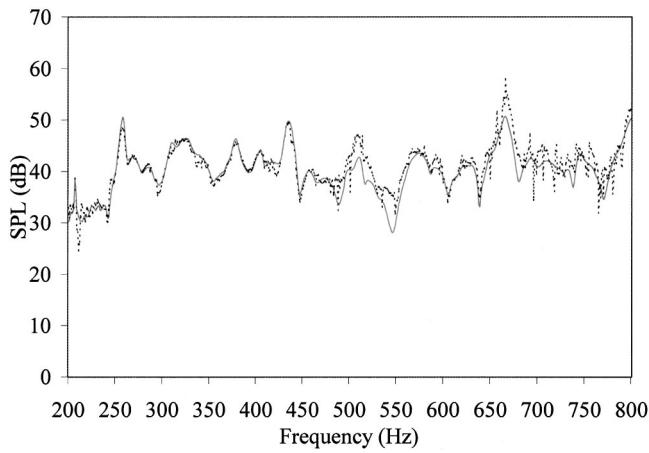


FIG. 9. Comparison of acoustic pressure spectra in the middle of the rear edge on the buck top surface. Solid line: Reconstructed; Dashed line: Measured.

cedures that may lead to optimal expansion numbers for the given test setup and measurements.<sup>3</sup> Such procedures are described as follows.

First, we set  $n=1$ , which corresponds to  $J=(n+1)^2=4$  (see Table I), solve a  $4 \times 4$  matrix equation (5) to determine the corresponding coefficients  $C_j$ , use them to reconstruct acoustic pressures on the measurement surface, and calculate the  $\|L\|^2$ -norm of the errors with respect to the measured values at the same locations. Next, we increase the expansion number by one,  $J=J+1$ , and repeat the above steps to get the new  $\|L\|^2$ -norm of the errors. Because of the measurement uncertainties, the  $\|L\|^2$ -norm of the errors may fluctuate with respect to the expansion number  $J$ . Hence curve-fitting is employed to smooth this error curve.

Since the accuracy of reconstruction on the measurement surface increases monotonically with the expansion number, the slope of this smoothed  $\|L\|^2$ -norm curve approaches zero beyond certain value of  $J$ . However, the optimal expansion for reconstruction on the source surface always occurs at a non-zero slope of this curve, i.e.,  $\tan \alpha > \varepsilon_0$ . Hence we stop the iteration at a pre-selected value of  $\varepsilon_0$  and take the corresponding value of  $J_0$  as being optimized with respect to the given input.

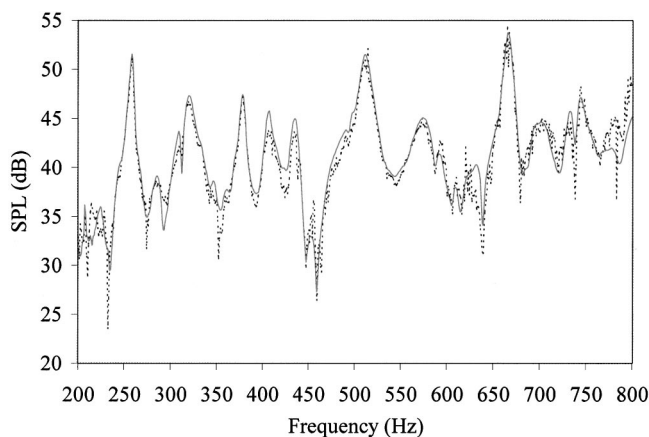


FIG. 10. Comparison of acoustic pressure spectra at the rear-left corner on the buck top surface. Solid line: Reconstructed; Dashed line: Measured.

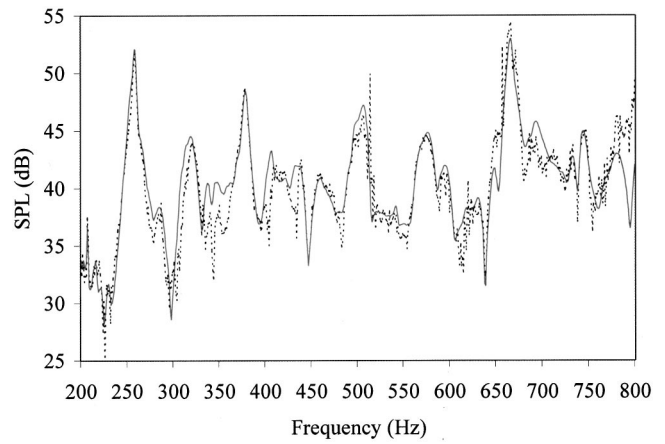


FIG. 11. Comparison of acoustic pressure spectra at the front-right corner on the buck top surface. Solid line: Reconstructed; Dashed line: Measured.

Obviously, this optimal expansion number  $J_0$  varies with the number of measurements and the excitation frequency. In general, the higher the frequency, the larger the value of  $J_0$ . Consequently, the efficiency of numerical com-

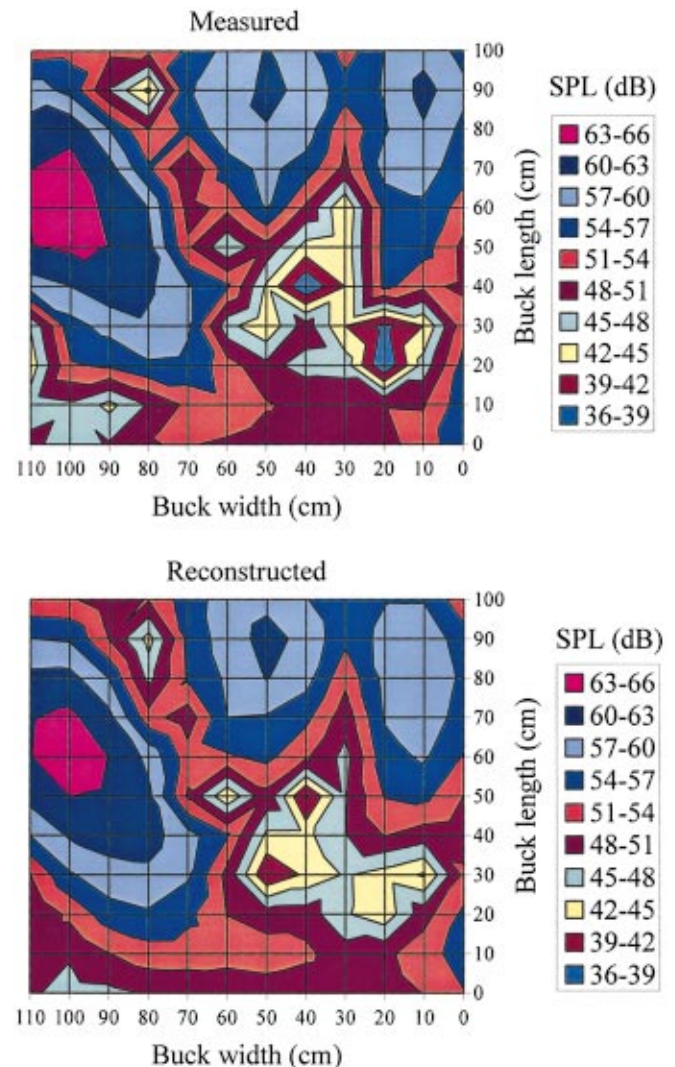


FIG. 12. Comparison of acoustic pressure distributions on the buck top surface subject to harmonic excitations at 372 Hz.

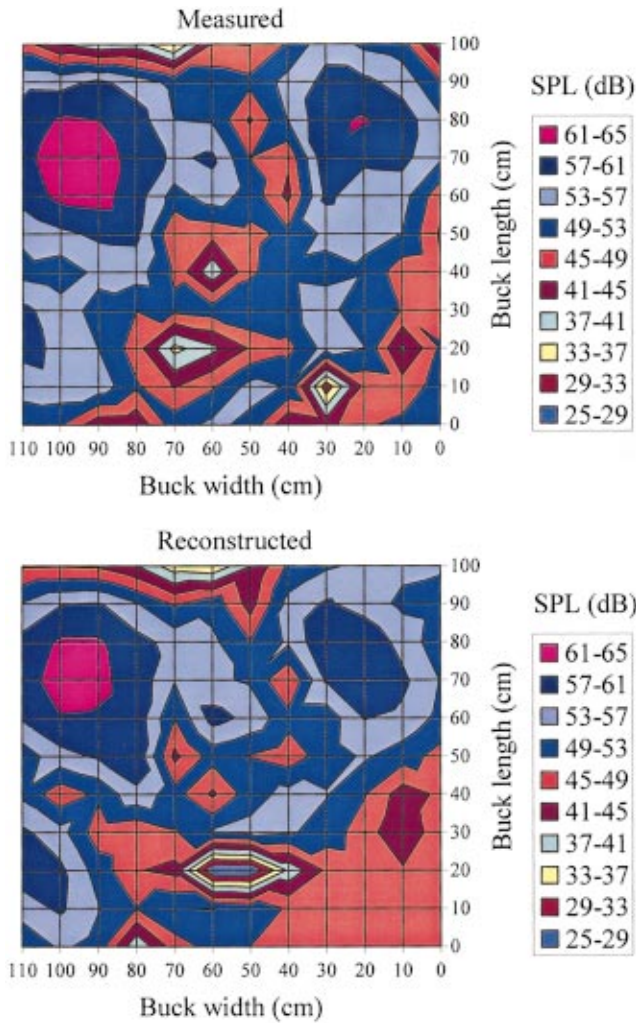


FIG. 13. Comparison of acoustic pressure distributions on the buck top surface subject to harmonic excitations at 488 Hz.

putations of the HELS method may deteriorate at high frequencies. Such an inefficiency difficulty exists in all expansion theories. Hence, other more effective methodology such as asymptotic expansion should be used.

In this paper reconstruction was conducted on an optimal expansion of basis functions. Because the matrices were relatively small, numerical computations were extremely fast. Also, we choose to focus our attention to the frequency range of 0 to 800 Hz, because the B&K Dual Channel Signal Analyzer Type 3550 displays 800 discrete line spectrum. In this way, we can validate the HELS method at every single frequency. For brevity, however, we only show some representative validation results below.

Figures 4–7 demonstrate comparisons of the reconstructed and measured acoustic pressures at 256, 323, 379, and 582 Hz, respectively, on the reconstruction surface, which was parallel to the vehicle buck top surface at 1 cm away, subject to random acoustic excitations. Ideally, the radiated acoustic pressure from a loudspeaker in an unbounded fluid medium would display a monopole character. It is interesting to note that this character was reflected somewhat in the surface acoustic pressure distribution at 256 Hz (see Fig. 4), where the amplitude was the highest at the cen-

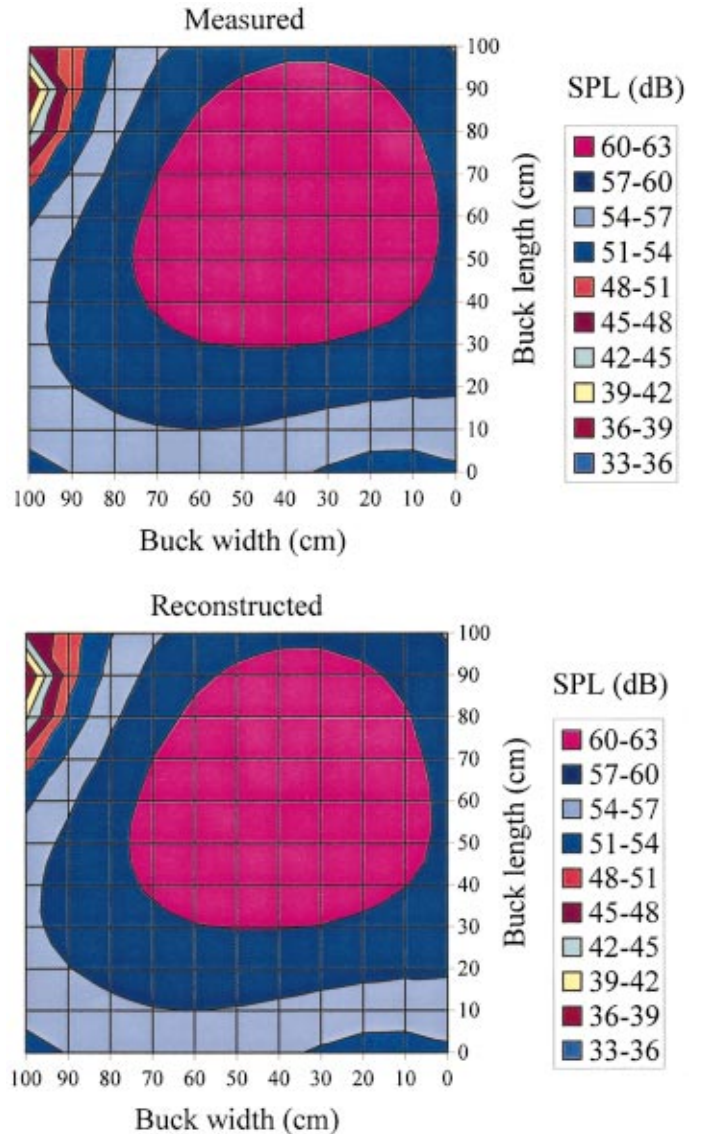


FIG. 14. Comparison of acoustic pressure distributions in the field subject to harmonic excitations at 576 Hz.

ter and decreased radially. This phenomenon may be attributed to the fact that the structure has a strong resonance at 256 Hz (see Fig. 3).

At 323 Hz, the characteristic of acoustic pressure distribution seemed to deviate significantly from that of a monopole. The amplitudes of acoustic pressures were seen to peak out in two circular regions and then reduced gradually outward (see Fig. 5). This may be due to the fact that the structure exhibits a large coupled resonance at 323 Hz (see Fig. 3). A similar phenomenon was observed at 379 Hz (see Fig. 6), at which the structure showed a small coupled resonance. As the frequency increased to 582 Hz, the acoustic pressure distribution became more and more complex. High pressures were observed to concentrate on several spots (see Fig. 7). The corresponding structural vibration may also become quite complicated.

Note that the normal component of the velocity distribution on the buck surface can be obtained by solving the Kirchhoff integral equation, given the surface acoustic pressure distribution. Hence, reconstruction of acoustic pressure

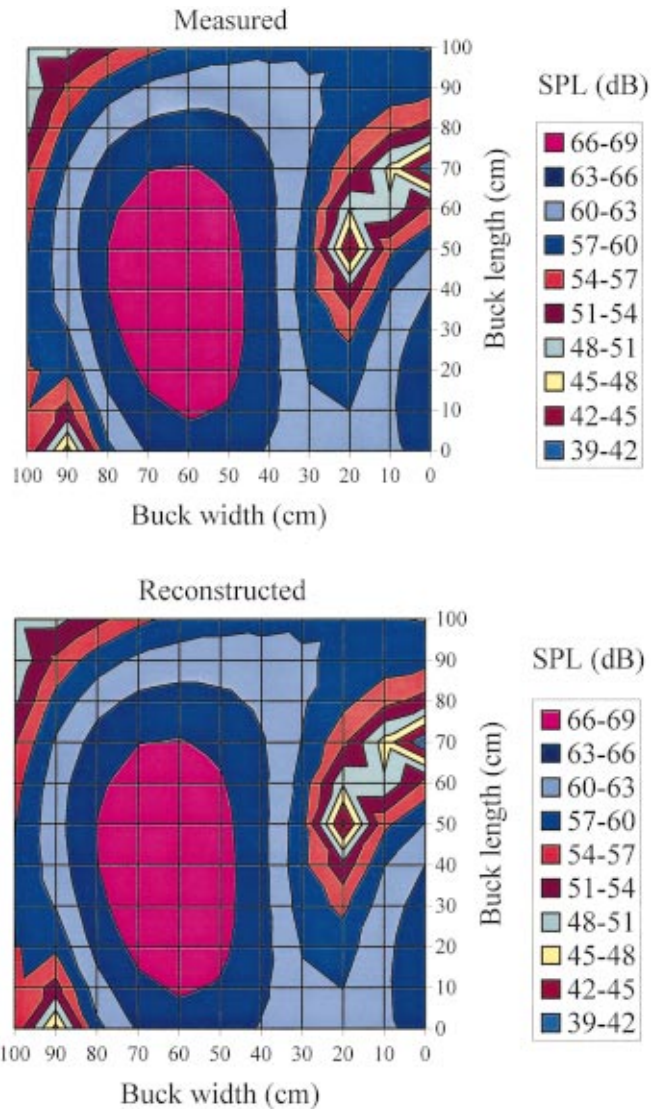


FIG. 15. Comparison of acoustic pressure distributions in the field subject to harmonic excitations at 644 Hz.

is often a first step in noise diagnostics. Figures 4–7 show that favorable agreement between the reconstructed and measured acoustic pressure distributions is obtained in each case.

Figures 8–11 depict comparisons of the reconstructed and measured acoustic pressure spectra around center, edge, and corners of the reconstruction surface. Note that in these plots the spectra started at 200 Hz, which was the cutoff frequency of the anechoic chamber. Once again, the reconstructed acoustic pressure spectra agreed well with the measured ones. However, the discrepancies between the reconstructed and measured spectra increased when the frequency was beyond 770 Hz, especially around the corners.

Next, we demonstrate validations of the second set of experiments, in which the buck was subject to harmonic excitations at 372, 488, 576, and 644 Hz simultaneously. Since in this case we were able to control the amplitudes of the excitation forces by using a B&K Vibration Exciter Controls Type 1050, the acoustic pressure signals were very stable and the input data were very consistent and repeatable. As a result, the accuracy of reconstruction was significantly improved.

For brevity, we only display two comparisons of reconstructed and measured acoustic pressure distributions at 372 and 488 Hz, which correspond to a small coupled resonance and nonresonance frequencies, respectively, at 1 cm above the buck top surface (see Figs. 12 and 13). Figures 14 and 15 demonstrate comparisons of reconstructed and measured acoustic pressure distributions at 576 and 644 Hz, which correspond to a large coupled resonance and anti-resonance frequencies, respectively, at 12 cm above the buck top surface. Because the input acoustic pressure signals are very consistent and stable, the accuracy of reconstruction is much higher than that under random excitation.

## V. CONCLUDING REMARKS

The Helmholtz Equation Least Squares (HELs) method is used successfully to reconstruct the radiated sound pressure fields from a nonspherical structure that has the geometry and dimensions of a real vehicle front end. Experiments are designed to simulate noise radiation from an engine of a passenger vehicle by installing a high fidelity loudspeaker inside the buck at the location of the engine. Both random and harmonic acoustic excitations are employed and the radiated acoustic pressures are measured over a planar surface above the buck. The data thus obtained are taken as input to the HELs formulation to reconstruct acoustic pressure fields both on the buck top surface and in the field. The reconstructed acoustic pressures are compared with those measured at the same locations. Satisfactory agreements between the reconstructed and measured acoustic pressure distributions and spectra are obtained. Results show that there exists an optimal number of expansion functions for a given set of measurements for which the best reconstruction of the acoustic pressure field can be obtained. This optimal number increases with the excitation frequency. Hence the efficiency of numerical computations of the HELs method may deteriorate in the high frequency regime. This difficulty is inherent in all expansion theories.

## ACKNOWLEDGMENTS

The authors would like to thank Dr. Yuhua Wu for his assistance in carrying out the numerical computations during this investigation. This work was supported by a grant from National Science Foundation, Grant No. CMS-9802847.

<sup>1</sup>Z. 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).

<sup>2</sup>S. F. Wu and A. Wang, "Noise diagnostic system," U.S. Patent Number: 5712805, 1998.

<sup>3</sup>S. F. Wu, "On reconstruction of acoustic pressure fields by using HELs method," *J. Acoust. Soc. Am.* **107**, 2511–2522 (2000).

<sup>4</sup>S. F. Wu and Y. Wu, "Reconstruction of radiated acoustic pressure fields from complex vibrating structures," in *Innovative Computational Methods Applied to Linear Structural Mechanics and Acoustics*, edited by S. F. Wu, Proceedings of the ASME Noise Control and Acoustic Division, NCA-Vol. 25, pp. 339–346 (1998).

<sup>5</sup>S. 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).

<sup>6</sup>K. Leach and S. F. Wu, "Visualization of sound radiation from a bowling ball" Proceedings of the ASME Noise Control and Acoustic Division, NCA-Vol. 26, pp. 209–215, Nashville, Tennessee, November 1999.

- <sup>7</sup>Brüel & Kjaer Technical Review, No. 1, 1995.
- <sup>8</sup>E. G. Williams, J. D. Maynard, and E. Skudrzyk, "Sound source reconstruction using a microphone array," *J. Acoust. Soc. Am.* **68**, 340–344 (1980).
- <sup>9</sup>J. D. Maynard, E. G. Williams, and Y. Lee, "Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH," *J. Acoust. Soc. Am.* **78**, 1395–1413 (1985).
- <sup>10</sup>W. A. Veronesi and J. D. Maynard, "Near field acoustic holography (NAH): II. Holographic reconstruction algorithms and computer implementation," *J. Acoust. Soc. Am.* **81**, 1307–1322 (1987).
- <sup>11</sup>LMS Acoustic Holography: Acoustic Troubleshooting by the Spatial Transformation of Sound Fields, LMS International, 1997.
- <sup>12</sup>W. A. Veronesi and J. D. Maynard, "Digital holographic reconstruction of sources with arbitrarily shaped surfaces," *J. Acoust. Soc. Am.* **85**, 588–598 (1989).
- <sup>13</sup>G.-T. Kim and B.-T. Lee, "3-D sound source reconstruction and field reproduction using the Helmholtz integral equation," *J. Sound Vib.* **136**, 245–261 (1990).
- <sup>14</sup>M. R. Bai, "Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries," *J. Acoust. Soc. Am.* **92**, 533–549 (1992).
- <sup>15</sup>B. K. 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).
- <sup>16</sup>E. G. Williams, "An overview of the inverse problem for sound reconstruction in interior spaces," *J. Acoust. Soc. Am.* **105**, 968 (1999).
- <sup>17</sup>D. S. Burnett, "A three-dimensional acoustic infinite element based on a prolate spheroidal multipole expansion," *J. Acoust. Soc. Am.* **96**, 2798–2816 (1994).
- <sup>18</sup>P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1986).
- <sup>19</sup>P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
- <sup>20</sup>I. N. Vekua, "On completeness of a system of metaharmonic functions," *Dokl. Akad. Nauk SSSR* **90**, 715–718 (1953).

# Wave propagation in micro-heterogeneous porous media: A model based on an integro-differential wave equation

Andrzej Hanyga and Vladimir E. Rok<sup>a)</sup>

*Institute for Solid Earth Physics, University of Bergen, Allégaten 41, 5007 Bergen, Norway*

(Received 26 July 1999; accepted for publication 18 February 2000)

A model hyperbolic partial differential equation with singular convolution operators and infinitely smooth solutions is studied. It is shown that short pulses, including finite-bandwidth pulses, propagate with a delay with respect to the wavefront. For a two-parameter family of such equations Green's functions are obtained in a simple self-similar form. As an application, it is demonstrated that the Gurevich–Lopatnikov dispersion law for a thin-layered porous medium can be approximated by a hyperbolic equation with singular memory. © 2000 Acoustical Society of America. [S0001-4966(00)00506-3]

PACS numbers: 43.20.Bi, 43.20.Hq, 43.20.Jr [ANN]

## INTRODUCTION

The dispersion relations of a porous medium are strongly affected by the presence of a viscous boundary layer in the pore fluid flow. The resulting attenuation turns out to be proportional to  $\omega^{1/2}$  in the high-frequency limit  $\omega \rightarrow \infty$ . In addition to Biot's original papers on poroelasticity, there is abundant literature on the frequency dependence of effective density and bulk modulus in poro-acoustic materials<sup>1–9</sup> with a similar frequency behavior of attenuation.

Similar dispersion laws have also been derived from diffusive relaxation.<sup>10–12</sup> Diffusive relaxation in a micro-heterogeneous medium involves a thermodynamic parameter  $\xi$  which depends on local stress, with a different dependence in the inclusions and the host medium. The parameter tends to assume a constant value by a diffusive process or heat conduction. For frequencies  $\omega \gg 1/\tau_R$ , where  $\tau_R$  is a characteristic time, the dispersion law for such materials can be expressed in the form<sup>11,13</sup>

$$k \cong \frac{1}{c_\infty} \omega + i\lambda \sqrt{-i\omega}, \quad (1)$$

where  $k$  denotes the complex wave number,  $c_\infty$  is the high-frequency limit of sound velocity and  $\lambda$  is a constant depending on high- and low-frequency limits of sound velocity and on  $\tau_R$ .

More general fractional-power dispersion laws appear in micro-heterogeneous porous media saturated by a viscous fluid. In the Gurevich–Lopatnikov theory<sup>14</sup> of fast wave dispersion in a thin-layered randomly heterogeneous porous medium the dominant attenuation mechanism is conversion of fast longitudinal waves to diffusive slow waves at the interfaces.

Fractional frequency-power dependence was also found in the theory of wave propagation in a stack of random thin porous layers based on the O'Doherty–Anstey approximation.<sup>15</sup> Dispersive media often exhibit strong frequency sensitivity above a threshold frequency. The thresh-

old frequency for a homogeneous porous medium is determined by the Biot frequency. The threshold frequency in the micro-heterogeneous porous medium can be lower by orders of magnitude.

As shown by Hanyga,<sup>16</sup> frequency-power laws can also be obtained as a high-frequency approximation of a universal mathematical model of wave propagation in porous media proposed by Wilson.<sup>2</sup> Applying the methods of fractional calculus<sup>17,18</sup> and its generalizations it is possible to construct finite-difference approximations for this model as well as for Wilson's model.<sup>16</sup>

Frequency-power laws have also been advocated for some real viscoelastic media<sup>19</sup> while singular memory effects have been studied for viscous fluids.<sup>20,21</sup>

One of the consequences of frequency-power dependence of attenuation in the high-frequency limit is the appearance of time convolution operators with singular kernels,  $K(t) \sim t^{-\alpha}$ ,  $0 < \alpha < 1$ , for  $t \rightarrow 0$ , in the equations of motion of the effective medium. As shown in the papers<sup>22,23</sup> for porous media and in Refs. 24–26, 21, 27 for a class of scalar equations, these singularities entail that all the derivatives of the wave field vanish at the wavefront. For viscoelastic media analogous results were obtained by Renardy,<sup>20,28</sup> Narain and Joseph,<sup>21</sup> and Prüss.<sup>29</sup> As a consequence of wavefront smoothness signals arrive with a delay with respect to the wavefronts. In this context the wavefront is defined in terms of the immediate response of the medium to a suddenly applied deformation while the additional delay depends on the singular part of the memory kernel. For two special classes of scalar PDEs with singular memory the solution has been obtained in an explicit form and the delay can be explicitly calculated for an originally delta-spiked signal.<sup>30,31</sup> For more general scalar equations and hyperbolic systems of equations with singular memory this is possible in the high frequency asymptotic approximation.<sup>32,31</sup>

These observations imply that a wave propagating in a medium with singular memory cannot be viewed merely as an elastic wave with amplitude attenuation. An illustrative case study in Ref. 30 demonstrates the difference between wave attenuation associated with singular and nonsingular hereditary effects.

<sup>a)</sup>On leave from Institute of Geosystems, Varshavskoye sh. 8, Moscow 113105, Russia.



High-frequency asymptotic solutions as well as explicit Green's functions for a medium with singular memory are expressed in terms of a series of explicit infinitely smooth causal functions.<sup>33,22,32</sup> In both cases the signal sent by the source is assumed to be a delta spike. The explicit form of the basis functions allows a quantitative characterization of the pulse delay but the scope of these two methods is limited to initially sharp signals. This raises the question of whether the effect persists for finite-bandwidth signals.

In this paper we have chosen an alternative approach in order to study wavefront smoothing and pulse delay for finite-bandwidth signals: exact numerical solutions of a sufficiently simple model equation. Numerical tests described below confirm a pulse delay superposed on the usual delay associated with finite wave speed.

In the last section it is shown that the complicated dispersion law derived by Gurevich and Lopatnikov<sup>14</sup> for the fast wave attenuation in a micro-heterogeneous porous medium can be approximated by a fractional-power law. Assuming a characteristic frequency of order of tens or hundreds of Hz pulse delay can be significant for seismic waves.

For a hereditary medium the initial conditions it is necessary to specify the initial data in terms of the history of the field variables prior to the initial moment. It is most convenient to assume that the system was at rest prior to the initial moment

$$u(t) = 0 \text{ for } t < 0 \quad (2)$$

(the more general case can be reduced to this special case by modifying the forcing function<sup>29</sup>).

In view of the initial condition Eq. (2) assumed throughout this paper we shall use the one-sided Laplace transformation and its inverse<sup>34,35</sup>

$$v(s) = L_{t \rightarrow s}[u(t)] = \int_0^\infty e^{-st} u(t) dt, \quad s \in \mathbb{C}, \quad (3)$$

$$L_{s \rightarrow t}^{-1}[v(s)] = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} e^{st} v(s) ds, \quad (4)$$

where it is assumed that all the singularities of the analytic function  $v(s)$  lie on the left of the vertical line  $[\sigma - i\infty, \sigma + i\infty]$ . The functions  $u(t)$  considered in this paper are bounded and hence  $\sigma$  can be set equal to 0.

The one-sided Laplace transformation is more convenient than the Fourier transformation for the initial conditions assumed here because it implicitly assumes the initial condition Eq. (2). It has commonly been used in the context hereditary models of wave propagation, including hereditary viscoelasticity.<sup>29,20,21,28</sup> Some formulas used below also appear more cumbersome in the Fourier domain [e.g., Eq. (15) below].

## I. MODEL EQUATIONS

The first-order causal linear equation for a wave  $u(t, x)$  propagating in the positive  $x$  direction and vanishing for  $t < 0$  corresponding to the dispersion relation Eq. (1) is

$$\partial_x u + \frac{1}{c_\infty} \partial_t u - \lambda \int_{-\infty}^t d\tau \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{t-\tau}} \partial_t u(t-\tau, x) = 0. \quad (5)$$

The upper integration limit is justified by causality.

In order to simplify the formulation of the problem we replace  $x, t$  by the nondimensional variables  $t/\tau_R, x/(c_\infty \tau_R)$ , retaining the notation  $t, x$  for the new variables.

The last term on the left-hand side of Eq. (5) can be rewritten in a compact form  $\lambda \phi_{1/2}(t) * \partial_t u(t, x)$ , where the asterisk denotes time convolution and the kernel  $\phi_{1/2}(t)$  is a special case of the more general kernel

$$\phi_{\alpha, \beta}(t) = \frac{t^{-\alpha}}{\Gamma(1-\alpha)} e^{-\beta t} \text{ for } t > 0 \quad (6)$$

with  $0 < \alpha < 1, \beta > 0$  and  $\phi_{\alpha, \beta}(t) = 0$  for  $t < 0$ . The constant  $\beta$  accounts for the fading memory effect, studied later. We shall also use an abbreviated notation  $\phi_\alpha(t) = \phi_{\alpha, 0}(t)$ .

In this paper we shall consider more general dispersion laws

$$k = \omega + \lambda \omega (\beta - i\omega)^{\alpha-1} \quad (7)$$

with the right-going waves satisfying the equation

$$\partial_t u + \partial_x u - \lambda \phi_{\alpha, \beta} * u = 0. \quad (8)$$

The second-order equation for  $u$  is obtained by multiplying Eq. (8) on the left by a corresponding operator vanishing on left-going waves<sup>24</sup>

$$\partial_t^2 u - \partial_x^2 u + (2\lambda \phi_{\alpha, \beta} + \lambda^2 \phi_{\alpha, \beta} * \phi_{\alpha, \beta}) * \partial_t^2 u = 0. \quad (9)$$

In 2D or 3D the operator  $\partial_x^2 u$  is replaced by the Laplacian  $\nabla^2 u$ .

Equation (9) is physically justified for  $\omega \gg 1/\tau_R$  and consequently its solutions faithfully represent the wave field at a distance  $< c_\infty \tau_R$  from the wavefront, or for  $t - x < 1$  in dimensionless variables.

## II. GREEN'S FUNCTIONS FOR ONE-DIMENSIONAL PROBLEMS

The Green's function is defined by the following initial-value problem:

$$(1 + K*) \partial_t^2 u - \nabla^2 u = 0, \quad (10)$$

$$u(0, \mathbf{x}) = 0, \quad \partial_t u(0, \mathbf{x}) = \delta(\mathbf{x}) \quad (11)$$

assuming Eq. (2), with the convolution kernel

$$K(t) = K_\alpha(t) \equiv 2\lambda \phi_\alpha(t) + \lambda^2 \phi_\alpha(t) * \phi_\alpha(t). \quad (12)$$

Solutions of more general initial-value and point source problems for Eq. (10) can be expressed in terms of the Green's function by applying the modified Duhamel principle, formulated in Ref. 25.

The Green's function in one spatial dimension was obtained in the papers in Refs. 25 and 26 (an improved presentation of these results can be found in the book of Dautray and Lions<sup>36</sup>). The original expression can be simplified by taking into account the invariance of the initial-value problem Eqs. (10)–(11) under the transformations  $\lambda \rightarrow \kappa \lambda, r \rightarrow \kappa^{1/(\alpha-1)} r, t \rightarrow \kappa^{1/(\alpha-1)} t$  for arbitrary  $\kappa > 0$ , where  $r = |x|$ .

After some straightforward algebra the Green's function obtained in these papers can be rewritten in terms of a function of two-dimensionless independent variables

$$G_{\alpha}^{(1)}(t, x; \lambda) = g_{\alpha}((t-r)/(\lambda r)^{1/\alpha}, (\lambda r)^{1/\alpha}/(2r)), \quad (13)$$

where

$$g_{\alpha}(\xi, \zeta) = \frac{1}{2} \int_0^{\xi} f_{\alpha}(\sigma) d\sigma + \zeta \int_0^{\xi} \phi_{\alpha}(\sigma) * f_{\alpha}(\sigma) d\sigma. \quad (14)$$

The function  $f_{\alpha}(\sigma)$  is defined by its inverse Laplace transform<sup>37</sup>

$$f_{\alpha}(\sigma) = L_{s \rightarrow 0}^{-1}[\exp(-s^{\alpha})] \quad (15)$$

and its power series expansion is

$$f_{\alpha}(\sigma) = 2H(\sigma) \sum_{n=1}^{\infty} (-1)^{n-1} \sin(n\pi\alpha) \frac{1}{n!} \frac{\Gamma(n\alpha+1)}{\sigma^{n\alpha+1}}, \quad (16)$$

cf. Refs. 25, 24,  $H(\sigma)$  denotes the Heaviside step function. The existence of such a simple two-argument expression for the Green's function results from the fact that the parameter  $\lambda$  can be removed by the scaling  $x_i \lambda^{\gamma} \rightarrow x_i$ ,  $t \lambda^{\gamma} \rightarrow t$ , where  $\gamma = 1/(1-\alpha)$ .

For  $\alpha = 1/2$  the function Eq. (15) can be expressed in closed form

$$f_{1/2}(\sigma) = \frac{1}{2\sqrt{\pi}} \sigma_+^{-3/2} \exp(-1/(4\sigma)), \quad (17)$$

where  $\sigma_+^{\gamma} = H(\sigma)\sigma^{\gamma}$ . Closed-form expressions for  $f_{\alpha}$  also exist for  $\alpha = 1/3, 2/3$ .<sup>33,31</sup>

For general  $\alpha \in ]0, 1[$  the method of steepest descent yields the following asymptotic expansion of the function  $f_{\alpha}(\sigma)$  for small  $\sigma$ :

$$f_{\alpha}(\sigma) \sim H(\sigma) \frac{\alpha^{1/(2-2\alpha)}}{\sqrt{2\pi(1-\alpha)}\sigma^{(2-\alpha)/(2-2\alpha)}} \times \exp\left[-\alpha^{1/(1-\alpha)} \frac{1-\alpha}{\alpha} \sigma^{-\alpha/(1-\alpha)}\right], \quad (18)$$

cf. Refs. 24, 25. Because of nonuniform convergence of the expansion Eq. (16) the asymptotic expansion is often more useful. For  $\alpha = 1/2$  Eq. (18) reduces to the exact formula Eq. (17).

Equation (14) can be simplified with the help of the identity Eq. (A1)

$$g_{\alpha}(\xi, \zeta) = \frac{1}{2} \int_0^{\xi} f_{\alpha}(\sigma) d\sigma + \frac{1}{\alpha} \zeta \int_0^{\xi} \sigma f_{\alpha}(\sigma) d\sigma, \quad (19)$$

reducing the number of integrations. The function  $g_{1/2}(\xi, \zeta)$  can be expressed in closed analytic form (Appendix B)

$$g_{1/2}(\xi, \zeta) = \frac{1}{2} (1-2\zeta) H(\xi) \operatorname{erfc}(1/2\sqrt{\xi}) + \frac{2\zeta\xi_+^{1/2}}{\sqrt{\pi}} \exp(-1/(4\xi)). \quad (20)$$

The corresponding Green's function is

$$G_{1/2}^{(1)}(x, t) = \lambda \frac{\sqrt{t-r}}{\sqrt{\pi}} \exp\left(-\frac{(\lambda r)^2}{4(t-r)}\right) + \frac{1}{2} (1-\lambda^2 r) \operatorname{erfc}\left(\frac{\lambda r}{2\sqrt{t-r}}\right) \quad (21)$$

with  $r = |x|$ .

From Eq. (21) it is immediately seen that the Green's function vanishes with all its derivatives at the wavefront  $r = t$ . The function  $f_{\alpha}(\sigma)$  vanishes at  $\sigma = 0$  with all its derivatives, as can be seen from Eq. (18).

### III. GREEN'S FUNCTIONS FOR TWO- AND THREE-DIMENSIONAL PROBLEMS

Green's functions for three-dimensional problems can be obtained from the one-dimensional one, Eq. (13), by the following formula:<sup>28</sup>

$$G_{\alpha}^{(3)}(t, \mathbf{x}; \lambda) = -\frac{1}{2\pi r} \partial_r G_{\alpha}^{(1)}(t, r; \lambda)|_{r=|\mathbf{x}|}. \quad (22)$$

Substituting Eq. (13) in Eq. (22) and taking into account Eq. (19) we have

$$G_{\alpha}^{(3)}(t, \mathbf{x}; \lambda) = \frac{1}{4\pi r} \left[ (\lambda r)^{-1/\alpha} \left(1 + \frac{t-r}{\alpha r}\right)^2 f_{\alpha}\left(\frac{t-r}{(\lambda r)^{1/\alpha}}\right) + \frac{1}{\alpha} \left(\frac{1}{\alpha} - 1\right) (\lambda r)^{-1/\alpha} \frac{1}{r^2} \times \int_0^{(t-r)/(\lambda r)^{1/\alpha}} \sigma f_{\alpha}(\sigma) d\sigma \right]. \quad (23)$$

The Green's function  $G_{\alpha}^{(2)}$  for two-dimensional problems can be obtained by integration of  $G_{\alpha}^{(3)}$  over the coordinate  $x_3$ . An alternative method of deriving the 2D Green's function is the Cagniard-de Hoop method applied in Ref. 38.

For  $r \equiv t-r \ll r$  the integral on the right-hand side of Eq. (23) is  $O[(\tau/r)^2]$  and

$$G_{\alpha}^{(3)}(r + \tau, \mathbf{x}; \lambda) = \frac{H(\tau)}{4\pi r} \frac{2\tau + \alpha r}{(\lambda r)^{1/\alpha} r^2} f_{\alpha}(\tau(\lambda r)^{-1/\alpha}) + O[(\tau/r)^2], \quad (24)$$

cf. Ref. 39.

For  $\lambda$  tending to zero,  $\lambda^{-1/\alpha} f_{\alpha}(\lambda^{-1/\alpha} \sigma) \rightarrow \delta(\sigma)$  in the sense of distributions. Indeed, the Laplace image  $\exp(-\lambda s^{\alpha})$  of the function on the left-hand side tends to 1. Consequently the Green's function  $G_{\alpha}^{(3)}(t, \mathbf{x}; \lambda)$  tends to  $\delta(t-r)/(4\pi r)$  in the distributions sense, as expected.

$G_{1/2}^{(3)}$  can be expressed in a very simple form<sup>32</sup>

$$G_{1/2}^{(3)}(t, \mathbf{x}; \lambda) = \frac{1}{4\pi r} [F_0(t-r, \lambda r) + 2\lambda F_1(t-r, \lambda r) + \lambda^2 F_3(t-r, \lambda r)], \quad (25)$$

where  $F_n$  are infinitely smooth causal functions defined for  $\xi > 0$  by the formulas

$$F_0(\tau, \xi) = \frac{\xi}{2\sqrt{\pi}} \tau_+^{-3/2} \exp(-\xi^2/(4\tau)), \quad (26)$$

$$F_1(\tau, \xi) = \frac{1}{\sqrt{\pi}} \tau_+^{-1/2} \exp(-\xi^2/(4\tau)), \quad (27)$$

$$F_2(\tau, \xi) = H(\tau) \operatorname{erfc}(\xi/(2\sqrt{\tau})), \quad (28)$$

with  $F_n(\xi) = 0$  for  $\xi < 0$ .

#### IV. THE FADING MEMORY EFFECT

The memory kernels should tend to zero as the delay time increases to infinity. The fading memory effect can be taken into account by multiplying the memory kernel by an exponential damping factor. The fading memory causes a corresponding decay of the wave field at the source after a momentary excitation, as follows from a theorem of Lokshin and Suvorova,<sup>26</sup> but it has a negligible effect on the pulse close to the wavefront, as demonstrated by asymptotic analysis<sup>30</sup> and corroborated by our numerical tests.

The following kernel has the fading memory property:

$$K(t) = \phi_{\alpha, \beta}(t) + \phi_{\alpha, \beta}(t) * \phi_{\alpha, \beta}(t). \quad (29)$$

Note that

$$K(t) = K_{\alpha}(t) e^{-\beta t}. \quad (30)$$

Green's function for Eq. (9) with the exponentially damped kernel Eq. (29) can be represented in the form

$$G_{1/2, \beta}^{(1)}(r, t) = \frac{1}{2} H(t-r) \left\{ \left[ \int_0^{t-r} d\tau f_{1/2, \beta}(r, \tau) \times \left( 1 + \frac{2\tau}{r} \right) \right] - I_{\beta}(r, t) \right\}, \quad (31)$$

where

$$I_{\beta}(r, t) = \beta \int_0^{t-r} d\tau \left( \sqrt{\frac{\tau}{\pi}} \exp(-\beta\tau) * f_{1/2, \beta}(r, \tau) \right) \quad (32)$$

and

$$f_{1/2, \beta}(r, t) = L_{s \rightarrow t}^{-1} \left[ \exp \left( -\frac{rs}{\sqrt{s+\beta}} \right) \right] \\ = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} ds \exp \left( st - \frac{rs}{\sqrt{s+\beta}} \right), \quad (33)$$

where  $\sigma > 0$ .

It is easy to show that

$$f_{1/2, 0}(r, t) = f_{1/2} \left( \frac{t-r}{r^2} \right) \quad (34)$$

and Eq. (31) with  $\beta = 0$  is equal to Eq. (25).

The only cut in the complex plane of the exponential function in the integrand of Eq. (33) starts at the branching point  $-\beta$  and follows the negative part of real axis. Our computation of Eqs. (33) and (31) was carried out numerically by deforming the Bromwich contour in the complex  $s$  plane to a fastest descent path  $\Gamma$  (Fig. 1) and reducing the complex integral to a real one. The contour  $\Gamma$  is defined by the equation

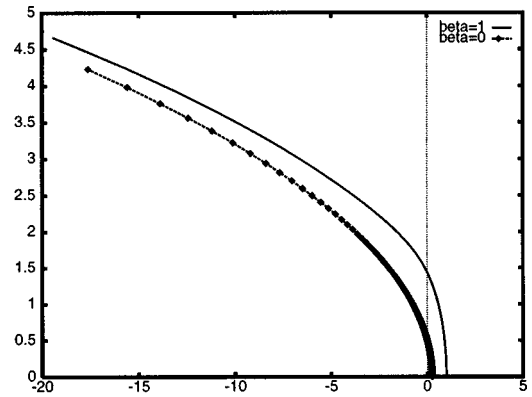


FIG. 1. Upper half of the contour  $\Gamma$  for  $\beta = 0.0, 1.0$ .

$$\operatorname{Im} \left( ts - \frac{rs}{\sqrt{s+\beta}} \right) = 0. \quad (35)$$

In the polar coordinates  $\rho, \varphi$  contour  $\Gamma$  is given by the equation  $\rho = \rho_{\Gamma}(\varphi)$ , where the function  $\rho_{\Gamma}$  can be determined from the equation

$$2t\rho^{3/2} \cos \left( \frac{\varphi}{2} \right) - r\rho - r\beta = 0. \quad (36)$$

If  $\beta \neq 0$ , then it is possible to substitute  $\rho/\beta = R$  and  $r/(2t\sqrt{\beta}) = p$ , and reduce the number of variables in Eq. (36)

$$R^{3/2} \cos \left( \frac{\varphi}{2} \right) - pR - p = 0. \quad (37)$$

The integral on the right side of Eq. (33) is reduced to the following real integral:

$$f_{1/2, \beta}(r, t) = \frac{\beta}{\pi} e^{-\beta t} \int_0^{\pi} d\varphi \{ [R(\varphi) + \eta(\varphi)] / [1 + 3R(\varphi)^{-1}] \} \\ \times \exp [t\beta \{ \eta(\varphi) - R(\varphi) \} / [1 + R(\varphi)^{-1}]], \quad (38)$$

where  $\eta(\varphi) = \cos(\varphi) + 2\cos(\varphi/2)$  and  $R(\varphi)$  is the positive root of Eq. (37).

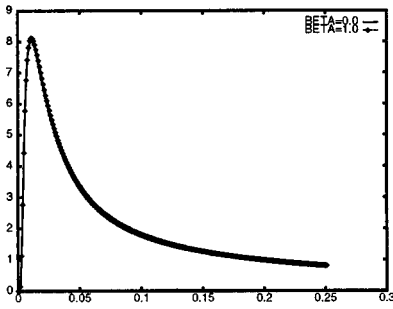
Equation (37) is semi-cubic and it is possible to determine its unique positive solution in closed form. For numerical integration of Eq. (33) we have chosen numerical solution of Eq. (37) because it is more accurate and effective.

One can see from Fig. 2 that the difference between  $G_{1/2}^{(1)}(r, r+\tau)$  and  $G_{1/2, \beta}^{(1)}(r, r+\tau)$  is very small for  $\beta = 1$  and  $\beta = 10$  if  $r \equiv t-r$  is not too large. Consequently, close to the wavefront the fading memory factor  $e^{-\beta t}$  can be safely neglected.

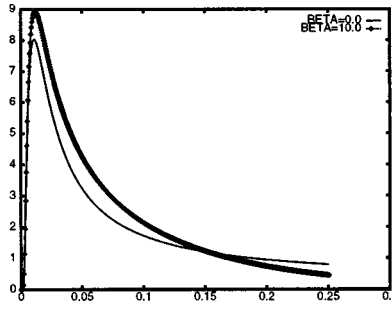
#### V. PROPAGATION OF FINITE BANDWIDTH PULSES

Let us now consider propagation of a finite-bandwidth pulse through a medium with a singular memory kernel.

Equation (9) is linear and causal. Consequently its solution for an arbitrary impulse having a defined front in space and starting point in time can be expressed in terms of Green's function. The Green's functions constructed above have the following behavior at the point  $r = 0$  for  $t \geq 0$



$\beta = 0.0$  and  $\beta = 1.0$ .



$\beta = 0.0$  and  $\beta = 10.0$ .

FIG. 2. Time derivatives of Green's functions for  $\beta=0.0, 1.0$  and  $10.0$  and  $\alpha=0.5$ .

$$G_{\alpha,\beta}^{(1)}(0,t) = \frac{1}{2} \left[ H(t) + \int_0^t d\tau \phi_{\alpha,\beta}(\tau) \right]. \quad (39)$$

Thus, Green's function  $G_{\alpha,\beta}^{(1)}$  can be considered as the solution of Eq. (9) with the boundary conditions Eq. (39) and

$$G_{\alpha,\beta}^{(1)}(r,t) \rightarrow 0 \text{ when } r \rightarrow \infty. \quad (40)$$

The last condition is satisfied because  $G_{\alpha,\beta}^{(1)} = 0$  for  $r > t$ . The convolution product of Green's functions  $G$  and any given arbitrary smooth and integrable function  $g(t)$  satisfies for  $t > 0$  Eq. (9) and the boundary condition

$$(G_{\alpha,\beta}^{(1)} * g)(0,t) = \frac{1}{2} \left[ \int_0^t d\tau g(\tau) + \int_0^t d\tau (\phi_{\alpha,\beta} * g)(0,\tau) \right]. \quad (41)$$

We can take  $G_{\alpha,\beta}^{(1)} * g$  as a particular solution of Eq. (9) satisfying the conditions Eqs. (40) and (41). The function

$$v(r,t) = \partial_t (G_{\alpha,\beta}^{(1)} * g) = (\partial_t G_{\alpha,\beta}^{(1)} * g) \quad (42)$$

is also the solution of Eq. (9) satisfying the boundary condition

$$v(0,t) = \frac{1}{2} [g(t) + (\phi_{\alpha,\beta} * g)(0,t)]. \quad (43)$$

$v(0,t)$  is the excitation applied at  $r=0$  and  $v(r,t)$  is the wave emitted by this excitation.

It is obvious from the expression on the right-hand side of Eq. (42), that the excitation function  $v(0,t)$  is at least as smooth as  $g$ . Moreover if the support of  $g$  is finite then the support of  $v(0,t)$  is also finite.

Let us choose  $g(t)$  in the form

$$g(t) = H \left( 1 - \left| 1 - \frac{t}{T} \right| \right) \sin \left( \frac{\pi n}{T} t \right) \exp \left( - \frac{1}{1 - (1 - t/T)^2} \right), \quad (44)$$

where  $n$  is an integer. This function is causal, infinitely differentiable, vanishes when  $t \rightarrow 0+0$  and  $t \rightarrow 2T-0$ , and its support is a finite interval  $[0, 2T]$ . If the function  $g$  is extended periodically its spectrum lies above the frequency  $\pi/T$  and the dominant frequency is  $\pi n/T$ . The spectrum decreases exponentially for higher frequencies. If the input signal is periodically continued its spectrum has the shape shown in Fig. 3. Consequently the signal has an effectively finite bandwidth. The function  $v(0,t)$  automatically enjoys similar properties.

Substitution of Eq. (44) into the second term of the boundary condition Eq. (43) yields a convolution integral with a singular integrand because of the singularity of functions  $\phi_{\alpha,\beta}$ . The singularity is integrable and the integral can be evaluated by integration by parts. The final form of the boundary condition is

$$v(0,t) = \frac{1}{2} g(t) + \frac{1}{2\Gamma(2-\alpha)} \times \int_0^t d\tau (t-\tau)^{1-\alpha} \partial_\tau [g(\tau) e^{-\beta(t-\tau)}]. \quad (45)$$

The integral on the right-hand side of Eq. (45) is well suited for numerical computations.

## VI. NUMERICAL TESTS

In Fig. 4 the output signal for  $\alpha = 1/2$  at  $x = 0.25$  is compared with the input signal  $g(t)$ . The input signal is shifted in time in such a way that it starts at the time of the wavefront arrival. The output signal is vertically scaled by a factor  $V$  in order to match the heights of the highest peaks of the input and output signals. The additional time shift is thus

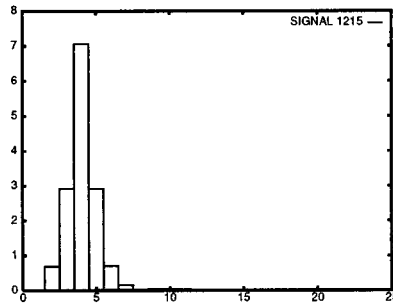
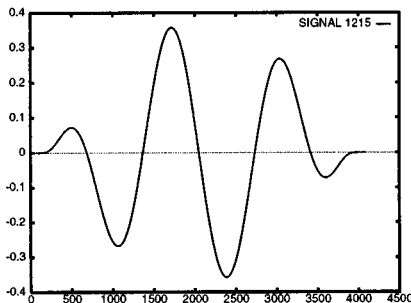
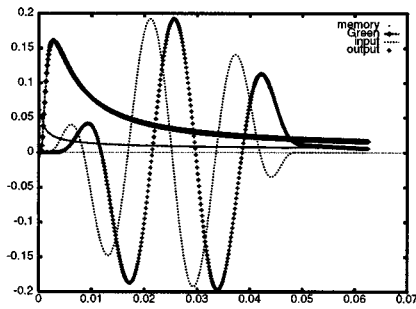
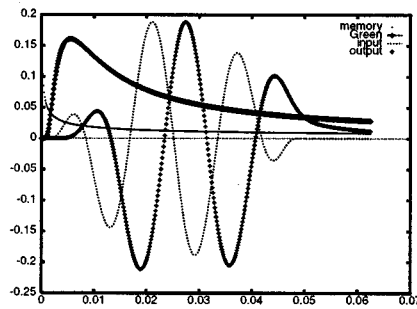


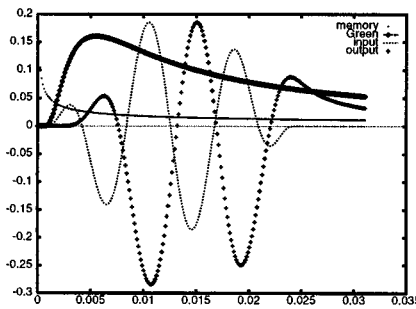
FIG. 3. A typical input signal and the spectrum of its periodic counterpart.



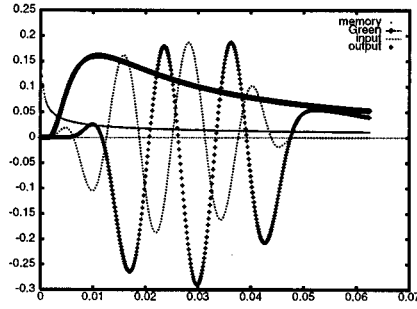
$T = 0.1, \omega_M = 15, V = 5.65.$



$T = 0.05, \omega_M = 30, V = 11.90.$



$T = 0.025, \omega_M = 60, V = 36.84.$



$T = 0.025, \omega_M = 120, V = 65.71.$

FIG. 4. Input and output signals for  $\alpha=0.5$  and  $x=0.25$ .

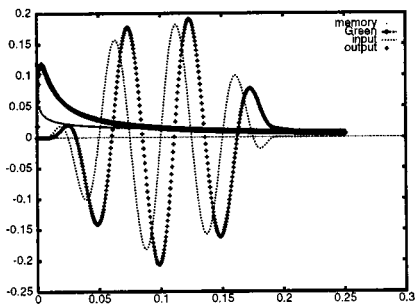
easy to see. The symbol  $\omega_M$  denotes the dimensionless frequency corresponding to the maximum in the spectrum of the input signal.

The memory function (vertically downscaled 1:500) and Green's function (downscaled 1:50) are shown too.

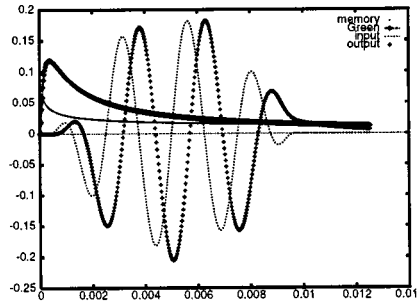
In Fig. 5 a similar comparison is made for  $\alpha=1/3$ .

Time delay of the peaks of Green's functions are clearly visible. The peak of the explicit Green's function for  $\alpha = 1/2$  arrives with a delay  $\Delta T = (\lambda r)^2/6$  after the wavefront. Asymptotic theory developed in Ref. 31 yields a time delay of  $16(\lambda r)^3/27$  for the  $t^{-1/3}$  singularity.

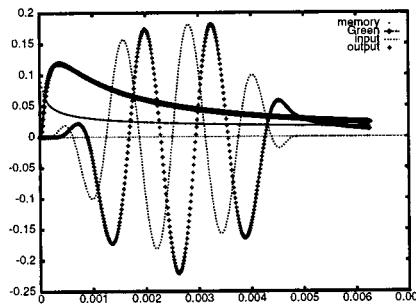
Time delay of the finite-bandwidth signal can be esti-



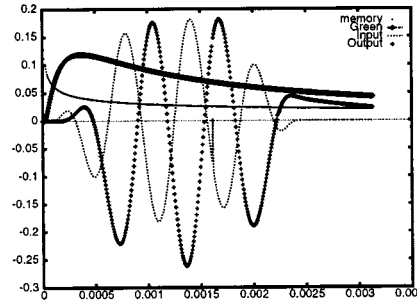
$T = 0.01, \omega_M = 150.$



$T = 0.005, \omega_M = 300.$

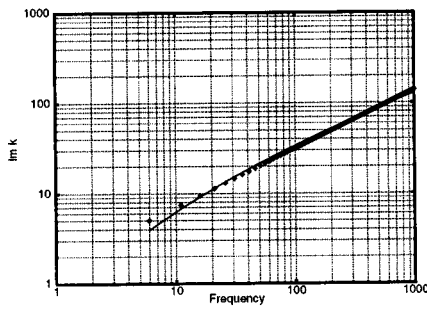


$T = 0.0025, \omega_M = 600.$

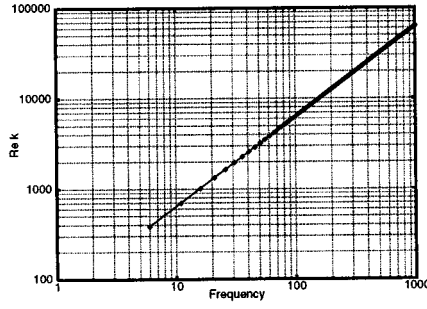


$T = 0.00125, \omega_M = 1200.$

FIG. 5. Input and output signals for  $\alpha=1/3$  and  $x=0.25$ .



Im  $k$  vs  $\omega/\omega_0$ .



Re  $k$  vs  $\omega/\omega_0$ .

FIG. 6. Comparison of Gurevich–Lopatnikov dispersion (solid line) with the  $\omega^{0.65}$  power law (dotted line);  $k$  denotes the complex wave number.

ated by comparing the position of the highest peak in the input and output signal. In order to facilitate the comparison, the output signal has been rescaled by a factor  $V$ . The delay is slightly frequency dependent. As expected, time delay is less pronounced for the weaker singularity ( $t^{-1/3}$  in our case).

The scale factors in the figure captions indicate the attenuation of the signal during its propagation.

The parameter  $\beta$  has little relevance for the signal propagating with the wavefront and consequently can be set equal to 0 for our purposes. A positive  $\beta$  is, however, required for the fading memory usually assumed in viscoelasticity. A positive  $\beta$  also ensures exponential decrease of the wave field at the source.

## VII. COMPARISON WITH THE GUREVICH–LOPATNIKOV THEORY OF MICRO-HOMOGENEOUS POROUS MEDIA

We shall demonstrate that, for all its simplicity, the dispersion relation associated with Eq. (9) with the memory kernel  $K_\alpha$  provides a very good approximation the Gurevich–Lopatnikov dispersion formula.<sup>14</sup>

According to Biot’s theory of fluid-saturated poroelastic media<sup>40</sup> memory effects with  $K(t) \sim t^{-1/2}$  should appear for frequencies  $\omega > \omega_B$ , where

$$\omega_B = \frac{\eta\phi}{\kappa\rho_f} \quad (46)$$

denotes the characteristic Biot frequency,  $\phi$  is the porosity,  $\kappa$  is the permeability and  $\eta, \rho_f$  denote the dynamic fluid viscosity and density. Biot’s characteristic frequency is rather large (typically of order of  $10^5$  Hz).

Biot’s model assumes that the medium is homogeneous at the sub-wavelength scale. For fast wave propagation in a micro-heterogeneous poroelastic medium scattering and fast-to-slow wave conversion has to be accounted for. The second effect can be described in terms of longitudinal wave propagation through a stack of random thin poroelastic layers with fast-to-slow conversion at the layer boundaries. According to the results of Refs. 14, 15 such a model exhibits another characteristic frequency

$$\omega_0 = \frac{\kappa N}{\eta a^2}, \quad (47)$$

where  $a$  denotes the autocorrelation length of the heterogeneities,

$$N = M \frac{K + 4\mu/3}{K + 4\mu/3 + \sigma M}, \quad (48)$$

$$M = \left( \frac{\sigma - \phi}{K_s} + \frac{\phi}{K_f} \right)^{-1}, \quad (49)$$

$\sigma = 1 - K/K_s$  and  $K, K_s, K_f, \mu$  denote the bulk modulus of the elastic skeleton, the bulk modulus of the elastic grains, the bulk modulus of the fluid and the shear modulus of the skeleton. The characteristic frequency  $\omega_0$  can be of order 10–100 Hz for fluid-saturated porous rocks, several orders of magnitude below Biot’s characteristic frequency  $\omega_B$ .<sup>14,41</sup>

For signals with the bandwidth in the range  $\omega_0 < \omega < \omega_B$  the real and imaginary part of the wave number vector derived from the Gurevich–Lopatnikov theory can be approximated by the causal frequency-power dispersion law

$$c_\infty k = \omega + i\lambda(-i\omega)^\alpha. \quad (50)$$

Very good matching is obtained for  $\alpha = 0.65$ , as can be seen from Fig. 6 where the frequencies range from  $3\omega_0$  to  $800\omega_0$ ,  $\omega_0 = 10$  Hz and  $\lambda = 0.055$ .

The comparison indicates that Eq. (9) with the memory kernel  $K(t) = K_{0.65}(t)$  matches the Gurevich–Lopatnikov model very closely over a wide range of frequencies.

## VIII. CONCLUSIONS

Explicit estimates of the pulse delay for the case of delta-spiked signals can be derived from explicit Green’s functions, such as Eq. (25) for  $K(t) \sim \text{const } t^{-1/2}$  [see Ref. 31 for  $K(t) \sim \text{const } t^{-2/3}, \text{const } t^{-1/3}$ ]. Numerical tests presented in this paper show that the pulse delay persists for finite-bandwidth signals, exhibiting some dependence on the dominant frequency of the pulse. The delay may be observable for the effective model of a heterogeneous porous medium constructed by Gurevich and Lopatnikov.

It has been shown that fading memory has little relevance for the wavefront aspects of the pulse propagation (cf. also Ref. 30).

A comparison with the Gurevich–Lopatnikov theory shows that for seismic frequencies the effective medium can be described by Eq. (9) with a memory kernel that has a singularity  $t^{-0.65}$ .

We conclude that a relatively simple equation with self-similarity properties gives a faithful representation of pulse propagation in complicated media with pulse delay effects.

## ACKNOWLEDGMENTS

Part of V. E. Rok's research was carried out at the University of Bergen and sponsored from a grant of the European Commission under the Contract No. JOF3-CT95-0019. The paper is a contribution to the Joule II project "Reservoir Oriented Delineation Technology." The authors are indebted to Malgorzata Seredyńska for a critical review of the first draft, to Boris Gurevich and to an anonymous reviewer for critical remarks concerning the presentation.

## APPENDIX A: AN IDENTITY

In order to derive Eq. (19) from Eq. (14) the following identity is needed:<sup>39</sup>

$$\phi_\alpha(\sigma) * f_\alpha(\sigma) = \frac{1}{\alpha} \sigma f_\alpha(\sigma). \quad (\text{A1})$$

In order to prove Eq. (A1) we note that the Laplace transform of the left-hand side is

$$\lambda s^{\alpha-1} \exp(-\lambda s^\alpha) \equiv -\frac{1}{\alpha} \frac{d}{ds} \exp(-\lambda s^\alpha),$$

which is the Laplace transform of the right-hand side of Eq. (A1).

## APPENDIX B: DERIVATION OF $g_{1/2}$

The function  $g_{1/2}$  can be evaluated explicitly. To this effect we shall prove the identity:

$$\int_0^\xi \sigma f_{1/2}(\sigma) d\sigma = -\frac{1}{2} \int_0^\xi f_{1/2}(\sigma) d\sigma + 2\xi^2 f_{1/2}(\xi). \quad (\text{B1})$$

For the proof of the identity we note that the Laplace image of the left-hand side is  $\alpha s^{\alpha-2} e^{-s^\alpha}$ . On the other hand, the Laplace image of the function  $\xi^2 f_\alpha(\xi)$  is

$$(d/ds)^2 e^{-s^\alpha} \equiv \alpha[\alpha s^{2\alpha-2} + (1-\alpha)s^{\alpha-2}] e^{-s^\alpha}.$$

For  $\alpha = 1/2$  the right-hand side becomes

$$\frac{1}{2}(\frac{1}{2}s^{-1} + \frac{1}{4}s^{-3/2})e^{-s^{1/2}}.$$

The corresponding source function is

$$\frac{1}{4} \int_0^\xi f_{1/2}(\sigma) d\sigma + \frac{1}{2} \int_0^\xi \sigma f_{1/2}(\sigma) d\sigma,$$

which proves the identity Eq. (B1).

The final formula Eq. (20) for  $g_{1/2}$  is obtained by combining Eq. (B1) with the formulas

$$f_{1/2}(\sigma) = \frac{1}{2\sqrt{\pi}} \sigma_+^{-3/2} \exp(-1/(4\sigma)), \quad (\text{B2})$$

$$\int_0^\xi f_{1/2}(\sigma) d\sigma = H(\xi) \operatorname{erfc}(1/(2\sqrt{\xi})) \quad (\text{B3})$$

(Ref. 42, 29.3.82–83).

<sup>1</sup>D. L. Johnson, J. Koplík, and R. Dashen, *J. Fluid Mech.* **176**, 379 (1987).

<sup>2</sup>D. K. Wilson, *J. Acoust. Soc. Am.* **94**, 1136 (1992).

<sup>3</sup>D. K. Wilson, *Appl. Acoust.* **50**, 171 (1997).

<sup>4</sup>J.-F. Allard and Y. Champoux, *J. Acoust. Soc. Am.* **91**, 3346 (1992).

<sup>5</sup>J. F. Allard, M. Henry, J. Tizianel, L. Kelders, and W. Lauriks, *J. Acoust. Soc. Am.* **104**, 2525 (1998).

<sup>6</sup>J. F. Allard, M. Henry, J. Tizianel, L. Kelders, and W. Lauriks, *J. Acoust. Soc. Am.* **105**, 3021 (1999).

<sup>7</sup>M. R. Stinson and Y. Champoux, *J. Acoust. Soc. Am.* **91**, 685 (1992).

<sup>8</sup>Y. Champoux and M. R. Stinson, *J. Acoust. Soc. Am.* **92**, 1120 (1992).

<sup>9</sup>T. Yamamoto and A. Tugut, *J. Acoust. Soc. Am.* **83**, 1744 (1988).

<sup>10</sup>M. Kelbert and I. Sazonov, *Pulses and Other Wave Processes in Fluids: An Asymptotic Approach to Initial Problems* (Kluwer Academic, Dordrecht, 1996).

<sup>11</sup>M. Y. Kelbert and I. Y. Chaban, *Izv. Ak. Nauk, ser. Mechanics of fluids and gases* **5**, 153 (1986).

<sup>12</sup>S. Hunklinger and W. Arnold, "Ultrasonic properties of glasses at low temperatures," in *Physical Acoustics*, edited by P. M. Warren and R. N. Thurston, Vol. 17 (Academic, New York, 1976), pp. 155–215.

<sup>13</sup>V. E. Rok, "Near-front spatial profile of wave impulses in microinhomogeneous media," in Extended Abstracts EAGE 57th Conference and Technical Exhibition, Glasgow, 29 May–2 June 1995, p. B008, 1995.

<sup>14</sup>B. Gurevich and S. Lopatnikov, *Geophys. J. Int.* **121**, 933 (1995).

<sup>15</sup>S. Gelinsky, S. A. Shapiro, T. Müller, and B. Gurevich, *Int. J. Solids Struct.* **35**, 4739 (1998).

<sup>16</sup>A. Hanyga, "Simple memory models of attenuation in complex viscoporous media," in Proceedings of the 1st Canadian Conference on Non-linear Solid Mechanics, Victoria, BC, June 16–20, 1999, Vol. 2, pp. 420–436, 1999.

<sup>17</sup>I. Podlubny, *Fractional Differential Equations* (Academic, San Diego, 1998).

<sup>18</sup>F. Mainardi, "Fractional calculus: Some basic problems," in *Fractals and Fractional Calculus in Continuum Mechanics*, Wien, 1997, Springer, CISM Courses and Lectures, no. 378.

<sup>19</sup>T. L. Szabo, *J. Acoust. Soc. Am.* **97**, 14 (1994).

<sup>20</sup>M. Renardy, *Rheol. Acta* **21**, 251 (1982).

<sup>21</sup>A. Narain and D. D. Joseph, *Rheol. Acta* **21**, 228 (1982).

<sup>22</sup>F. Mainardi, G. Servizi, and G. Turchetti, *J. Geophys.* **43**, 83 (1977).

<sup>23</sup>A. Hanyga, "Asymptotic theory of wave propagation in viscoporelastic media," in *Theoretical and Computational Acoustics '97*, edited by Y.-C. Teng, E.-C. Shang, Y.-H. Pao, M. H. Schultz, and A. D. Pierce, Singapore, 1999, World-Scientific, Proc. 3rd Int. Conf. on Computational and Theoretical Acoustics, Newark, NJ, July 14–18, 1997.

<sup>24</sup>A. A. Lokshin and V. E. Rok, *Russian Math. Surveys* **33**, 243 (1978).

<sup>25</sup>A. A. Lokshin and V. E. Rok, *Dokl. Akad. Nauk SSSR* **239**, 1305 (1978).

<sup>26</sup>A. A. Lokshin and Y. V. Suvorova, *Mathematical Theory of Wave Propagation in Media with Memory* (Moscow University, Moscow, 1982) (in Russian).

<sup>27</sup>V. E. Rok, *Geoinformatika-98* **3**, 63 (1998) (in Russian).

<sup>28</sup>M. Renardy, W. J. Hrusa, and J. A. Nohel, *Mathematical Problems in Viscoelasticity* (Longman Scientific & Technical, Essex, 1987).

<sup>29</sup>J. Prüss, *Evolutionary Integral Equations* (Birkhäuser Verlag, Basel, 1993).

<sup>30</sup>A. Hanyga and M. Seredyńska, *Geophys. J. Int.* **137**, 319 (1999).

<sup>31</sup>A. Hanyga and M. Seredyńska, "Asymptotic wavefront expansions in hereditary media with singular memory kernels," to appear in *Q. Appl. Math.* (2000).

<sup>32</sup>A. Hanyga and M. Seredyńska, *Wave Motion* **30**, 175 (1999).

<sup>33</sup>P. W. Buchen and F. Mainardi, *J. Mec.* **14**, 597 (1975).

<sup>34</sup>D. V. Widder, *The Laplace Transform* (Princeton University Press, Princeton, 1946).

<sup>35</sup>G. Doetsch, *Einführung in Theorie und Anwendung der Laplace Transformation* (Birkhäuser Verlag, Basel, 1958).

<sup>36</sup>R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology* (Springer-Verlag, Berlin, 1992), Vol. 5.

<sup>37</sup>H. Pollard, *Bull. Am. Math. Soc.* **52**, 908 (1946).

<sup>38</sup>S. L. Lopatnikov, A. A. Lokshin, and V. E. Rok, *Izv. Akad. Nauk SSSR, ser. Mekhanika Tverdogo Tela* **5**, 188 (1990).

<sup>39</sup>V. E. Rok, "Time-domain representation of waves in media with frequency power law of dispersion," in Extended Abstracts EAGE 58th Conference and Technical Exhibition, Amsterdam, 3–7 June 1996, p. C008, 1996.

<sup>40</sup>M. A. Biot, *J. Appl. Phys.* **33**, 1482 (1962).

<sup>41</sup>S. Gelinsky and S. A. Shapiro, *Geophys. J. Int.* **128**, F1 (1997).

<sup>42</sup>M. Abramowitz and I. Stegun, *Mathematical Tables* (Dover, New York, 1970).

# Directive line source model: A new model for sound diffraction by half planes and wedges

Penelope Menounou<sup>a)</sup>

*Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712-1063*

Ilene J. Busch-Vishniac<sup>b)</sup>

*Whiting School of Engineering, Johns Hopkins University, Baltimore, Maryland 21218*

David T. Blackstock<sup>c)</sup>

*Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712-1063*

(Received 20 September 1999; revised 17 December 1999; accepted 18 February 2000)

A new method termed Directive Line Source Model (DLSM) is presented for predicting the diffracted field produced by a sound wave incident on a rigid or pressure release half plane. In the new method the edge of the half plane is modeled as an infinite set of directive point sources continuously distributed along the edge. Because DLSM is fast, simple and intuitive, it is a promising tool for the study of diffraction. It can be applied for several types of incident radiation: omnidirectional cylindrical and spherical waves, plane waves, as well as waves from directional sources. Wedges may also be treated. Finally, DLSM can handle diffraction by an arbitrarily shaped edge profile, for example, a half plane having an edge that is jagged instead of straight. Results for plane, cylindrical and spherical incident waves, as well as for arrays of line and point sources, are presented and their agreement with known analytical solutions is demonstrated. Predictions based on DLSM compare favorably with experimental data. © 2000 Acoustical Society of America.

[S0001-4966(00)00406-9]

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

## INTRODUCTION

Rigorous solutions for diffraction by a half plane began to appear more than a century ago. Solutions have been obtained in both the frequency and time domain and by a variety of methods (see Bowman and Senior<sup>1</sup> for a summary). However, researchers continue to work on diffraction, mainly because the rigorous solutions are restricted by certain assumptions: the half plane is infinitely thin, its edge is straight and infinitely long, the incident field is produced by simple sources. Moreover, the solutions are different for plane, cylindrical and spherical incident waves. An additional motivation for study is presented by the fact that the rigorous solutions are hard to evaluate computationally. Alternatives and appropriate approximations are therefore constantly in demand. In the present work a new method is presented, which, besides being easy to compute, can be applied for several types of incident waves, as well as for arbitrarily shaped diffracting edges.

The typical geometry for the half plane diffraction problem is shown in Fig. 1, where the sound field is produced by a point source  $S$  located at  $(r_0, \phi_0, z_0)$  and the receiver  $A$  can be any point  $(r, \phi, z)$ . The notation is based on a cylindrical coordinate system having the edge of the half plane as its  $z$ -axis. The radial distance  $r$  is measured from the edge of the half plane and the angle  $\phi$  is measured from the surface of the half plane facing the source. For the rest of this paper, the source is assumed to be located at  $z_0=0$ . The sound from the

source can be either normally incident on the half plane ( $S$  and  $A$  in the same plane normal to the half plane  $z_0=z=0$ ) or obliquely incident ( $z_0=0 \neq z$ ). The shortest distance that the sound travels to reach the receiver by diffraction is  $L = \bar{s} + \bar{r}$ , where  $\bar{s} = SN$ ,  $\bar{r} = NA$  and  $N$  is the point where the least diffraction path intersects the edge. The distance between source and receiver is denoted by  $R_1$ , and the distance from the image source to the receiver by  $R_2$ . Based on geometrical acoustics, in region I [shown in Fig. 1(b)] both incident and reflected waves are present, in II only the incident wave is present, and III is a shadow zone. The dividing line at  $\phi = \pi + \phi_0$  ( $B_i$ ) that separates regions II and III establishes a *shadow boundary* for the incident wave and the dividing line at  $\phi = \pi - \phi_0$  ( $B_r$ ) establishes a shadow boundary for the reflected wave.

The new method presented in this work is termed Directive Line Source Model (DLSM). The model is based on appropriate approximate forms of known exact solutions in the frequency domain (that by Sommerfeld<sup>2,3</sup> for plane incident waves, by Carslaw<sup>4,5</sup> for cylindrical incident waves, and by MacDonald<sup>6,7</sup> for cylindrical and spherical incident waves). The DLSM is generalized to address equally both the frequency and time domain, as well as several types of incident radiation. Results obtained by DLSM are compared with experimental data, as well as with known exact solutions in the time domain (that by Friedlander<sup>8</sup> for plane and cylindrical incident waves, and by Biot, Tolstoy,<sup>9</sup> and Medwin<sup>10</sup> for spherical incident waves).

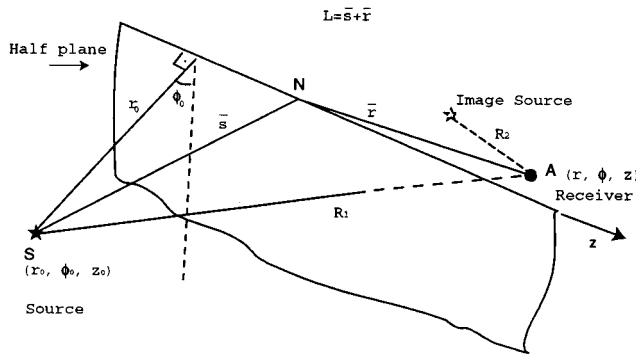
More specifically the contents of the present work are as follows: In Sec. I common characteristics of the diffracted field produced by plane, cylindrical and spherical incident

<sup>a)</sup>Electronic mail: menounou@mail.utexas.edu

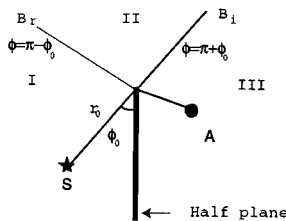
<sup>b)</sup>Electronic mail: ilenebv@jhu.edu

<sup>c)</sup>Electronic mail: dtb@mail.utexas.edu





(a) perspective view



(b) side view

FIG. 1. Half plane—source—receiver configuration.

waves are outlined. In Sec. II the common characteristics are used to develop a unified representation for the diffracted field. The diffracted field is shown to be equivalent to re-radiation from an infinite set of directive point sources, regardless of the type of the incident wave. The applicability of DLSM is discussed for a variety of other cases including pressure release half planes, wedges, and half planes having an arbitrarily shaped edge profile. In Sec. III results obtained by DLSM are compared with known solutions for plane, cylindrical and spherical incident waves, as well as for incident waves from directional sources, such as arrays of point or line sources. Finally, results for spherically spreading incident waves on a rigid half plane are compared with experimental data.

## I. COMMON CHARACTERISTICS OF DIFFRACTION BY PLANE, CYLINDRICAL AND SPHERICAL INCIDENT WAVES ON A HALF PLANE

Although the exact solutions for diffraction by plane, cylindrical and spherical time harmonic incident waves have different mathematical formulations and have been derived by different methods, all share some common characteristics.

The diffracted field  $P_d$  ( $p_d = P_d e^{-i\omega t}$ ) produced by a plane, cylindrical and spherical time harmonic incident wave can be written<sup>3,7,11</sup> in terms of Fresnel integrals  $F(w)$ , as shown in Table I. The expressions are exact for plane incident waves and approximate for cylindrical and spherical

TABLE I. Diffracted field in terms of Fresnel integrals  $F$ —exact for plane incident waves, approximate for cylindrical and spherical waves. The solutions consist of two terms: the first is associated with the incident field (parameters:  $\rho_1$ ,  $m_1$ ,  $R_1$ ), the second with the reflected field (parameters:  $\rho_2$ ,  $m_2$ ,  $R_2$ ).

Type of incident wave	Diffracted wave $P_d$	Assumptions
Plane	$P_d = -\text{sgn}(\rho_1) e^{-ikr \cos(\phi - \phi_0)} \left( \frac{1-i}{2} \right) [F(\infty) - F( \rho_1 )]$ $-\text{sgn}(\rho_2) e^{-ikr \cos(\phi + \phi_0)} \left( \frac{1-i}{2} \right) [F(\infty) - F( \rho_2 )]$	None
Cylindrical	$P_d = \text{sgn}(m_1) \frac{2}{i\pi} \sqrt{\frac{\pi}{2}} \frac{e^{ikR_1}}{\sqrt{k(L+R_1)}} [F(\infty) - F( m_1 )]$ $+ \text{sgn}(m_2) \frac{2}{i\pi} \sqrt{\frac{\pi}{2}} \frac{e^{ikR_2}}{\sqrt{k(L+R_2)}} [F(\infty) - F( m_2 )]$	$kL \gg 1$
Spherical	$P_d = \text{sgn}(m_1) \frac{k e^{-i(\pi/4)}}{\sqrt{kL}} \frac{e^{ikR_1}}{\sqrt{k(L+R_1)}} [F(\infty) - F( m_1 )]$ $+ \text{sgn}(m_2) \frac{k e^{i(\pi/4)}}{\sqrt{kL}} \frac{e^{ikR_2}}{\sqrt{k(L+R_2)}} [F(\infty) - F( m_2 )]$	$kL \gg 1$

$$F(w) = \int_0^w e^{i\pi\tau^2/2} d\tau$$

$$\rho_1 = 2\sqrt{\frac{kr}{\pi}} \cos \frac{\phi - \phi_0}{2}, \quad \rho_2 = 2\sqrt{\frac{kr}{\pi}} \cos \frac{\phi + \phi_0}{2}$$

$$m_1 = -\text{sgn} \left( \cos \frac{\phi - \phi_0}{2} \right) \sqrt{k(L-R_1)}, \quad m_2 = -\text{sgn} \left( \cos \frac{\phi - \phi_0}{2} \right) \sqrt{k(L-R_2)}$$

TABLE II. Asymptotic forms of the solutions given in Table I. The forms hold for receiver locations away from the shadow boundaries and the edge of the half plane.

Type of incident wave	Diffracted wave $P_d$	Assumptions
Plane	$P_d = -\frac{i}{4} \sqrt{\frac{2}{kr\pi}} e^{ikr-i(\pi/4)} D(\phi_0; \phi)$	$ \rho_1 ,  \rho_2  \gg 1$
Cylindrical	$P_d = -\frac{e^{ikr}}{\sqrt{2\pi kr}} \frac{e^{ikr_0}}{\sqrt{2\pi kr_0}} D(\phi_0; \phi)$	$ m_1 ,  m_2  \gg 1$ $kL \gg 1$
Spherical	$P_d = -\frac{i}{2} \frac{e^{ikL}}{\sqrt{2\pi kL}} \frac{e^{-i(\pi/4)}}{\sqrt{rr_0}} D(\phi_0; \phi)$	$ m_1 ,  m_2  \gg 1$ $kL \gg 1$
$D(\phi_0; \phi) = \frac{1}{2} \left[ \sec\left(\frac{\phi - \phi_0}{2}\right) + \sec\left(\frac{\phi + \phi_0}{2}\right) \right]$		

waves. For cylindrical and spherical incident waves the expressions in Table I are valid if the length of the shortest source-edge-receiver path ( $L$ ) is more than one wavelength. In all cases the diffracted field is represented by two terms: one associated with the incident field and one associated with the reflected field.

For an important special case the expressions in Table I can be greatly simplified. Let  $w \gg 1$ , where  $w$  is the argument of the Fresnel integrals (the physical meaning is that the receiver is neither close to the edge of the half plane nor close to the shadow boundaries). An asymptotic expansion of  $F(w)$  may be used<sup>3,7,11</sup> to yield the expressions in Table II. Here each diffracted field seems to have a directivity function

$$D(\phi_0; \phi) = \frac{1}{2} \left[ \sec\left(\frac{\phi - \phi_0}{2}\right) + \sec\left(\frac{\phi + \phi_0}{2}\right) \right]. \quad (1)$$

Observe that  $D(\phi_0; \phi)$  is a simple function of just the angular positions of source ( $\phi_0$ ) and receiver ( $\phi$ ). Unlike other directivity functions,  $D(\phi_0; \phi)$  does not depend on the frequency of the incident wave. The following should be noted regarding  $D(\phi_0; \phi)$ : (i) The same directivity function is recovered from solutions developed by totally different methods (for example, the Sommerfeld solution for plane waves and the MacDonald solution for spherical incident waves); (ii) The same directivity function can be recovered from solutions in the time domain as well (for example, the Friedlander solution for plane and cylindrical waves and the BTM solution for spherical waves as shown in Sec. III below); and (iii) The directivity function is the same regardless of the type of the incident wave. It should also be noted that  $D(\phi_0; \phi)$  becomes infinite at the shadow boundaries and so do the solutions in Table II. This is an artifact of the approximations used; the exact solutions yield finite values at the shadow boundaries.

## II. DIRECTIVE LINE SOURCE MODEL

In the present section the Directive Line Source Model (DLSM) is developed. It is based on the approximate forms of the exact analytical solutions presented in the previous section. In Sec. A we present a unified representation of the

diffracted field produced by plane, cylindrical and spherical incident waves. The diffraction problem is transformed into a fully quantified problem of radiation from a directive line source. In Sec. B the transformation is developed further. The diffraction problem becomes a problem of radiation from directive point sources continuously distributed along the edge of the half plane. Conditions for the validity of DLSM are outlined in Sec. C. The results are then modified/generalized for diffraction by pressure release half planes (Sec. D) and wedges (Sec. E). A discussion is then given for half planes having an arbitrary edge profile, e.g., jagged (Sec. F). Finally, Sec. G includes a discussion about the advantages of the DLSM and its relation to other models of diffraction.

### A. The unified representation

In this section a unified representation of the diffracted field produced by plane, cylindrical and spherical incident waves is presented. Furthermore, the problem of diffraction is transformed into a more tractable radiation problem.

Regardless of the type of incident wave, each factor in front of the directivity function  $D(\phi_0; \phi)$  in the solutions shown in Table II contains the following asymptotic form of the Hankel function:

$$H_0^{(1)}(kr) \approx \sqrt{\frac{2}{kr\pi}} e^{ikr-i\pi/4}, \quad \text{for } kr \gg 1. \quad (2)$$

The diffracted field for plane, cylindrical and spherical incident waves can thus be written in the following unified form:

$$P_d = -\frac{i}{4} A_f H_0^{(1)}(k\rho) D(\phi_0; \phi), \quad (3)$$

where both the amplitude factor  $A_f$  and the effective distance  $\rho$  in the argument of the Hankel function depend on the type of incident wave. A summary is given in Table III. Equation (3) and Table III reveal that for all types of incident waves the diffraction solution  $P_d$  is equivalent to radiation from a line source, as indicated by the Hankel function  $H_0^{(1)}(k\rho)$ , but modified by the directivity function  $D(\phi_0; \phi)$ . The line source lies on the edge of the half plane in the case of plane and cylindrical incident waves ( $\rho = r$ ). In the case of spherical incident waves the radiation appears to come from a line source lying distance  $L$  from the receiver ( $\rho = L$ ). Therefore, the line source appears to be at the edge of the half plane, but the receiver is located not at its physical point  $A$  but at distance  $L$  from the edge. Except for the common factor  $-i/4$ , the amplitude of the Hankel function is different for the three cases. For cylindrical waves the amplitude includes an additional Hankel function  $H_0^{(1)}(kr_0)$ , which represents the additional cylindrical spreading and phase shift undergone before the wave reaches the edge of the half plane. For spherical incident waves the amplitude includes the factor  $1/\sqrt{rr_0}$ , which suggests spherical spreading, i.e., the incident wave spreads spherically before the radiation associated with diffraction starts. Note that the distance  $1/\sqrt{rr_0}$  is the apparent distance before line source radiation begins; it is not the literal distance from the source to the edge.

TABLE III. Factors in the frequency domain representation of the diffracted signal [Eq. (3)] for plane, cylindrical and spherical time harmonic incident waves.

Type of incident wave	Diffracted wave $P_d$			Assumptions
Plane	$P_d = -\frac{i}{4}$	$H_0^{(1)}(kr)$	$D(\phi_0; \phi)$	$ \rho_1 ,  \rho_2  \gg 1$ $kr \gg 1$
Cylindrical	$P_d = -\frac{i}{4}$	$H_0^{(1)}(kr_0)$ $H_0^{(1)}(kr)$	$D(\phi_0; \phi)$	$ m_1 ,  m_2  \gg 1$ $kr, kr_0, kL \gg 1$
Spherical	$P_d = -\frac{i}{4}$	$\frac{1}{\sqrt{rr_0}}$ $H_0^{(1)}(kL)$	$D(\phi_0; \phi)$	$ m_1 ,  m_2  \gg 1$ $kL \gg 1$
General expression Eq. (3)	$P_d = -\frac{i}{4}$	$A_f$ $H_0^{(1)}(k\rho)$	$D(\phi_0; \phi)$	

The transformation of the diffraction problem into a more tractable radiation problem can now be fully quantified. The attributes of the new problem are as follows: (a) the radiation comes from a coherent line source located on the edge of the half plane, (b) the receiver in the radiation problem is located at distance  $\rho$  from the line source, where  $\rho$  depends on the type of the incident wave, (c) the amplitude of the line source depends on the type of the incident wave, and (d) for any type of incident wave the directivity of the line source is  $D(\phi_0, \phi)$ .

### B. Radiation from directive point sources along the edge of the half plane

In this section the transformation of the diffraction problem is further developed into a problem of radiation from directive point sources continuously distributed along the edge of the half plane. A time harmonic incident wave is considered first, the impulse response solution is then given, and the result is finally generalized to apply for an arbitrary incident signal.

The Hankel function in Eq. (3) may be expressed by an integral as follows:<sup>12</sup>

$$i\pi H_0^{(1)}(k\rho) = \int_{-\infty}^{\infty} \frac{e^{ik\sqrt{\rho^2+z^2}}}{\sqrt{\rho^2+z^2}} dz = I_h(\rho). \quad (4)$$

The integrand has a very useful interpretation for the transformed radiation problem. It represents the contribution of a point source of infinitesimal length  $dz$  located at  $z$  along the edge of the half plane, and a distance  $\sqrt{\rho^2+z^2}$  away from the receiver. The unified representation of the transformed radiation problem may therefore be written as a line integral over directive point sources along the edge of the half plane:

$$P_d = -\frac{A_f}{4\pi} \int_{\text{edge}} \frac{e^{ik\sqrt{\rho^2+l^2}}}{\sqrt{\rho^2+l^2}} D(\phi_0; \phi) dl. \quad (5)$$

The unified representation of the diffracted field in the frequency domain has its counterpart in the time domain. The expressions in Table IV are the time domain counterparts (obtained via inverse Fourier Transform) of the expressions in Table III and represent the diffracted signal produced by an impulse incident signal. As in the frequency domain, the expressions in Table IV can be written in a unified way as follows:

TABLE IV. Factors in the time domain representation of the diffracted signal [Eq. (6)] for plane, cylindrical and spherical incident impulse waves.  $H$  is the Heaviside function and  $K$  is the complete elliptic integral of the first kind.

Type of incident wave	Diffracted wave $P_d$			
Plane	$p_d = -\frac{1}{4\pi}$		$\frac{2c}{\sqrt{c^2t^2-r^2}} H(ct-r)$	$D(\phi_0; \phi)$
Cylindrical	$p_d = -\frac{1}{4\pi}$		$\frac{4}{\pi} \frac{c}{\sqrt{c^2t^2-r_0^2}} H(ct-r_0) * \frac{c}{\sqrt{c^2t^2-r^2}} H(ct-r)$	$D(\phi_0; \phi)$
	$= -\frac{1}{4\pi}$		$\frac{8c}{\pi} \frac{1}{\sqrt{(ct+r_0-r)(ct-r_0+r)}} K\left(\sqrt{\frac{(ct-r_0-r)(ct+r_0+r)}{(ct+r_0-r)(ct-r_0+r)}}\right)$	$D(\phi_0; \phi)$
Spherical	$p_d = -\frac{1}{4\pi}$	$\frac{1}{\sqrt{rr_0}}$	$\frac{2c}{\sqrt{c^2t^2-L^2}} H(ct-L)$	$D(\phi_0; \phi)$
General expression Eq. (6)	$p_d = -\frac{1}{4\pi}$	$A_t$	$I_\delta(\rho)$	$D(\phi_0; \phi)$

$$p_d = -\frac{A_t}{4\pi} I_\delta(\rho) D(\phi_0; \phi), \quad (6)$$

where  $I_\delta$ , the amplitude factor  $A_t$ , and the distance  $\rho$  are given for each type of incident wave in Table IV. Equation (6) is the impulse response function of our problem.

As in the frequency domain,  $I_\delta$  is related to the contributions coming from point sources along the edge of the half plane. In the case of 1D and 3D propagation the shape of the incident impulse remains unchanged as it travels from the source to the edge of the half plane. The integral  $I_\delta$  represents the contributions from an infinite set of impulsive point sources  $[\delta(t)]$  continuously distributed along the edge

$$\begin{aligned} I_\delta(\rho)^{\text{plane spherical}} &= \int_{-\infty}^{\infty} \frac{\delta\left(t - \frac{\sqrt{\rho^2 + z^2}}{c}\right)}{\sqrt{\rho^2 + z^2}} dz \\ &= \frac{2c}{\sqrt{c^2 t^2 - \rho^2}} H(ct - \rho). \end{aligned} \quad (7)$$

In the case of 2D propagation, however, the shape of the incident impulse changes as it travels from the source to the edge. At the time the incident signal reaches the edge the impulse is equal to  $c/2\pi\sqrt{c^2 t^2 - r_0^2} H(ct - r_0)$ . Thus  $I_\delta$  is the integral over an infinite set of impulsive point sources along the half plane, each one emitting the signal  $c/2\pi\sqrt{c^2 t^2 - r_0^2} H(ct - r_0)$  instead of  $\delta(t)$ . The contribution from the point sources is therefore represented by the convolution integral in Table IV. More specifically, it can be shown that the convolution integral in Table IV is equal to the summation of the contributions of an infinite set of point sources along the edge, each emitting the signal  $c/2\pi\sqrt{c^2 t^2 - r_0^2} H(ct - r_0)$

$$\begin{aligned} I_\delta(\rho)^{\text{cylindrical}} &= \frac{4}{\pi} \int_{-\infty}^{+\infty} \frac{1}{2\pi \sqrt{\left(t - \frac{\sqrt{\rho^2 + z^2}}{c}\right)^2 - \frac{r_0^2}{c^2}}} \frac{1}{\sqrt{\rho^2 + z^2}} dz. \end{aligned} \quad (8)$$

Finally, the diffracted signal caused by diffraction of a plane, cylindrical or spherical arbitrary source signal  $f(t)$  can be found by the convolution integral of  $f(t)$  with the impulse response [Eq. (6)]. This convolution integral can also be expressed in terms of a line integral along the edge of the half plane as follows:

$$p_d = -\frac{A_t}{4\pi} \int_{\text{edge}} \frac{\bar{f}\left(t - \frac{\sqrt{\rho^2 + l^2}}{c}\right)}{\sqrt{\rho^2 + l^2}} D(\phi_0; \phi) dl, \quad (9)$$

where  $\bar{f}(t)$  is the incident signal  $f(t)$  at the source for plane and spherical incident waves or the incident signal at the edge of the half plane for cylindrical incident waves.

### C. Region of validity of the DLSM

The DLSM is based on an interpretation of the unified representation of the diffracted field, where assumptions

about the position of source and receiver have been made. The assumptions, indicated in the last column of Table III, are of three types: (i) one that allows expression of the exact solution in terms of Fresnel integrals (no restriction needed for plane waves,  $kL \gg 1$  for cylindrical and spherical waves), (ii) one that allows use of the asymptotic form of the Fresnel integrals, ( $|\rho_{1,2}| \gg 1$  for plane waves,  $|m_{1,2}| \gg 1$  for cylindrical and spherical waves), and (iii) one that allows use of the asymptotic form of the Hankel function ( $kr \gg 1$  for plane waves,  $kr$  and  $kr_0 \gg 1$  for cylindrical waves and no additional restriction for spherical waves). All the assumptions have the same physical meaning: The receiver must be away from the shadow boundaries  $B_i$  and  $B_r$  (at  $\phi = \pi + \phi_0$  and  $\phi = \pi - \phi_0$ , respectively), and both source and receiver should be more than one wavelength away from the edge of the half plane. Non time harmonic waves can also be treated by DLSM [see Eqs. (6) and (9)]. In this case the proximity to the edge is not measured in wavelengths; other characteristic distances are used instead (for example, the spatial duration of a transient source signal, or the correlation length of a noise source signal).

### D. Pressure release half planes

The DLSM may also be applied for a pressure release half plane. The only difference in the results is in the directivity function. For a pressure release half plane the directivity function<sup>1</sup> is

$$D(\phi_0; \phi) = \frac{1}{2} \left[ \sec\left(\frac{\phi - \phi_0}{2}\right) - \sec\left(\frac{\phi + \phi_0}{2}\right) \right]. \quad (10)$$

In general, the diffracted field  $p_d$  (for any type of incident wave time harmonic or non time harmonic) is represented by two terms, one associated with the incident field ( $p_{di}$ ) and one associated with the reflected field ( $p_{dr}$ ) (see for example Table I for time harmonic waves). The first term in the directivity function, which includes the difference of the angles  $\phi - \phi_0$ , is associated with the incident field. The second term, which includes the sum  $\phi + \phi_0$ , is associated with the reflected field. If the half plane is rigid,  $p_d = p_{di} + p_{dr}$  and the directivity is described by Eq. (1). If the half plane is pressure release,  $p_d = p_{di} - p_{dr}$ , which leads to Eq. (10).

### E. Wedges

Diffraction by a wedge may also be treated by DLSM. The unified representation of the diffracted field produced by plane, cylindrical and spherical incident waves on a rigid wedge is the same as presented in Secs. A and B. Again only the directivity function is different; it is<sup>1</sup>

$$\begin{aligned} D_\nu(\phi_0; \phi) &= \frac{1}{\nu} \sin \frac{\pi}{\nu} \left[ \left( \cos \frac{\phi - \phi_0}{\nu} - \cos \frac{\pi}{\nu} \right)^{-1} \right. \\ &\quad \left. - \left( \cos \frac{2\pi - \phi - \phi_0}{\nu} + \cos \frac{\pi}{\nu} \right)^{-1} \right], \end{aligned} \quad (11)$$

where the  $\nu\pi = 2\pi - 2\Omega$  and  $2\Omega$  is the angle of the wedge. To obtain the half plane result, take the limit as the wedge angle becomes zero, in which case  $\nu = 2$ . Equation (11) then reduces to Eq. (1). If the wedge is made from pressure re-

lease material, the directivity function is the same as in Eq. (11) but with a plus sign instead of the minus sign between the two reciprocal terms.

## F. Jagged edges

Perhaps the most important characteristic of the DLSM is that it provides a flexible model for diffraction by arbitrarily shaped edges, as for example by jagged or piecewise smooth edges. DLSM transforms the diffraction problem into radiation from point sources along the edge of the half plane. It seems reasonable to assume that we can transform the problem of diffraction by a jagged edge half plane into radiation of point sources along the jagged edge of the half plane. The model provides insight into the physical phenomenon without considerable increase in the complexity of the calculations. It has been shown that DLSM can be applied successfully to predict the diffracted field behind a barrier having a jagged top edge.<sup>13,14</sup> The results are important in a new design of noise barriers.

## G. Comparison of DLSM to existing methods for diffraction

In this section the advantages of DLSM are outlined and its relation to existing methods of diffraction is examined.

The DLSM appears to be a promising tool for the study of diffraction by half planes. Application of DLSM is fast, simple and intuitive. It can handle plane, cylindrical and spherical incident waves, while the exact solutions have been developed for a specific type of incident wave. Furthermore, as shown in Sec. III, the method can also handle directional sources. Finally, DLSM seems to be a promising tool for the study of diffraction by edges of arbitrary profiles.

The calculations involved in DLSM are much simpler than the ones contained in the exact theory. For example, the integral in Eq. (9) can be evaluated analytically for simple waveforms  $f(t)$  such as square pulses or N waves. Although, in principle, knowledge of the impulse response function [as for example the Friedlander solution<sup>8</sup> or the Biot–Tolstoy–Medwin (BTM) solution<sup>10</sup>] is sufficient to compute the diffracted field by an arbitrary incident wave  $f(t)$ , the numerical evaluation of the convolution between  $f(t)$  and the impulse response function is usually cumbersome. Among the considerations for the numerical evaluation of the convolution integral are truncation of the very slowly decaying tail in the impulse response function, numerical treatment of the singularity (infinite value) at the front of the impulse response function, and the “wrap around effect” when a circular convolution algorithm is used.

The idea of representing the diffracted field as a line integral along the edge of the diffracting object is by no means new. Line integral representations began to appear at the beginning of the century and they have been used in several fields including optics, acoustics and electromagnetics.

The first line integral representations were based on the Kirchhoff theory of diffraction.<sup>15</sup> Rubinowitz (1924) was the first to reduce the surface integrals emerging in the Kirchhoff theory to a line integral along the edge of the diffracting

object, where the diffracting object was an aperture in a plane screen. Keller *et al.*<sup>16</sup> evaluated explicitly this line integral in the high frequency limit. Years later, Embleton<sup>17</sup> presented the Kirchhoff solution for diffraction by a half plane as an integral along the edge of the half plane. Compared to all of the abovementioned line integral representations, DLSM has the advantage that incorporates fewer restrictions. Solutions based on the Kirchhoff theory are valid for receiver locations close to the shadow boundaries  $B_i$  and  $B_r$ , and for high frequencies. Not subject to any of these restrictions, DLSM is valid in almost the entire field including points deep in the shadow zone, which is of major interest.

Line integral representations that do not have the restrictions of the Kirchhoff theory can also be found in the literature. Compared to them, DLSM has the advantage that it addresses equally both the time and frequency domain, as well as all three types of incident radiation. One of the most widely used methods in electromagnetics applications is the method of equivalent edge currents.<sup>18,19</sup> Equivalent edge currents are the fictitious currents that must be placed on the edge of the half plane to produce the known diffracted field. The edge contributions are therefore represented as a line integral along the edge of the half plane. The method is restricted to the frequency domain and to plane incident waves. DLSM can be thought of as an acoustical analog of the equivalent edge currents that addresses equally both the frequency and time domain, as well as all three types of incident radiation. Closely related to the method of equivalent edge currents is the method of incremental length diffraction coefficients, which was introduced by Mitzner<sup>20,21</sup> (see also Ref. 19 and references therein). The incremental length diffraction coefficients are coefficients that must be applied to each infinitesimal element of the edge (modeled as a point source) to generate the diffracted field that must be added to the Kirchhoff solution to produce the entire diffracted field. The diffraction coefficients, unlike DLSM, have been derived for time harmonic plane incident waves. Gautesen<sup>22</sup> expressed (in the high frequency limit) the sound field in the vicinity of the edge of an obstacle as a line integral along the diffracting edge. Because DLSM is invalid in the vicinity of the edge, the Gautesen representation and DLSM complement each other. Recently, Svensson *et al.*<sup>23</sup> expressed the BTM solution as a line integral along the edge of the half plane. Their formulation is very much like DLSM having the additional advantage of being valid for all receiver locations. However, because it is based on the BTM solution, it is applicable only to spherically spreading incident impulses.

Finally, DLSM is compared with other unified representations. The most often cited unified interpretation of the diffracted field can be found in the work of Keller<sup>24–26</sup> and Pierce.<sup>27,28</sup> Both authors interpret the diffracted wave, regardless of the type of incident wave, as a ray scattered from the edge of the half plane. Their starting point is the asymptotic form of the exact solution for plane incident waves (Keller), and for spherical incident waves (Pierce) (see Table II). Although the unified representation presented here and those of Keller and Pierce are based on the same

equations, they provide different interpretations of diffraction. In the present work the diffracted field is interpreted as sound radiated from the edge of the half plane (not as a diffracted ray). This interpretation has two advantages: (i) No high frequency assumption is required. (ii) We transform the diffraction problem into radiation by point sources along the edge of the half plane. By plausible extension, (ii) allows us to treat diffraction by arbitrarily shaped edges, for example, jagged or piecewise smooth edges without the implication of infinite field values at the caustics regions.

In conclusion, the work presented here synthesizes two components: (i) the line source radiation from the edge of the half plane and (ii) the unified representation for diffraction by time harmonic and non time harmonic plane, cylindrical and spherical incident waves. Previous works address these components separately, while the present work combines the two components in a simple and intuitive model.

Finally, it should be noted that the abovementioned references are a selection of methods related to DLSM. They do not constitute an exhaustive list of methods for studying diffraction. For example, infinite series solutions<sup>1,29,30</sup> or the hybrid method of progressing and oscillatory waves<sup>31</sup> have been omitted.

### III. RESULTS

In the present section results from the DLSM are presented and compared with known analytical solutions for omnidirectional spherical and cylindrical waves, plane waves, and waves from directional sources.

#### A. Omnidirectional spherical waves

In this section results obtained by DLSM for spherically spreading incident waves are compared with the BTM solution,<sup>9,10</sup> which is an exact analytical solution for an impulsive point source. Comparison with the BTM solution is important for two reasons. First, Biot and Tolstoy's method of solution is completely different from the method used to obtain the equations in Table II and DLSM. Good agreement between the DLSM and BTM solutions indicates that DLSM does not depend on the derivation method. Second, the BTM interpretation for diffraction by a half plane is different from that presented here. The comparison also shows that DLSM is an accurate alternative interpretation that requires far fewer calculations.

The analytical comparison between the two methods is presented first for an impulsive point source. Then numerical comparisons are presented for an N-shaped source signal. It is shown that agreement between DLSM and BTM becomes better as source and receiver move away from the edge of the half plane and away from the shadow boundaries. The case of oblique incidence is considered next, and finally the comparison of the two methods is shown for receiver locations in front of and behind the half plane.

The BTM solution for an impulsive point source located at  $(r_0, \phi_0, z_0)$  is as follows:<sup>9,10</sup>

$$p_d^{\text{BTM}} = -\frac{c}{2\pi} e^{-y/2} \frac{1}{rr_0 \sinh y} M, \quad ct > L, \quad (12)$$

where the parameter  $y$  is defined by

$$y = \cosh^{-1} F, \quad (13)$$

$$F = \frac{c^2 t^2 - [r^2 + r_0^2 + (z - z_0)^2]}{2rr_0} = \frac{c^2 t^2 - L^2 + 2rr_0}{2rr_0},$$

and  $M$  is given by

$$M = \frac{\sin\left[\frac{\pi \pm \phi \pm \phi_0}{2}\right]}{1 - 2e^{-y/2} \cos\left[\frac{\pi \pm \phi \pm \phi_0}{2}\right] + e^{-y}}. \quad (14)$$

BTM used the term  $\sin(a \pm b \pm c)$  to represent the sum of four terms:  $\sin(a+b+c)$ ,  $\sin(a+b-c)$ ,  $\sin(a-b+c)$ , and  $\sin(a-b-c)$ . With appropriate rearrangement, Eq. (14) becomes

$$M = \frac{2(1 + e^{-y}) \cos\frac{\phi \pm \phi_0}{2}}{(1 + e^{-y})^2 + 4e^{-y} \left[ \cos^2\frac{\phi \pm \phi_0}{2} - 1 \right]}, \quad (15)$$

where the familiar term  $\cos((\phi \pm \phi_0)/2) = \cos((\phi - \phi_0)/2) + \cos((\phi + \phi_0)/2)$  appears. The front of the diffracted signal arrives at the receiver location  $(r, \phi, z)$  at time given by  $ct = L$ . Just behind the front, the time  $t$  can be written as  $t = \tau_0 + \tau$  with  $c\tau_0 = L$  and  $\tau \ll \tau_0$ . With that assumption the parameter  $F$  becomes

$$F = \frac{2rr_0 + 2c^2\tau_0\tau + c^2\tau^2}{2rr_0} \approx 1 + \frac{c^2\tau_0\tau}{rr_0}, \quad (16)$$

which in turn yields

$$\sinh y = \sqrt{\left(1 + \frac{c^2\tau_0\tau}{rr_0}\right)^2 - 1} \approx \sqrt{\frac{2c^2\tau_0\tau}{rr_0}}. \quad (17)$$

The parameter  $M$  just behind the front, where  $e^{-y} \rightarrow 1$ , becomes

$$M = \frac{2 \cdot 2 \cos\frac{\phi \pm \phi_0}{2}}{2^2 + 4 \left[ \cos^2\frac{\phi \pm \phi_0}{2} - 1 \right]} = \frac{1}{\cos\frac{\phi \pm \phi_0}{2}} = D(\phi; \phi_0). \quad (18)$$

The diffracted pulse  $p_d$  [Eq. (12)] thus becomes

$$p_d^{\text{BTM}} \approx -\frac{1}{2\pi} \frac{1}{\sqrt{2\tau_0\tau}} \frac{1}{\sqrt{rr_0}} D(\phi_0; \phi). \quad (19)$$

Consider now the impulse response solution developed by DLSM for a spherically spreading incident wave:

$$p_d^{\text{DLSM}} = -\frac{1}{4\pi} \frac{1}{\sqrt{rr_0}} I_\delta D(\phi_0; \phi), \quad (20)$$

where the integral  $I_\delta$  represents the integral along the edge of the half plane assuming that each point along the edge radiates a Dirac function  $\delta(t)$ ,

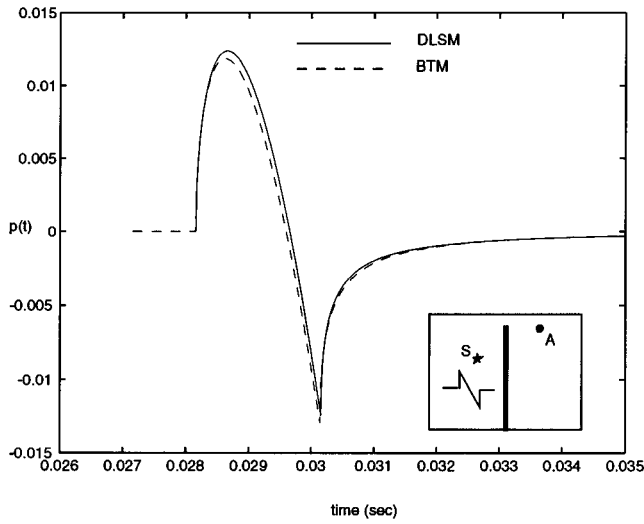


FIG. 2. Comparison of DLSM and BTM results for diffraction of a spherically spreading N wave by a rigid half plane. Source and receiver in a plane normal to the half plane ( $z=z_0=0$ ). The N wave has unity amplitude and duration  $2T=2$  ms. Point source  $S$  located at  $(r_0, \phi_0)=(5 \text{ m}, 30^\circ)$ ; receiver  $A$  at  $(r, \phi)=(5 \text{ m}, 270^\circ)$ .

$$I_\delta = \int_{-\infty}^{\infty} \frac{\delta\left(t - \frac{\sqrt{L^2 + z^2}}{c}\right)}{\sqrt{L^2 + z^2}} dz. \quad (21)$$

Just behind the front the time  $t$  can be written as  $t = \tau_0 + \tau$ , where  $\tau \ll \tau_0$  and  $c\tau_0 = L$ . Equation (21) becomes

$$I_\delta = \int_{-\infty}^{\infty} \frac{\delta\left(t - \frac{\sqrt{L^2 + z^2}}{c}\right)}{\sqrt{L^2 + z^2}} dz = \frac{2c}{\sqrt{c^2 t^2 - L^2}} \approx \frac{2}{\sqrt{2\tau\tau_0}}, \quad (22)$$

and therefore the diffracted field obtained by DLSM becomes:

$$p_d^{\text{DLSM}} \approx -\frac{1}{2\pi} \frac{1}{\sqrt{rr_0}} \frac{1}{\sqrt{2\tau\tau_0}} D(\phi_0; \phi), \quad (23)$$

which is the asymptotic form of the Biot–Tolstoy–Medwin exact solution [Eq. (19)].

The good agreement between DLSM and the exact analytical solution is shown in Fig. 2 for the case of an N wave. The numerical results for the DLSM (solid line) have been obtained by evaluating Eq. (9). If  $f(t)$  in Eq. (9) is an N wave, the integration can be performed analytically and  $p_d$  is given in a closed form expression. Evaluating the closed form expression is much simpler than actually performing the numerical integration. The exact solution (dashed line) has been computed by numerically convolving the BTM solution with the incident N wave. The evaluation of the exact solution involved far more calculations than the DLSM solution (see remarks in Sec. II G). The source is located at  $(r_0, \phi_0, z_0) = (5 \text{ m}, 30^\circ, 0 \text{ m})$ . The N wave has unity amplitude and time duration  $2T = 2$  ms. The spatial duration of the N wave is  $sd = 2cT = 0.646 \text{ m}$ . The receiver is located at  $(r, \phi, z) = (5 \text{ m}, 270^\circ, 0 \text{ m})$ , within the region of validity of

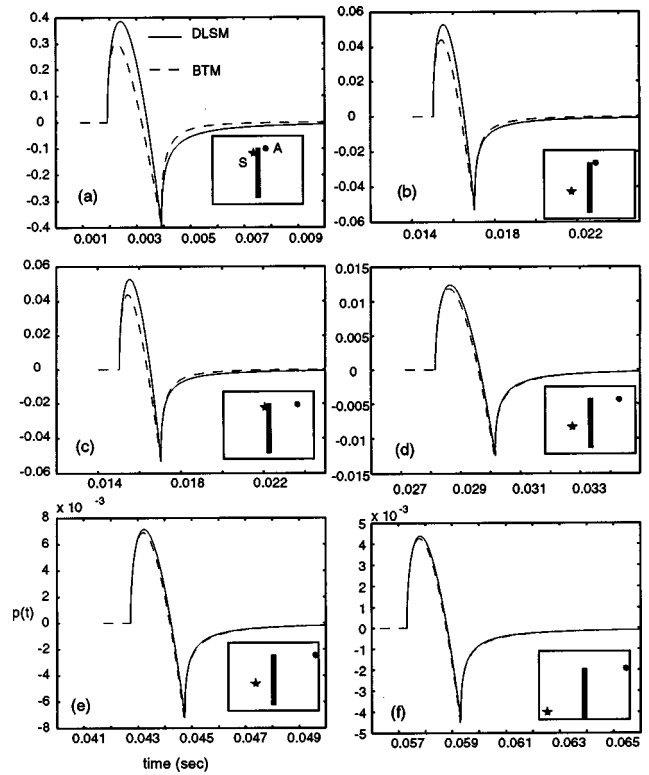


FIG. 3. Comparison between DLSM and BTM results for a spherically spreading N wave, as a function of source/receiver proximity to the edge. DLSM prediction becomes better as proximity distance increases. Source and receiver in a plane normal to the half plane in all cases ( $z_0 = z = 0$ ). Point source  $S$  at  $(r_0, \phi_0, z_0) = (r_0, 30^\circ, 0)$  and receiver  $A$  at  $(r, \phi, z) = (r, 270^\circ, 0)$ , where  $r = 0.5 \text{ m}$ ,  $r_0 = 0.5 \text{ m}$  [(a)];  $r = 0.5 \text{ m}$ ,  $r_0 = 5 \text{ m}$  [(b)];  $r = 5 \text{ m}$ ,  $r_0 = 0.5 \text{ m}$  [(c)];  $r = 5 \text{ m}$ ,  $r_0 = 5 \text{ m}$  [(d)];  $r = 10 \text{ m}$ ,  $r_0 = 5 \text{ m}$  [(e)]; and  $r = 10 \text{ m}$ ,  $r_0 = 10 \text{ m}$  [(f)].

DLSM, i.e., away from the edge of the half plane ( $L = 10 \text{ m} \gg sd$ ) and away from the shadow boundary  $B_i$  ( $\phi_{B_i} = 210^\circ \ll \phi = 270^\circ$ ).

DLSM has been developed under the assumption that source and receiver are not in the proximity of the edge. In Fig. 3 the time waveforms predicted by DLSM are compared with the BTM solution for various source and receiver locations. It can be observed that the agreement between DLSM and the exact analytical solution becomes better as source and receiver move away from the edge.

The second assumption for the validity of DLSM is that the receiver not to be close to the shadow boundary. Figure 4 shows that indeed the model gives accurate predictions for receiver locations deep into the shadow zone. Agreement between DLSM and the exact analytical solution is observed to be better as the receiver moves away from the shadow boundary.

Next consider oblique incidence, that is,  $z \neq z_0 = 0$ . Figure 5 shows that agreement between the DLSM and the exact analytical solution is also very good in this case. It should be noted that the DLSM is easily applied when the incidence is oblique rather than normal. The only change is in the arithmetic value of the shortest source-edge-receiver path  $L$ .

The use of DLSM is not limited to receiver positions in the shadow zone. Positions in front of and above the half plane may be treated as well. Comparisons for receiver po-

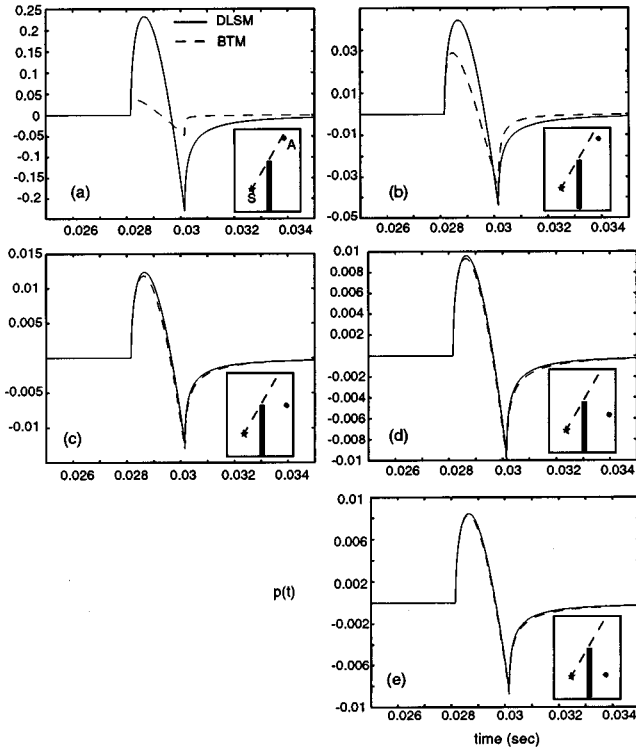


FIG. 4. Comparison between DLSM and BTM results for a spherically spreading N wave, as a function of receiver proximity to the shadow boundary (dashed line in inset). DLSM prediction becomes better as receiver moves away from the shadow boundary. Source and receiver in a plane normal to the half plane for all cases ( $z_0 = z = 0$ ). Point source  $S$  at  $(r_0, \phi_0, z_0) = (5 \text{ m}, 30^\circ, 0)$ ; receiver  $A$  at  $(r, \phi, z) = (5 \text{ m}, \phi, 0)$ , where  $\phi = 212^\circ$  [(a)];  $\phi = 222^\circ$  [(b)];  $\phi = 270^\circ$  [(c)];  $\phi = 300^\circ$  [(d)]; and  $\phi = 330^\circ$  [(e)]; shadow boundary at  $\phi = 210^\circ$ .

sitions in all quadrants are presented in Fig. 6. Agreement with the exact analytical solution is good. Note the sign change of the time waveform as the receiver moves across the boundary  $B_i$ , the continuation of the half plane in the free space above it, and across the boundary  $B_r$ .

### B. Omnidirectional cylindrical waves

The results obtained by DLSM are now compared with known exact analytical solutions for cylindrical incident waves in what follows. The analytical comparison is presented first for the impulse response solution. A comparison of numerical results for a cylindrically spreading N wave is then given.

Consider Friedlander's impulse response solution for a cylindrically spreading wave incident on a rigid half plane:<sup>8</sup>

$$p_d^{\text{Friedlander}} = -\frac{c}{\pi} \frac{1}{\sqrt{c^2 t^2 - R_1^2}} - \frac{c}{\pi} \frac{1}{\sqrt{c^2 t^2 - R_2^2}} \quad ct > L, \quad (24)$$

where  $R_1$ ,  $R_2$  and  $L$  are defined in Fig. 1 [ $R_1^2 = r^2 + r_0^2 - 2rr_0 \cos(\phi - \phi_0)$ ,  $R_2^2 = r^2 + r_0^2 - 2rr_0 \cos(\phi + \phi_0)$ ,  $L = r + r_0$ ]. The front of the diffracted signal arrives at a point  $(r, \phi)$  at time given by  $ct = L$ . At that time the parameter  $G = (c^2 t^2 - r^2 - r_0^2) / (2rr_0)$  obtains its minimum value  $G = 1$ . Just behind the front,  $ct - L$  is small and therefore  $G$

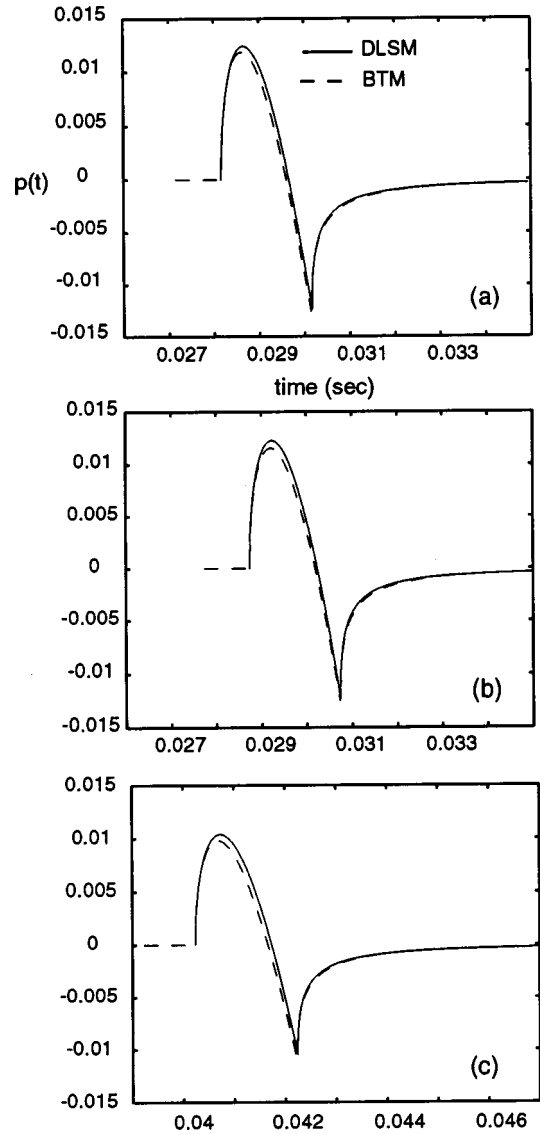


FIG. 5. Comparison between DLSM and BTM results for normal [source and receiver in a plane normal to the half plane ( $z_0 = z = 0$ ) shown in (a)] and oblique incidence [shown in (b) and (c)]. Point source at  $(r_0, \phi_0, z_0) = (5 \text{ m}, 30^\circ, 0)$ ; receivers at  $(r, \phi, z) = (5 \text{ m}, 270^\circ, z)$ , where  $z = 0 \text{ m}$  [(a)];  $z = 2 \text{ m}$  [(b)]; and  $z = 10 \text{ m}$  [(c)].

$-1$  is small as well. Thus the first term of the diffracted signal in Eq. (24) becomes

$$\begin{aligned} E^* &= \frac{c}{\pi} \frac{1}{\sqrt{c^2 t^2 - r^2 - r_0^2 + 2rr_0 \cos(\phi - \phi_0)}} \\ &= \frac{c}{\pi \sqrt{2rr_0}} \frac{1}{\sqrt{G + \cos(\phi - \phi_0)}} \\ &= \frac{c}{\pi \sqrt{2rr_0}} \frac{1}{\sqrt{G + 2 \cos^2\left(\frac{\phi - \phi_0}{2}\right) - 1}} \\ &\approx \frac{c}{2\pi \sqrt{rr_0}} \frac{1}{\cos\left(\frac{\phi - \phi_0}{2}\right)}. \end{aligned} \quad (25)$$



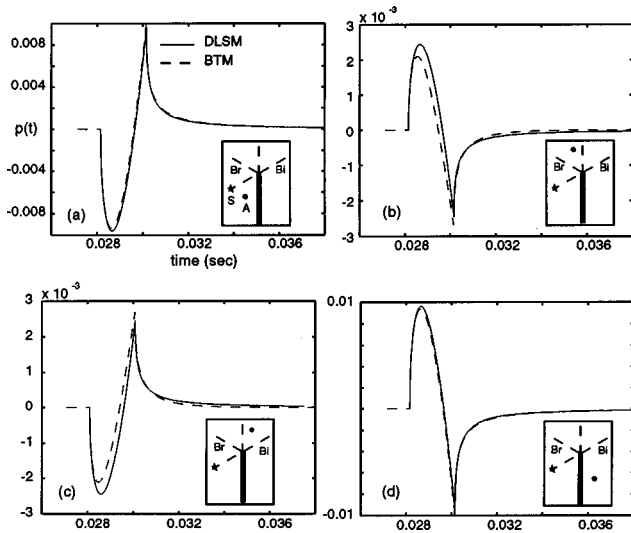


FIG. 6. Comparison between DLSM and BTM results for receiver positions in front of, above and behind the half plane. Source and receiver in a plane normal to the half plane. Point source at  $(r_0, \phi_0, z_0) = (5 \text{ m}, 60^\circ, 0)$ ; receivers at  $(r, \phi, z) = (5 \text{ m}, \phi, 0)$ , where  $\phi = 30^\circ$  [(a)];  $\phi = 170^\circ$  [(b)];  $\phi = 190^\circ$  [(c)]; and  $\phi = 330^\circ$  [(d)].

The second term of the diffracted signal yields a similar asymptotic form and thus the diffracted signal becomes

$$p_d^{\text{Friedlander}} \approx \frac{c}{2\pi\sqrt{rr_0}} D(\phi_0; \phi). \quad (26)$$

Consider now the impulse response solution developed by DLSM for a cylindrically spreading incident wave:

$$p_d^{\text{DLSM}} = -\frac{D(\phi_0; \phi)}{4\pi} \frac{8c}{\pi} \frac{1}{\sqrt{(ct+r_0-r)(ct-r_0+r)}} \times K\left(\sqrt{\frac{(ct-r_0-r)(ct+r_0+r)}{(ct+r_0-r)(ct-r_0+r)}}\right), \quad (27)$$

where  $K$  is the complete elliptic integral of the first kind. At the front of the diffracted signal  $ct = r + r_0$  and thus the expression in front of  $K$  in Eq. (27) becomes

$$-\frac{D(\phi_0; \phi)}{4\pi} \frac{8c}{\pi} \frac{1}{2\sqrt{r_0 r}}. \quad (28)$$

The modulus of the elliptic integral becomes zero and therefore

$$K(0) = \frac{\pi}{2}. \quad (29)$$

Substituting Eqs. (28) and (29) into the DLSM solution [Eq. (27)] we obtain

$$p_d^{\text{DLSM}} \approx \frac{c}{2\pi\sqrt{rr_0}} D(\phi_0; \phi), \quad (30)$$

which is the approximate form of Friedlander's exact solution [Eq. (26)].

The good agreement between DLSM and the exact analytical solution is shown in Fig. 7 for the case of an N wave. The source is located at  $(r_0, \phi_0) = (10 \text{ m}, 30^\circ)$ . The N wave

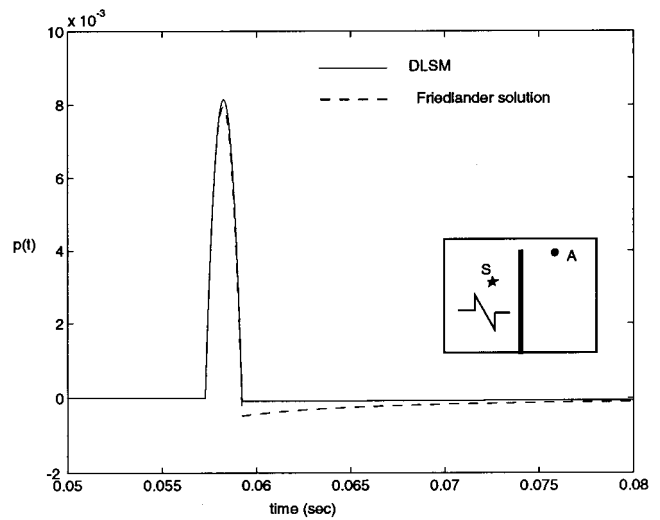


FIG. 7. Comparison between DLSM and Friedlander solution for diffraction of a cylindrically spreading N wave by a rigid half plane. The N wave has unity amplitude and duration  $2T = 2 \text{ ms}$ . Line source  $S$  at  $(r_0, \phi_0) = (10 \text{ m}, 30^\circ)$ ; receiver  $A$  at  $(r, \phi) = (5 \text{ m}, 270^\circ)$ .

is of unity amplitude and time duration  $2T = 2 \text{ ms}$ . The spatial duration of the pulse is  $sd = 0.646 \text{ m}$ . The receiver is located at  $(r, \phi) = (10 \text{ m}, 270^\circ)$ , within the region of validity of DLSM, i.e., away from the edge of the half plane ( $r \gg sd$ ) and away from the shadow boundary  $B_i$  ( $\phi_{B_i} \ll \phi$ ). Note that the shape of the diffracted signal is substantially different from that in Fig. 2. In our case the incident signal is cylindrically spreading and its shape changes as it propagates from the source to the edge. The spherically spreading incident signal in Fig. 2 on the other hand does not undergo this change of shape.

### C. Plane incident waves

The results obtained by DLSM are compared with known exact analytical solutions for plane incident waves in this section. The analytical comparison is presented first for a plane unit step function. A comparison of numerical results for a plane square pulse is then given.

Friedlander's solution for a plane unit step function is<sup>8</sup>

$$p_d^{\text{Friedlander}} = -\frac{1}{\pi} \tan^{-1} \left( \sec \frac{\phi - \phi_0}{2} \sqrt{\frac{ct-r}{2r}} \right) - \frac{1}{\pi} \tan^{-1} \left( \sec \frac{\phi + \phi_0}{2} \sqrt{\frac{ct-r}{2r}} \right), \quad (31)$$

where the solution holds for  $ct > r$ , i.e., after the front of the diffracted wave arrives at the receiver. Immediately behind the front,  $ct - r$  is small and  $(ct - r)/r$  is also small, for which case Eq. (31) becomes

$$p_d^{\text{Friedlander}} \approx -\frac{1}{\pi} \sqrt{\frac{ct-r}{2r}} D(\phi_0; \phi). \quad (32)$$

If the front of the diffracted wave arrives at the receiver at  $\tau_0 = r/c$ , immediately behind the front we have  $ct = c\tau_0 + c\tau$ , where  $\tau \ll \tau_0$ . Equation (32) then becomes

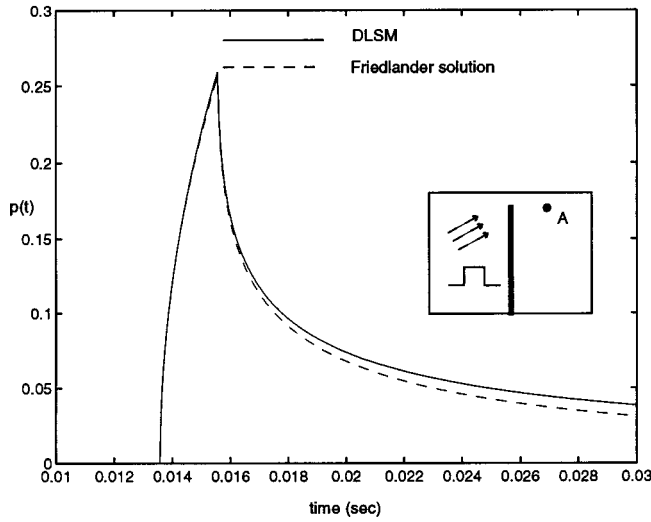


FIG. 8. Comparison between DLSM and Friedlander solution for diffraction of a plane square pulse by a rigid half plane. Square pulse of unity amplitude, duration  $2T=2$  ms, incident at angle  $\phi_0=30^\circ$ ; receiver A located at  $(r, \phi)=(5 \text{ m}, 270^\circ)$ .

$$p_d^{\text{Friedlander}} \approx -\frac{1}{\pi} \sqrt{\frac{\tau}{2\tau_0}} D(\phi_0; \phi). \quad (33)$$

Consider now the DLSM for a plane unit step function:

$$p_d^{\text{DLSM}} = -\frac{I_{\text{us}}(r)}{4\pi} D(\phi_0; \phi), \quad (34)$$

where  $I_{\text{us}}(r)$  is the integral along the edge of the half plane assuming that each point along the edge radiates a unit step wave. The integration can be performed analytically to yield

$$I_{\text{us}}(r) = 2 \ln \left[ \frac{\sqrt{c^2 t^2 - r^2} + ct}{r} \right], \quad ct > r. \quad (35)$$

Immediately behind the front of diffracted signal, the argument of the logarithm is a small number and therefore Eq. (35) becomes

$$I_{\text{us}}(r) = 2 \left[ \frac{\sqrt{c^2 t^2 - r^2} + ct}{r} - 1 \right] = 2 \left[ \frac{\sqrt{c^2 t^2 - r^2} + ct - r}{r} \right]. \quad (36)$$

Immediately behind the front  $ct-r$  is a small number and can be ignored compared to  $\sqrt{c^2 t^2 - r^2}$  in the numerator. By substituting for  $ct=c\tau_0+c\tau$ ,  $c\tau_0=r$  and assuming  $\tau \ll \tau_0$ , we obtain

$$I_{\text{us}}(r) \approx 2\sqrt{2} \sqrt{\frac{\tau}{\tau_0}}. \quad (37)$$

Substituting Eq. (37) into the DLSM solution [Eq. (34)] we obtain

$$p_d^{\text{DLSM}} \approx -\frac{1}{\pi} \sqrt{\frac{\tau}{2\tau_0}} D(\phi_0; \phi), \quad (38)$$

which agrees with Eq. (33), the asymptotic form of Friedlander's exact solution.

The results from DLSM and the exact analytical solution are compared in Fig. 8 for a plane square pulse. The angle of incidence is  $\phi_0=30^\circ$ . The square pulse has unity amplitude

and time duration  $2T=2$  ms. The spatial duration of the pulse is  $sd=0.646$  m. The receiver is located at  $(r, \phi)=(5 \text{ m}, 270^\circ)$ , within the region of validity of DLSM, i.e., away from the edge of the half plane ( $r \gg sd$ ) and away from the shadow boundary  $B_i$  ( $\phi_{B_i} \ll \phi$ ).

#### D. Directional sources: Array of line and point sources

The DLSM can be applied for omnidirectional as well as for directional sources like pistons or arrays of simple sources. One advantage of DLSM is that directional sources can be handled without increase in the complexity of the calculations.

Consider, for example,  $n$  time harmonic point sources forming an array of equally spaced elements. If the array is located several wavelengths away from the edge, the array is considered as a directional point source having directivity

$$\Theta(\theta) = \frac{\sin\left(n \frac{kd}{2} \sin \theta\right)}{n \sin\left(\frac{kd}{2} \sin \theta\right)}, \quad (39)$$

where  $d$  is the distance between the elements of the array,  $n$  is the number of the elements, and  $\theta$  is the angle measured from the axis of the array [see Fig. 9(a)]. The directivity pattern of an array of three elements is shown in Fig. 9(a) by dashed lines for  $d=0.1$  m and  $f=10$  kHz ( $kd=9.12$ ). The diffracted signal is computed according to DLSM as the contribution of directive point sources continuously distributed along the edge of the half plane. The signal emitted from the point sources on the edge must be multiplied by the arithmetic value of the directivity function of the array [Eq. (39)] at point  $N[\Theta(\theta_N)]$ , where  $N$  is the point on the edge where the least diffraction path intersects the edge (see Fig. 1). The results are compared with a superposition of the diffracted fields produced by the elements taken simply and calculated by the spherical wave formula given in Table II. Figure 9 shows the comparison; it is seen to be very good.

The comparisons are repeated for an array of three time harmonic line sources oriented as shown in Fig. 10. The space between the elements is the same as in the previous example but the frequency is half ( $kd=4.6$ ). As a result the directivity pattern of the array has fewer nulls. It can be observed that the agreement in this case is also good.

The same good agreement between DLSM and analytical solutions has been observed for a wide range of frequencies and different orientations of the array.

#### E. Comparison with experimental data

Results obtained by DLSM are now compared with measurements carried out by Rosenberg.<sup>32</sup> The source was an electric spark, which produced short N-shaped pulses, approximately  $50 \mu\text{s}$  in duration. The half plane was a sheet of plywood with a metal strip at the top to serve as the edge. Figure 11 shows the arrangement. The pulse was short enough that diffracted signals from the ends of the plywood sheet did not interfere diffracted signals from the top edge;

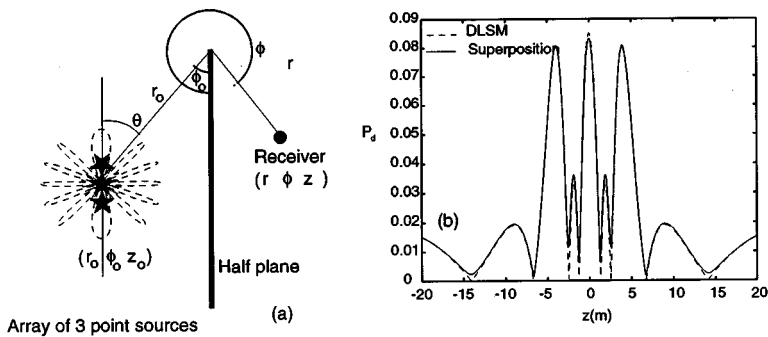
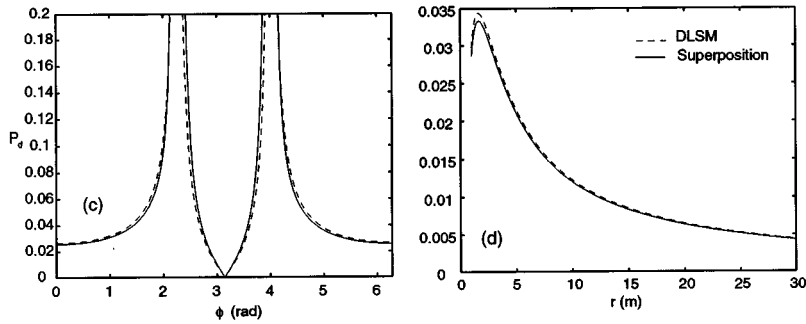


FIG. 9. Comparison between DLSM and other analytical solutions for diffraction of an array of three time harmonic point sources on a rigid half plane for a wide range of receiver locations. The distance  $d$  between the elements of the array is 0.1 m and the frequency is 10 kHz. In the application of the DLSM, the array is considered as one directive point source with directivity pattern shown with dashed lines in (a).



the plywood and metal top therefore approximate a half plane. The spark source was located at a fixed height,  $H_s = 60$  cm, above the table, and at a fixed distance,  $D_s = 60$  cm, from the screen. The edge of the half plane was 94 cm above the table surface. The microphone was placed in various positions specified by the height  $H$  from the table, the distance  $D$  from the half plane, and the distance  $Z$  from the source in the direction along the edge (see Fig. 11). A set of experiments was conducted for many receiver positions by varying  $H$ ,  $D$ , and  $Z$ . The experimental data available to

us were for the following values:  $H$  from 70 to 90 cm in 10 cm increments,  $D$  from 20 to 60 cm also in 10 cm increments, and  $Z = (D_s + D)\tan a$ , where  $a$  ranges from  $0^\circ$  to  $30^\circ$  in 10 degree increments. In Fig. 11 the measured signals (solid line) are compared with the waveforms predicted by DLSM (dashed line) at various receiver locations, for normal ( $Z=0$ ), as well as for oblique incidence ( $Z \neq 0$ ). Overall, the diffracted signals obtained by DLSM are in reasonably good agreement with experimental signals at various receiver locations.

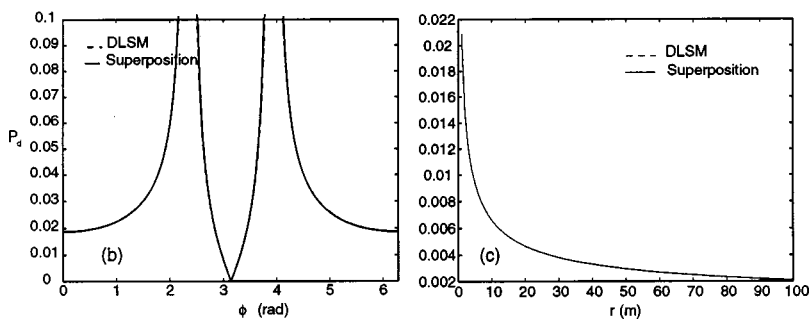
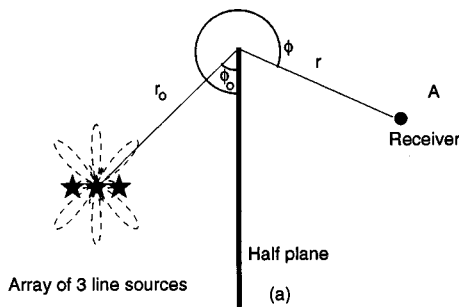


FIG. 10. Comparison between DLSM and other analytical solutions for diffraction of an array of three time harmonic line sources on a rigid half plane for a wide range of receiver locations. The distance  $d$  between the elements of the array is 0.1 m and the frequency is 5 kHz. In the application of the DLSM, the array is considered as one directive line source with directivity pattern shown with dashed lines in (a).

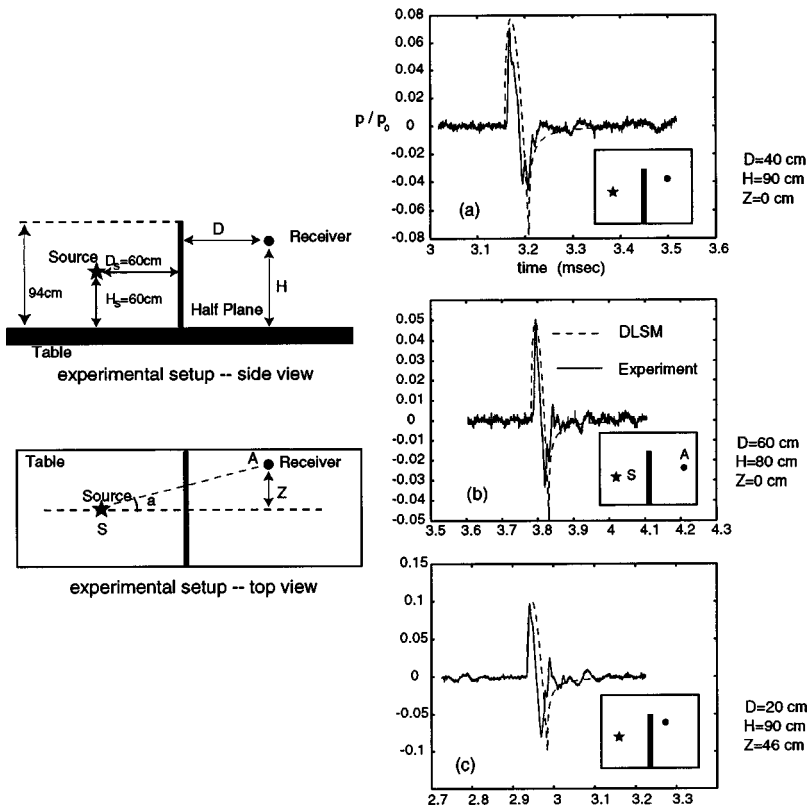


FIG. 11. Comparison between DLSM and experimental data at various receiver locations, for normal [source and receiver in a plane normal to the half plane  $Z=0$  (a),(b)] and oblique incidence [(c)].

#### IV. SUMMARY

In this paper a new method termed Directive Line Source Model is presented for predicting the diffracted field produced by a sound wave incident on a half plane. The model is based on a unified representation of diffraction by plane, cylindrical and spherical incident waves. The DLSM transforms a difficult diffraction problem into a much easier to solve radiation problem. Its advantages and relation to other methods of diffraction is given. Application to many cases—rigid half planes, pressure release half planes, wedges, and half planes having jagged edges—is discussed. Results for plane, cylindrical and spherical incident waves, as well as for arrays of line and point sources are presented. Agreement with known exact analytical solutions is demonstrated. Finally, results obtained from the DLSM compare favorably with experimental measurements.

#### ACKNOWLEDGMENTS

The work presented here was supported by the Texas Advanced Technology Program and the F. V. Hunt Postdoctoral Research Fellowship from the Acoustical Society of America.

<sup>1</sup>J. J. Bowman and T. B. A. Senior, *Electromagnetic and Acoustic Scattering by Simple Shapes*, edited by J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi (North-Holland, Amsterdam, 1969).  
<sup>2</sup>A. Sommerfeld, "Mathematische Theorie der Diffraction," *Math. Ann.* **47**, 317–374 (1896).  
<sup>3</sup>A. Sommerfeld, *Vorlesungen über Theoretische Physik*, Vol. 4, Optik (Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1959).  
<sup>4</sup>H. S. Carslaw, "Some multiform solutions of the partial differential equations of physics and mathematics and their applications," *Proc. London Math. Soc.* **30**, 121–163 (1899).

<sup>5</sup>H. S. Carslaw, "Diffraction of waves by a wedge of any angle," *Proc. London Math. Soc.* **18**, 1291–306 (1920).  
<sup>6</sup>H. M. MacDonald, *Electric Waves* (Cambridge University Press, Cambridge, England, 1902).  
<sup>7</sup>H. M. MacDonald, "A class of diffraction problems," *Proc. London Math. Soc.* **14**, 410–427 (1915).  
<sup>8</sup>F. G. Friedlander, *Sound Pulses* (Cambridge University Press, London, 1958).  
<sup>9</sup>M. A. Biot and I. Tolstoy, "Formulation of wave propagation in infinite media by normal coordinates with an application to diffraction," *J. Acoust. Soc. Am.* **29**, 381–391 (1957).  
<sup>10</sup>H. Medwin, "Shadowing by finite noise barriers," *J. Acoust. Soc. Am.* **69**, 1060–1064 (1981).  
<sup>11</sup>P. C. Clemmow, "A note on the diffraction of a cylindrical wave by a perfectly conducting half-plane," *Q. J. Mech. Appl. Math.* **3**, 377–384 (1950).  
<sup>12</sup>I. S. Gradshteyn and I. W. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965).  
<sup>13</sup>P. Menounou, "Theoretical study of diffraction by straight and ragged edge barriers," Ph.D. Dissertation, The University of Texas at Austin (1998).  
<sup>14</sup>P. Menounou, D. T. Blackstock, and I. J. Busch-Vishniac, "Jagged-edge noise barriers," in *Proceedings of the 16th International Congress on Acoustics*, pp. 97–98, Seattle, Washington, June 1998.  
<sup>15</sup>M. Born and E. Wolf, *Principles of Optics*, 3rd ed. (Pergamon, New York, 1965), pp. 375–382.  
<sup>16</sup>J. B. Keller, R. M. Lewis, and B. D. Seckler, "Diffraction by an Aperture II," *J. Appl. Phys.* **28**, 570–579 (1957).  
<sup>17</sup>T. F. W. Embleton, "Line integral theory of barrier attenuation in the presence of the ground," *J. Acoust. Soc. Am.* **67**, 42–45 (1980).  
<sup>18</sup>A. Michaeli, "Elimination of infinities in equivalent edge currents, Part I: Fringe current components," *IEEE Trans. Antennas Propag.* **34**, 912–918 (1986).  
<sup>19</sup>T. B. A. Senior and J. L. Volakis, *Approximate boundary conditions in electromagnetics*, Institute of Electrical Engineers, Electromagnetic Waves Series 41, London (1995).  
<sup>20</sup>K. M. Mitzner, "Incremental length diffraction coefficients," Aircraft Division, Northrop Corporation Technical Report No. AFAL-TR-73-296 (1974).  
<sup>21</sup>E. F. Knott, "The relationship between Mintzer's ILDC and Michaeli's equivalent currents," *IEEE Trans. Antennas Propag.* **33**, 112–114 (1985).

- <sup>22</sup> A. K. Gautesen, "Integral representations for edge diffraction," *J. Acoust. Soc. Am.* **72**, 1942–1946 (1982).
- <sup>23</sup> U. P. Svensson, R. I. Fred, and J. Vanderkooy, "An analytic secondary model of edge diffraction impulse responses," *J. Acoust. Soc. Am.* **106**, 2331–2344 (1999).
- <sup>24</sup> J. B. Keller, "Diffraction by an aperture," *J. Appl. Phys.* **28**, 426–444 (1957).
- <sup>25</sup> J. B. Keller, "The geometrical theory of diffraction," *J. Opt. Soc. Am.* **52**, 116–130 (1962).
- <sup>26</sup> S. N. Karp and J. B. Keller, "Multiple diffraction by an aperture in a hard screen," *Opt. Acta* **8**, 61–72 (1961).
- <sup>27</sup> A. D. Pierce, "Diffraction of sound around corners and over wide barriers," *J. Acoust. Soc. Am.* **55**, 941–955 (1974).
- <sup>28</sup> A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, American Institute of Physics, New York, 1989).
- <sup>29</sup> I. Tolstoy, "Diffraction by a hard truncated wedge and a strip," *IEEE J. Ocean Eng.* **14**, 4–16 (1989).
- <sup>30</sup> I. Tolstoy, "Exact, explicit solutions for diffraction by hard sound barriers and seamounts," *J. Acoust. Soc. Am.* **85**, 661–669 (1989).
- <sup>31</sup> E. Heyman and L. B. Felsen, "Traveling wave and SEM representations for transient scattering by a circular cylinder," *J. Acoust. Soc. Am.* **79**, 230–238 (1986).
- <sup>32</sup> E. J. Rosenberg, "Continued investigation of noise reduction by a random edge barrier," MS thesis, The University of Texas at Austin (1997).

# Limits of time-reversal focusing through multiple scattering: Long-range correlation

Arnaud Derode,<sup>a)</sup> Arnaud Tourin, and Mathias Fink

Laboratoire Ondes et Acoustique, Université Denis Diderot-Paris 7, ESPCI, 10 rue Vauquelin, 75005 Paris, France

(Received 25 October 1999; accepted for publication 17 March 2000)

Experimental results of time-reversal focusing in a high-order multiple scattering medium are presented and compared to theoretical predictions based on a statistical model. The medium consists of a random collection of parallel steel rods. An ultrasonic source (3.2 MHz) transmits a pulse that undergoes multiple scattering and is recorded on an array. The time-reversed waves are sent by the array back to the source through the scattering medium. The quality of temporal focusing is very well predicted by a simple statistical model. However, for thicker samples, persistent temporal side-lobes appear. We interpret these side-lobes as a consequence of the growing number of crossing paths in the sample due to high-order multiple scattering. As to spatial focusing, the resolution is practically independent from the array's aperture. With a 16-element array, the resolution was found to be 30 times finer than in a homogeneous medium. Resolutions of the order of the wavelength (0.5 mm) were attained. These results are discussed in relation with the statistical properties of time-reversal mirrors in a random medium. © 2000 Acoustical Society of America. [S0001-4966(00)05406-0]

PACS numbers: 43.20.Fn, 43.30.Ft, 43.60.Gk [DEC]

## INTRODUCTION

The property of time-reversal invariance in the wave equation provides an elegant and robust method to focus an acoustic wave through an inhomogeneous medium.<sup>1</sup> Usually it requires an array of piezoelectric transducers (although time-reversal can also be achieved on a single transducer) that are used as a time-reversal mirror, i.e., a device that can record a waveform  $f(t)$ , time-reverse it and send back ( $f - t$ ) into the same medium. In a time-reversal experiment, a source transmits a short pulse that propagates and may be scattered or reflected by inhomogeneities within the medium; the resulting pressure wave is recorded on an array. When the array sends back the time-reversed waves, they propagate as if they were traveling backwards in time and refocus both spatially and temporally onto the source. In recent years, this principle has been successfully applied to a variety of domains: nondestructive evaluation of materials,<sup>2</sup> underwater acoustics,<sup>3</sup> and medicine.<sup>4,5</sup>

Particularly, it has been shown that time-reversal focusing was a very robust method that could still refocus a wave even in the presence of very high-order multiple scattering.<sup>6</sup> In a typical experiment such as the one depicted in Fig. 1, a multiple scattering sample is placed between the source and the array. The whole setup is in a water tank. The scattering sample is a set of parallel steel rods (diameter 0.8 mm) randomly distributed in the plane. The sample thickness is  $L = 15$  mm. There are roughly 19 rods per  $\text{cm}^2$  so the typical distance between two rods is 2.3 mm. When the source sends a short pulse ( $2\frac{1}{2}$  cycles of a 3.2-MHz sine wave) the transmitted signal shows clear evidence of multiple scattering: it spreads over more than 100  $\mu\text{s}$ , i.e.,  $\sim 150$  times the initial

pulse duration. Yet, the time-reversal process is able to recreate a very short signal back on the source, as is shown in Fig. 2. Moreover, it was shown<sup>5</sup> that the spatial resolution of the time-reversed beam around the source was much better in a multiple scattering medium than in a homogeneous medium, and this remains true even if the received signals are digitized on a single bit!<sup>7</sup>

Time-reversal mirrors take advantage of high-order multiple scattering, whereas one usually tries to get rid of it in medicine or nondestructive evaluation. A question then arises: which amount of multiple scattering should there be to achieve the best focusing? From previous results,<sup>5</sup> one may think that the higher the order of multiple scattering, the better.

In this paper, we will show that when multiple scattering is not too strong, a simple statistical model can predict the efficiency of the time-reversal process as a function of the length of the time-reversed signals and the number of elements on the array. Then we will show that when multiple scattering becomes too strong, the quality of focusing becomes poorer due to internal correlations. We interpret the physical origin of these correlations as the result of crossing paths within the scattering sample due to the high order of scattering involved. The paper is organized in two sections. In Sec. I, the quality of temporal focusing at the source will be studied both theoretically and experimentally. In Sec. II we will deal with spatial focusing around the source.

## I. TEMPORAL FOCUSING ON THE SOURCE

In this section, we will examine the properties of the signal recreated at the source location, as the one plotted in Fig. 2(b). There are essentially two variables in a time-reversal experiment:  $N$ , the number of active elements on the array, and  $[T_1, T_2]$  the time-reversal window selected in the

<sup>a)</sup>Electronic mail: arnaud.derode@espci.fr

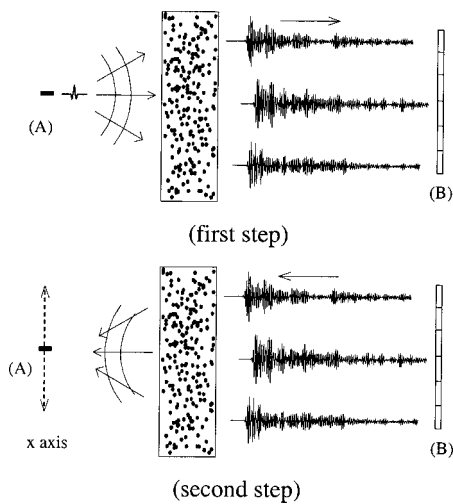


FIG. 1. Time-reversal focusing. In the first step the source (A) transmits a short pulse that propagates through the rods. The scattered waves are recorded on a 128-elements array (B). In the second step,  $N$  elements of the array ( $0 < N < 128$ ) retransmit the time-reversed signals through the rods. The piezoelectric element (A) is now used as a detector, and measures the signal reconstructed at the source position. It can also be translated along the  $x$  axis while the same time-reversed signals are transmitted by B, in order to measure the directivity pattern.

scattered signals. The parameter that will serve us as a measurement of the quality of temporal focusing will be the signal-to-noise ratio (SNR) defined as the peak amplitude divided by the standard deviation of the surrounding noise. We will build a simple model to predict the value of the SNR in a multiple scattering sample, given  $N$  and  $[T_1, T_2]$ .

Imagine that the source transmits a pulse  $\delta(t)$  through the sample. As a consequence, the array records  $N$  impulse responses  $h_i(t)$ ,  $i = 1, \dots, N$ . Due to reciprocity, the propagation impulse response from the source to the  $i$ th element is the same as that from the  $i$ th element to the source. Therefore, the total signal  $r(t)$  recreated on the source after time-reversal writes as

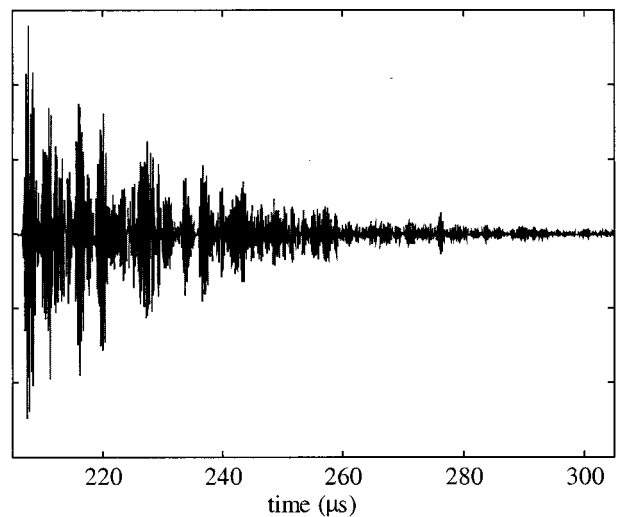
$$r(t) = \sum_{i=1}^N h_i(t) \otimes h_i(-t) = \sum_{i=1}^N c_i(t), \quad (1)$$

where  $c_i(t)$  denotes the contribution of the  $i$ th element of the time-reversal mirror to the final signal  $r(t)$ .

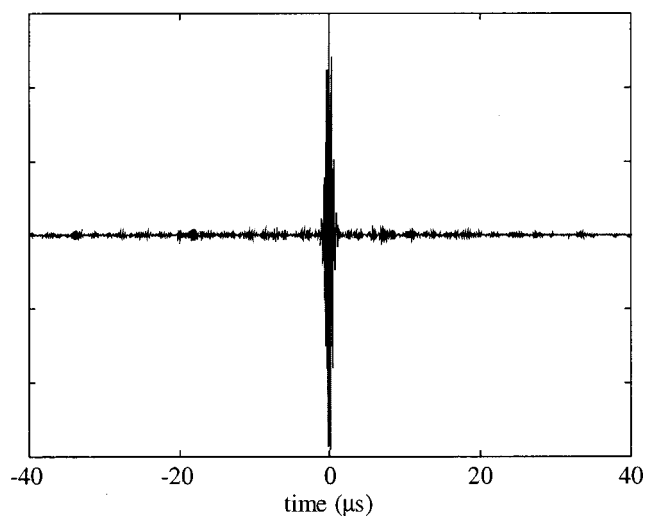
This signal results in a double summation: one over time [the convolution of each impulse response  $h_i(t)$  by its time-reversed self] and one over space [the addition of the contributions  $c_i(t)$  coming from the  $N$  elements of the array].

Since the positions of the scatterers are random, we will treat the signal  $r(t)$  as a random signal. In fact,  $r(t)$  writes as a double sum (in space and in time) of random variables. As a consequence, one of the key issues in this article will be the correlation properties of the impulse responses  $h_i(t)$  both in time and in space.

As an example, Fig. 3 represents some of the contributions  $c_i(t)$  and their summation  $r(t)$  for  $N=122$ ; these results were obtained through a 15-mm-thick sample. Each of the  $c_i(t)$  is maximum and positive at time  $t=0$ , so the summation of all the  $c_i(t)$  to form  $r(t)$  necessarily gives rise to a peak at time  $t=0$ , the amplitude of the peak growing proportionally with  $N$ ; outside the peak, if we assume that con-



(a)



(b)

FIG. 2. (a) Signal transmitted through the rods ( $L=15$  mm), recorded by the array element no. 64. (b) Signal recreated at the source position after time-reversal of the scattered signals by  $N=122$  elements of the array.

tributions from different elements  $c_i(t)$  and  $c_j(t)$  are not correlated, the standard deviation should only rise as  $\sqrt{N}$ , therefore the SNR should also increase as the square root of the number of transducers. With a similar argument, we can expect the SNR to rise proportionally to  $\sqrt{\Delta T}$  if we expand the size of the time-reversal window  $\Delta T = T_2 - T_1$ .

Actually, the dependence of the SNR on  $\Delta T$  raises the issue of temporal correlation between  $h_i(t_1)$  and  $h_i(t_2)$ , whereas the dependence of the SNR on  $N$  is linked to the spatial correlation between  $c_i(t)$  and  $c_j(t)$ .

To begin with, we will only consider the problem of temporal correlation, i.e., the evolution of the SNR as a function of  $\Delta T$ . To that end we will keep  $N=1$ , and study the time-reversal process on a single transducer.

#### A. $N=1$ , $\Delta T$ variable

At a time  $t$  outside the peak, the theoretical expression of the SNR is

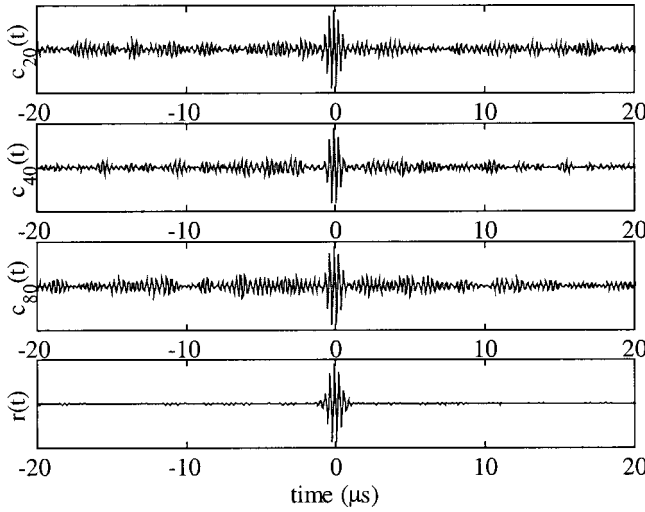


FIG. 3. Contributions  $c_i(t)$  of three array elements, Nos. 20, 40 and 80. Each contribution is maximum and positive at time  $t=0$ . The total signal  $r(t)$  results from the summation of  $N=122$  contributions. On the vertical scale, each signal has been normalized by its maximum.

$$\text{SNR}(t) = \frac{\langle r(0) \rangle}{\sqrt{\langle r^2(t) \rangle - \langle r(t) \rangle^2}}, \quad (2)$$

where the brackets  $\langle \rangle$  denote an ensemble average.

In order to calculate the theoretical value of the SNR, we need a random approach to model the scattered signals such as the one plotted in Fig. 2(a). When the source transducer is excited by an electric Dirac pulse, the signal received on one element of the array is the impulse response of the total system source+rods+receiver, and randomness occurs because of the rods. We will follow a naive approach that consists in treating this random impulse response as a ‘‘shot noise’’  $h(t)$ , i.e., a series of replica with random arrival times:<sup>8</sup>

$$h(t) = \sum_j e(t-t_j). \quad (3)$$

Here  $e(t)$  is the zero mean acousto-electrical impulse response of the source convoluted by that of the receiver. The arrival times  $t_j$  are supposed to be independent random variables, and represent the numerous multiple scattering paths. If the number of arrival times per time unit is large, that is to say, if there is a large number of scattering paths with the same length, the signal  $h(t)$  is a sum of many independent random variables, and can be treated as one realization of a Gaussian process (central limit theorem). Furthermore, we assume that the duration  $E$  of  $e(t)$  is much smaller than that of the distribution of arrival times—typically,  $E < 1 \mu\text{s}$  whereas the scattered signal is spread on several tens of microseconds.

These assumptions allow us to consider the scattered signal  $h(t)$  as a nonstationary Gaussian signal with zero mean and variance  $\sigma^2(t)$ , whose autocorrelation function is written

$$\langle h(t_1)h(t_2) \rangle = \sigma(t_1)\sigma(t_2)\rho(t_2-t_1), \quad (4)$$

where  $\rho$  is the correlation coefficient between  $h(t_1)$  and  $h(t_2)$ .

As long as the arrival times are independent, it can easily be shown that

$$\rho(\tau) = \frac{\int e(\theta)e(\tau+\theta) d\theta}{\int e^2(\theta) d\theta} \propto e(\tau) \otimes e(-\tau). \quad (5)$$

So the correlation properties of the scattered signal only depend on  $e(t)$ : Therefore, the correlation time  $\tau_c$  of the scattered signal  $h(t)$  is of the same order as the duration  $E$  of  $e(t)$ . Finally, we assume that  $\sigma(t)$  is a slowly varying function compared to  $e(t)$ , which implies that

$$\tau_c \ll \tau_\sigma, \quad (6)$$

$\tau_\sigma$  being some characteristic width of  $\sigma(t)$ .

After the time-reversal, the signal reconstructed on the source is

$$r(t) = h(t) \otimes h(-t) W(-t) = \int h(\theta) W(\theta) h(t+\theta) d\theta, \quad (7)$$

where  $W(\theta)$  defines the time-reversal window: it is equal to 1 for  $T_1 < \theta < T_2$  and 0 otherwise. We can study the first two statistical moments of  $r(t)$ .

Following the assumptions mentioned earlier, the expected value of  $r(t)$  is straightforwardly

$$\langle r(t) \rangle = \int \langle h(\theta)h(t+\theta) \rangle W(\theta) d\theta = \rho(t) \int_{T_1}^{T_2} \sigma^2(\theta) d\theta. \quad (8)$$

On average, the signal recreated on the source after time-reversal is proportional to  $\rho(t)$  and is therefore maximum at time  $t=0$ . Its amplitude increases with  $\Delta T$ , and its typical duration is  $\tau_c$ . This is not surprising since time-reversal and backpropagation are formally analogous to a correlation integral, and indeed  $r(t)$  can be viewed as a statistical estimator of the autocorrelation function of  $h(t)$ .

Knowing the expected value of  $r(t)$  is not sufficient, we also have to calculate its variance in order to evaluate by how much  $r(t)$  is likely to fluctuate around its mean value.

After some computation,<sup>6</sup> the variance of  $r(t)$  outside the central peak ( $|t| > E$ ) is given by

$$\langle r^2(t) \rangle - \langle r(t) \rangle^2 = \int \rho^2(\tau) d\tau \int_{T_1}^{T_2} \sigma^2(\theta)\sigma^2(t+\theta) d\theta. \quad (9)$$

Hence, the theoretical value of the SNR is

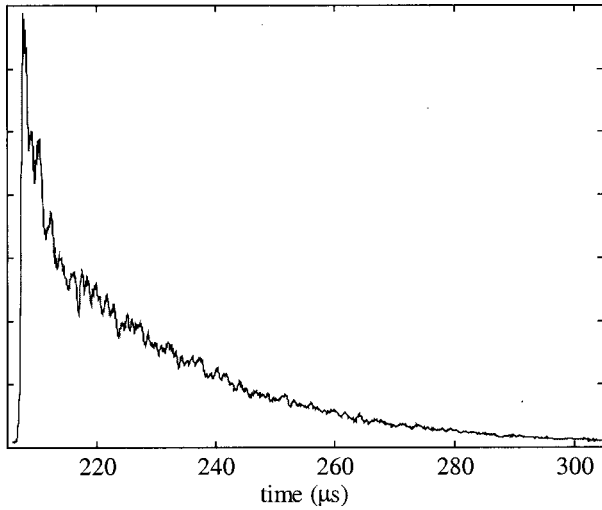
$$\text{SNR}(t) = \frac{1}{\sqrt{\int \rho^2(\tau) d\tau}} \times \frac{\int_{T_1}^{T_2} \sigma^2(\theta) d\theta}{\sqrt{\int_{T_1}^{T_2} \sigma^2(\theta)\sigma^2(t+\theta) d\theta}}. \quad (10)$$

The integral in the first term of the right-hand side of Eq. (10) has the dimensions of a time. Actually, this integral can be used as a definition of the correlation time  $\tau_c$  of the scattered impulse response  $h(t)$ :

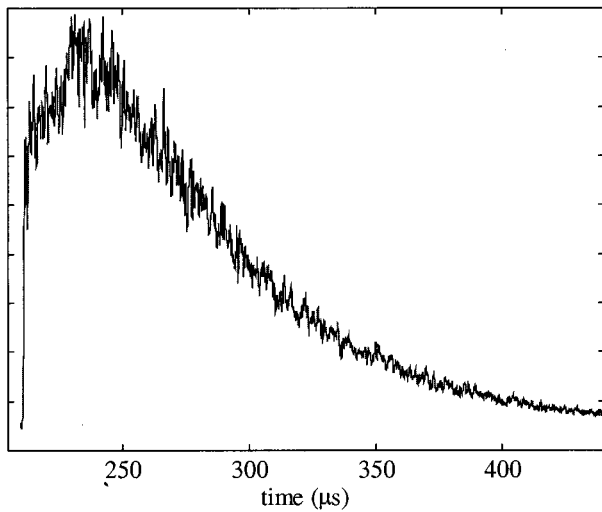
$$\tau_c = \int \rho^2(\tau) d\tau. \quad (11)$$

As to the second factor on the right-hand side of Eq. (10), it strongly depends on  $\sigma(\theta)$ , whose shape can be roughly inferred from the average envelope of the transmitted signal. In the simple case where  $\sigma(\theta)$  can be considered constant, then Eq. (10) greatly simplifies into





(a)



(b)

FIG. 4. Average envelope of the signals transmitted through multiple scattering samples,  $L=15$  mm (a) and  $L=40$  mm (b).

$$\text{SNR}(t) = \sqrt{\frac{\Delta T}{\tau_c}}. \quad (12)$$

This result has a very simple interpretation.  $\Delta T$  is the duration of the time-reversal window, and  $\tau_c$  is the correlation time of the scattered signal. Therefore, the ratio  $\Delta T/\tau_c$  gives the number of independent information (or “information grains”) within the time-reversal window. When we expand the time-reversal window, we add more and more new information (i.e., uncorrelated with previous data), and the resulting enhancement of the SNR is as usual proportional to the square root of the number of so-called “information grains.”

In reality,  $\sigma(\theta)$  cannot be strictly constant and usually decays with a characteristic time  $\tau_\sigma$ . In the time-reversal experiment we have carried out through a 15-mm-thick scattering sample, the average envelope (Fig. 4) of the transmitted signals decays exponentially with a decay time  $\tau_\sigma = 27.2 \mu\text{s}$ . Taking into account this exponential decay, we obtain the following expression for the SNR:

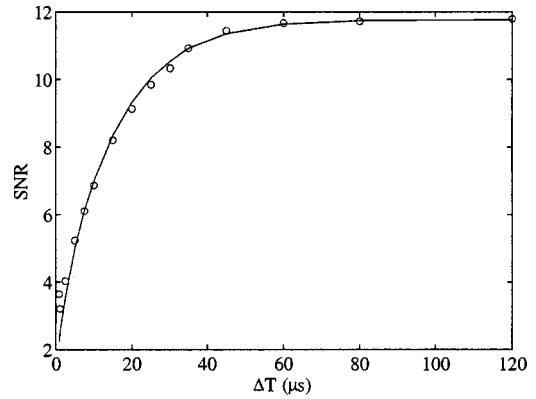


FIG. 5. Time-reversal through a 15-mm-thick scattering sample: signal-to-noise ratio versus duration of the time-reversal window. The circles are the experimental results, the continuous line represents the theoretical prediction for  $\tau_c=0.29 \mu\text{s}$ .

$$\text{SNR}(t) = \sqrt{\frac{2\tau_\sigma}{\tau_c}} \frac{\sinh(\Delta T/\tau_\sigma)}{\sqrt{\sinh(2\Delta T/\tau_\sigma)}} \exp(t/\tau_\sigma). \quad (13)$$

Note that this expression of the SNR is not stationary: the noise level decays exponentially with time, as the transmitted signals do. In practice, when the noise level is not stationary, the SNR is calculated by taking the root mean square value of the noise at a time  $t$  off the central peak, in a time window much smaller than the decay time  $\tau_\sigma$ .

We plotted the SNR as a function of the duration of the time-window  $\Delta T$  in Fig. 5; we compared it to the experimental results we obtained by time-reversal through a 15-mm-thick scattering sample. The experimental procedure was the following: the source and the time-reversal mirror were 31 cm apart, the scattering sample was placed 18 cm from the array. The scattered signals were digitized on 9 bits at a 20-MHz sampling rate. The beginning of the time-reversal window was taken at time  $T_1=211 \mu\text{s}$ , and  $\Delta T$  was varied from 1 to 120  $\mu\text{s}$ . For each value of  $\Delta T$ , the time-reversal process was performed separately by the 122 array elements, one at a time. For each contribution  $c_i(t)$ ,  $i=1,\dots,122$ , the SNR was calculated as the ratio of the peak amplitude ( $t=0$ ) to the standard deviation of the surrounding noise in an 8- $\mu\text{s}$  time-window centered at  $t=6 \mu\text{s}$  after the peak. The final experimental value of the SNR that is plotted in Fig. 5 is the average of the 122 SNR deduced from the 122 contributions  $c_i(t)$ .

In the theoretical expression of the SNR, it should be noted that the correlation time  $\tau_c$  cannot be known exactly. If we assume that the transmitted signal is a series of replica of the incoming signal ( $2\frac{1}{2}$  cycles of a 3.2-MHz sine wave), then, from Eqs. (5) and (11), we have  $\tau_c=0.27 \mu\text{s}$ . Yet, when comparing the theoretical and experimental results, the best agreement was obtained for a correlation time  $\tau_c = 0.29 \mu\text{s}$ , which seems to make sense since the convolution of the incoming signal by the acousto electrical response of the transducers tends to slightly increase the correlation time.

Finally, as we can see in Fig. 5, the agreement between the experimental value of the SNR and the theoretical prediction based on the “shot noise” model is excellent. It should be noted that we obtained the same agreement when

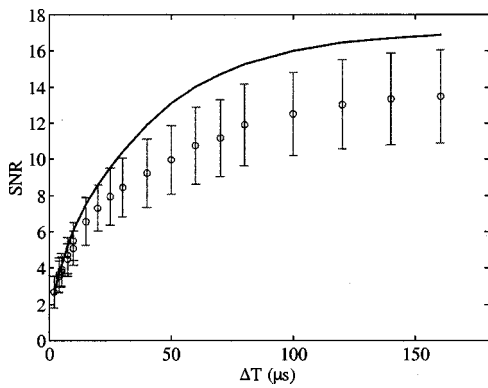


FIG. 6. Time-reversal through a 40-mm-thick scattering sample: signal-to-noise ratio versus duration of the time-reversal window. The circles are the experimental results ( $\pm$  one standard deviation), the continuous line represents the theoretical prediction for  $\tau_c = 0.29 \mu s$ .

evaluating the rms noise level surrounding the peak in a different time window.

So it seems that Eq. (10) can be used as a reliable prediction of the temporal sidelobe level in a time-reversal focusing experiment. Here, due to the exponential decay ( $\tau_\sigma = 27.2 \mu s$ ) of the transmitted signals, the SNR saturates and reaches its maximum value after  $\Delta T \sim 70 \mu s$ , and it is useless to expand the time-reversal window further.

How to obtain a better SNR? We need to have a larger decay time, and consequently longer time windows  $\Delta T$ , to record more “information grains.” In other words, we need to enlarge the order of multiple scattering in the sample. To that end, we can repeat the same experiment through a thicker sample ( $L = 40 \text{ mm}$ ). This time, the transmitted signals last much longer: the average envelope of the transmitted signals is plotted in Fig. 4; after the maximum, the characteristic decay time is  $\sim 76 \mu s$ .

The experimental procedure is the same as before. The beginning of the time-reversal window is  $T_1 = 220 \mu s$ , and its duration  $\Delta T$  varies from 1 to 160  $\mu s$ . The comparison between experimental and theoretical results derived from Eq. (10) is plotted in Fig. 6.

This time, we observe that the model fails at giving a correct prediction, especially at larger  $\Delta T$ . The experimental value of the SNR seems to be systematically and significantly lower than the theoretical prediction.

What can be the cause of this discrepancy? When the time-reversal window is enlarged, we record a larger part of the scattered signal and send it back into the medium. For instance, when  $\Delta T$  is doubled, then we send back twice as many “information grains” (provided the arrival times are uncorrelated), hence the SNR should increase by a factor of  $\sqrt{2}$ . If the SNR no longer increases when adding these information grains to the time-reversed signal, it can be either because their amplitude become negligible (as was the case for  $L = 15 \text{ mm}$ ), or that this information is not new, i.e., it is correlated with the earlier part of the scattered signal. In the case of the thicker sample ( $L = 40 \text{ mm}$ ), the decay time is too large to explain the saturation of the SNR. A possible explanation is that the “shot noise” model is not valid in this sample: at large  $\Delta T$ , the arrival times may no longer be considered independent. Even though the scatterers’ posi-

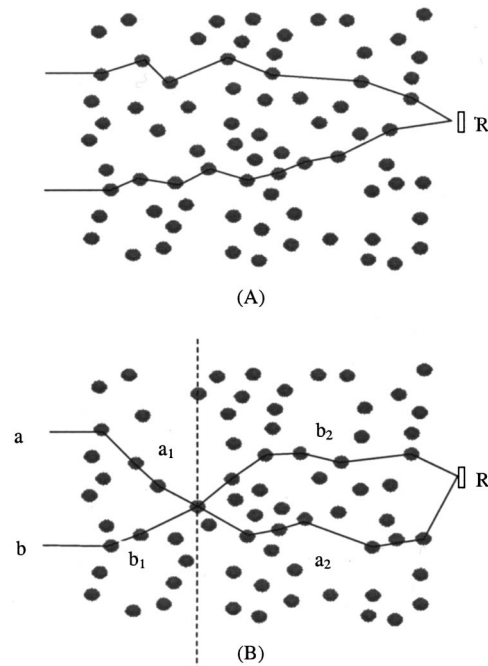


FIG. 7. Sketches of scattering paths arriving on the receiver  $R$ : (a) two independent paths and (b) two crossing paths that generate four possible arrival times on the receiver.  $a_1$  and  $a_2$  are the travel times corresponding to the first and second parts (before and after the crossing) of the path labeled  $a$ . The same convention applies to the path labeled  $b$ .

tions are uncorrelated, the multiple scattering paths may be somewhat correlated.

In optics, long-range correlations in the diffuse intensity transmitted through a highly scattering slab have been recently observed.<sup>9</sup> Yet the problem is different here, since we are not interested in the transmitted intensity but in adaptive time-reversal focusing.

A possible origin of the correlation in the arrival times is the growing number of crossing paths as the sample thickness increases. We give a simple interpretation of the phenomenon in Fig. 7. We keep the idea that the wave is scattered in the sample following every possible path, as a random walker that goes from one scatterer to another before getting out of the sample. Each path corresponds to an arrival time  $t_i$  in the scattered signal received at some point  $R$ . Consider the two paths in Fig. 7(a): they are totally independent of each other, and since there is no correlation between the scatterers, there is no correlation either between the arrival times  $t_1$  and  $t_2$ , as in a “shot noise.” But there are also more complicated paths, as the ones drawn in Fig. 7(b). In this example, two paths  $a$  and  $b$  are crossing in the sample, giving rise to four possible arrival times at point  $R$ :  $t_1$ ,  $t_2$ ,  $t_3$  and  $t_4$ . There are four arrival times, but only two couple of independent paths. Indeed, we have  $t_1 = a_1 + a_2$ ,  $t_2 = b_1 + b_2$ ,  $t_3 = a_1 + b_2$ ,  $t_4 = a_2 + b_1$ :  $t_1$  and  $t_2$  are still independent, as well as  $t_3$  and  $t_4$ , but  $t_1$  is correlated with  $t_3$  and  $t_4$ , and so is  $t_2$ .

Let us examine the consequence of this correlated paths on the temporal side-lobe level. For simplicity, imagine that the scattered signal is a series of impulses

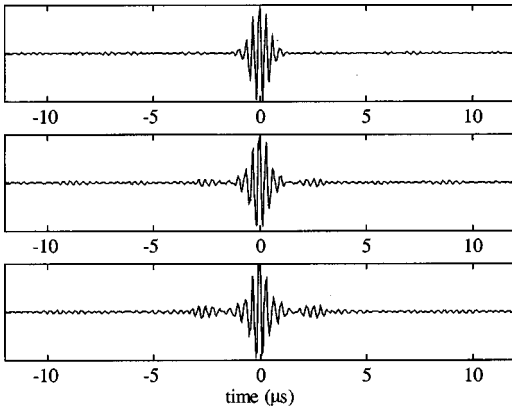


FIG. 8. Signal recreated on the source position after time-reversing  $N = 122$  scattered signals through three multiple scattering samples:  $L = 15$  mm,  $L = 40$  mm,  $L = 80$  mm. As the sample thickness increases, the secondary temporal side lobes grow, due to internal correlations within the scattering sample.

$$h(t) = \sum_{i=1}^{N_A} \delta(t - t_i), \quad (14)$$

where  $N_A$  is the number of arrival times and the  $t_i$  are random variables distributed over some interval  $[0, T]$ . The time-reversal process will give rise to a signal

$$h(t) \otimes h(-t) = \sum_{j=1}^{N_A} \sum_{i=1}^{N_A} \delta(t - t_i + t_j). \quad (15)$$

There are  $N_A^2$  terms in this expression: the  $N_A$  terms for which  $i = j$  will create a peak at time  $t = 0$ . The other terms ( $i \neq j$ ) will contribute to the noise surrounding the central peak, creating secondary peaks for all possible values of  $\delta t = t_i - t_j$ .

Now, if every pair  $(t_i, t_j)$  forms a couple of independent random variables for  $i \neq j$ , then the time differences  $\delta t$  will be distributed over the whole interval  $[-T, T]$ . Whereas if there is a correlation between  $t_i$  and  $t_j$  ( $i \neq j$ ), this will not be the case. For instance, if we have  $t_1 = a_1 + a_2$  and  $t_3 = a_1 + b_2$ , then  $\delta t = b_2 - a_2$  will only be distributed over an interval  $[-B, B]$ , where  $B$  is necessarily smaller than  $T$ . This implies that the contributions of the crossing paths to the “noise” tend to be gathered around the central peak instead of being spread all over the interval  $[-T, T]$ .

Hence, the SNR is decreased by the presence of crossing paths, and naturally this effect becomes more obvious when the number of crossing paths increases. In a random walk, the essential parameter is the transport mean free path  $l^*$ . As a physical interpretation,  $l^*$  is a characteristic distance after which the wave has lost the memory of its initial direction. In our samples, previous experiments showed that the transport mean free path is of the order of 4 mm.<sup>10</sup> When the sample thickness  $L$  is not too large compared to  $l^*$ , most paths going to the receiver do not cross each other. But as the sample thickness  $L$  increases, the paths traversing the slab get more and more intricate and the probability of crossing increases, inducing a correlation in the arrival times.

This has an important impact on the efficiency of a time-reversal experiment. Figure 8 presents the time-reversed signals we obtained at the source with a 122-element array

through various thicknesses: 15, 40, and 80 mm. For each value of  $L$ , all 122 signals were time-reversed and the resulting signal  $r(t)$  was plotted.

For  $L = 15$  mm, we see practically no temporal side lobes, and as we have seen previously, the SNR is very well predicted by a “shot noise” approach. For  $L = 40$  mm, a side lobe appears close to the peak: this side lobe is responsible for the discrepancy between the shot noise model and the experimental results. At  $L = 80$  mm, the same side lobe is present, at the same time ( $\sim 2.5 \mu\text{s}$  before and after the central peak) and with a stronger amplitude.

Remember that the signal recreated at the source writes

$$r(t) = \sum_{i=1}^N h_i(t) \otimes h_i(-t) = \sum_{i=1}^N \int h_i(\theta) h_i(t + \theta) d\theta, \quad (16)$$

hence  $r(t)$  can be viewed as a statistical estimate of the autocorrelation function of the scattered waves  $\langle h(\theta) h(t + \theta) \rangle$ . If the shot noise model was right and  $h(t)$  was actually a decorrelated series of replicas of  $e(t)$ , then  $\langle h(\theta) h(t + \theta) \rangle$  should be 0 for  $|t| > E$ , where  $E$  is the duration of  $e(t)$ . The existence of persisting temporal side lobes for large thickness is an indication of a correlation in the arrival times. We interpret this correlation as a consequence of the growing number of crossing paths.

In two samples with the same scatterer concentrations and different thicknesses ( $L = 40$  mm and  $L = 80$  mm), the first side lobe appears at the same time:  $\sim 2.5 \mu\text{s}$  from the main peak, which corresponds to a path length difference of 3.75 mm in water. This is probably related to the average distance between two scatterers (2.4 mm). Indeed, among all possible crossing paths, the simplest and probably more frequent is a “recurrent scattering:” the wave bounces between two scatterers. Given the radius of the scatterers (0.4 mm), the typical bouncing time is  $\sim 2 \mu\text{s}$ . Yet this interpretation has to be confirmed by studying the side lobes with a different density of scatterers.

Finally, the “shot noise” model was shown to be very accurate at  $L = 15$  mm, but failed at giving the correct prediction at a larger thickness. Should we just abandon the shot noise model and only use it when crossing paths can be neglected? In fact, the model can be modified to take into account the presence of correlations in the arrival times due to multiple scattering. We still assume that the impulse response  $h(t)$  is a series of replicas with random arrival times:

$$h(t) = \sum_j e(t - t_j), \quad (17)$$

but the arrival times  $t_i$  are not independent, their correlation function  $R(\tau)$  is no longer a delta function. Then the correlation coefficient  $\rho(\tau)$  of the scattered signal in Eq. (5) has to be replaced by

$$\rho(\tau) \propto e(\tau) \otimes e(-\tau) \otimes R(\tau). \quad (18)$$

This only affects the first term on the right-hand side of Eq. (10), i.e., the correlation time  $\tau_c$  of the scattered signal. The only effect of additional correlation due to crossing paths is to increase the correlation time. Indeed, if we take  $\tau_c = 0.45 \mu\text{s}$  instead to  $\tau_c = 0.29 \mu\text{s}$ , then the experimental re-

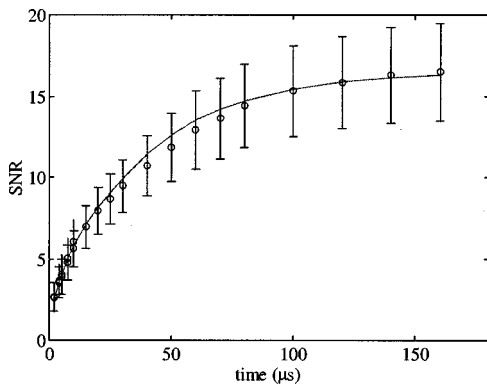


FIG. 9. Time-reversal through a 40-mm-thick scattering sample: signal-to-noise ratio versus duration of the time-reversal window. The circles are the experimental results ( $\pm$  one standard deviation), the continuous line represents the theoretical prediction for  $\tau_c = 0.45 \mu\text{s}$ .

sults obtained at  $L = 40 \text{ mm}$  are in good agreement with the prediction of the model (Fig. 9). However, it is clear that the assumption that  $h(t)$  is formed with identical replicas of the same signal  $e(t)$  is not very realistic; particularly, at very large times, frequency-dependent attenuation probably plays a more important role.

### B. $\Delta T$ constant, $N$ variable

So far, we have studied the SNR as a function of the duration  $\Delta T$  of the time-reversal window, with only one element performing the time-reversal operation. We are now going to keep  $\Delta T$  constant and study the SNR as a function of  $N$ , the number of active elements on the array. The signal recreated at the source results from the summation of  $N$  contributions from the active elements on the array:

$$r(t) = \sum_{i=1}^N h_i(t) \otimes h_i(-t) = \sum_{i=1}^N c_i(t). \quad (19)$$

As we have seen (Fig. 3), all these contributions add constructively at time  $t = 0$  to form the central peak. Outside the peak, the question is to know whether the  $c_i(t)$  and  $c_j(t)$  ( $i \neq j$ ) are correlated or not. In the ideal case where each  $c_i$  is fully decorrelated from every  $c_j$ , then the SNR should rise as  $\sqrt{N}$ . If  $c_i(t)$  and  $c_{i+m}(t)$  are partially correlated, as long as this correlation is short range and falls to zero after some length  $l_{\text{coh}}$ , the SNR should still increase proportionally to  $\sqrt{N/m_{\text{coh}}}$ , with  $m_{\text{coh}} = l_{\text{coh}}/\text{array pitch}$ . Once again, as the aperture of the array increases, it receives more and more “information grains,” and the SNR varies proportionally to the square root of the number of information grains.

Figure 10 presents the experimental values of the SNR as a function of  $\sqrt{N}$ , through  $L = 15 \text{ mm}$ . The time-reversal window begins at  $T_1 = 211 \mu\text{s}$  and its duration is  $\Delta T = 120 \mu\text{s}$ . As we did previously, the SNR was calculated as the ratio of the peak amplitude to the standard deviation of the surrounding noise close to the peak, in a  $8\text{-}\mu\text{s}$  time-window centered at  $t = 6 \mu\text{s}$  after the peak. For each value of  $N$ , the results presented in Fig. 10 are averaged over the  $123-N$  groups of  $N$  adjacent elements. Once again, the experimental results confirm what was expected: the SNR

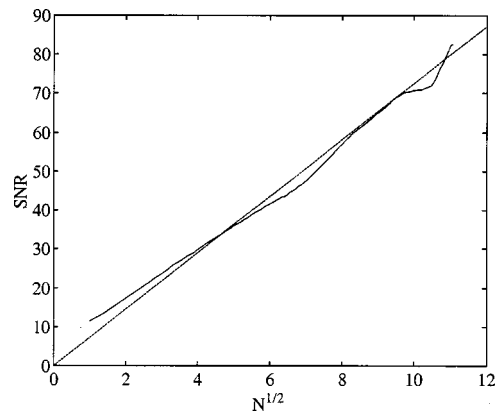


FIG. 10. Time-reversal through a 15-mm-thick scattering sample: signal-to-noise ratio versus square root of the number of active array elements.

grows linearly with  $\sqrt{N}$ , in agreement with the independent “information grains” theory. The slope of the linear fit yields  $m_{\text{coh}} \approx 2.5$ .

However, if we repeat the same experiment through a 40-mm-thick sample ( $T_1 = 220 \mu\text{s}$ ,  $\Delta T = 160 \mu\text{s}$ ), and calculate the SNR in the same time window as usual, there is a discrepancy between theory and experiments. The SNR starts rising like  $\sqrt{N}$ , then saturates and reaches a maximum value (Fig. 11). The effect is even more striking for a thicker sample; for  $L = 80 \text{ mm}$ , with  $N = 122$  elements, the SNR is only 1.5 times higher than with one element!

This saturation cannot be explained by a short-range correlation; it is due to a long-range spatial correlation between contributions  $c_i(t)$  and  $c_j(t)$ , even for  $i \neq j$  and  $|t| > E$ . As before, its physical origin may lie in the growing importance of crossing paths. As the sample thickness increases, the proportion of crossing paths rises, inducing a correlation  $R(\tau)$  in the arrival times. The expected value  $\langle r(t) \rangle$  of the signal recreated at the source is no longer  $e(t) \otimes e(-t)$ , but  $e(t) \otimes e(-t) \otimes R(t)$ . The correlation  $R(\tau)$  in the arrival times induces temporal side lobes, particularly around  $2.5 \mu\text{s}$  after the peak as we have pointed out.

Each transmitted signal  $h_i(t)$  can be regarded as a realization of the same random process “traversing a scattering

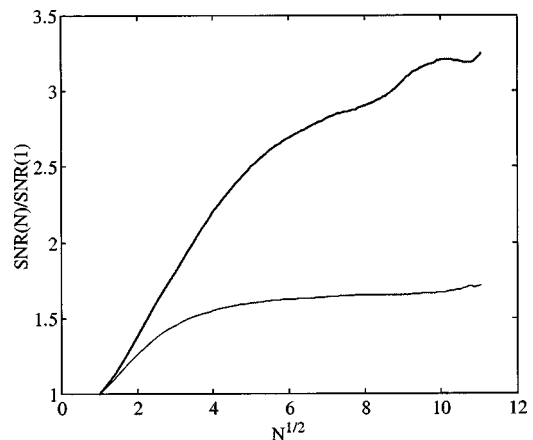


FIG. 11. Time-reversal through  $L = 40 \text{ mm}$  (thick line) and  $L = 80 \text{ mm}$  (thin line): signal-to-noise ratio versus square root of the number of active array elements. The results have been normalized by the SNR obtained for  $N = 1$ .

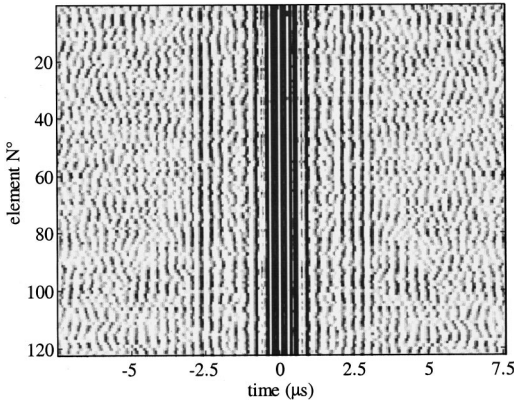


FIG. 12. Contributions  $c_i(t)$  from all elements on the array in a B-scan format, for  $L=80$  mm. The colors have been saturated. On both sides of the central peak, the temporal sidelobes are clearly visible on all elements of the array. The long-range spatial correlation around  $t\sim 2.5$   $\mu$ s from the central peak manifests itself as vertical lines. At later times, the contributions do not seem correlated.

sample with a given density of scatterers.” And each contribution  $c_i(t)$  can be regarded as an estimate of the expected value  $\langle r(t) \rangle$ . When all contributions  $c_i(t)$  ( $i=1,\dots,122$ ) are added, the side lobes are not averaged out; they are inherent to the multiple scattering process within the sample.

For  $L=80$  mm, the long-range correlation between the  $c_i(t)$  is particularly obvious in Fig. 12. The 122 contributions  $c_i(t)$  have been represented in a BSCAN mode: time is in the abscissa, and each line of the picture represents the contribution of one array element to the total signal  $r(t)$ . Among the noise surrounding the central peak, the temporal side lobes that were pointed out earlier are clearly visible on each line, at the same time. When adding all the signals, the temporal side lobes are not averaged out, hence the saturation of the SNR.

The saturation value of the SNR depends on the correlation coefficient between the  $N$  contributions  $c_i(t)$ . We have calculated a resemblance coefficient between  $c_i(t)$  and  $c_j(t)$  as a function of the “distance”  $m=|i-j|$  between the elements using the expression

$$\frac{1}{122-m} \sum_{i=1}^{122-m} \frac{\int_{t_1}^{t_2} c_i(t) c_{i+m}(t) dt}{\sqrt{\int_{t_1}^{t_2} c_i^2(t) dt} \sqrt{\int_{t_1}^{t_2} c_{i+m}^2(t) dt}}. \quad (20)$$

This coefficient was evaluated in 2.5- $\mu$ s-long time windows centered on  $t=0, 2.5, 5$  and  $10$   $\mu$ s. The results are plotted in Fig. 13 for  $L=80$  mm. At time  $t=0$ , all contributions  $c_i(t)$  are almost identical; all contributions are in phase and add constructively. Outside the peak, at time  $t=2.5$   $\mu$ s, the results show a plateau, as could be expected from the B-scan. The height of this plateau is  $\sim 45\%$ . Even at later times where apparently no temporal side lobes were visible on the B-scan, there is still a plateau with a diminishing height: 5% at  $t=5$   $\mu$ s, 3% at  $t=10$   $\mu$ s. These experimental results show that only after  $t\sim 35$   $\mu$ s can the  $c_i(t)$  be considered as spatially uncorrelated.

The height of the correlation plateau has a strong influence on the maximum value of the SNR, and explains why the SNR was only 1.5 times higher with  $N=122$  elements

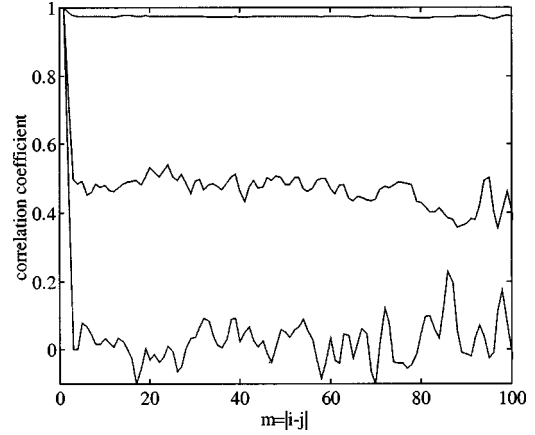


FIG. 13. Resemblance coefficient of the contributions  $c_i(t)$  and  $c_j(t)$ , as a function of  $m=|i-j|$ , at time  $t=0$  (central peak),  $t=2.5$   $\mu$ s and  $t=10$   $\mu$ s. Even 10  $\mu$ s after the peak, a plateau remains with an average value of 3%.

than with one element. Indeed, a simple computation on the sum of  $N$  correlated random variables indicates that when we increase the number of elements active on the array from 1 to  $N$ , the SNR should increase by a factor of

$$\frac{\sqrt{N}}{\sqrt{1 + 2 \sum_{m=1}^{N-1} (1 - m/N) \rho[m]}}, \quad (21)$$

$\rho[m]$  being the correlation coefficient between  $c_i(t)$  and  $c_{i+m}(t)$ . If the  $c_i$  were fully uncorrelated, then  $\rho[m]=0$  and Eq. (21) would give  $\sqrt{N}$ . If there is a short-range correlation with a typical length scale  $m_{\text{coh}}$ , Eq. (21) would grow as  $\sqrt{N}/m_{\text{coh}}$  as soon as  $N > m_{\text{coh}}$ . However, if there is a persistent long-range correlation, with a plateau such that  $\rho[m] = \alpha = \text{constant}$ , then even if  $N \rightarrow \infty$ , Eq. (21) tends to a maximum value  $1/\sqrt{\alpha}$ .

This result has an important consequence: it is useless to increase the number of transducers on a time-reversal mirror beyond a certain limit, depending on  $\alpha$ . This was obvious for  $\alpha=45\%$ : the SNR was only increased by a factor of  $1/\sqrt{0.45} \approx 1.5$  when using all 122 transducers instead of one. But even if the level of the correlation plateau may seem small (a few percent), it can be very limiting for the time-reversal process. When  $0 < \alpha \leq 1$ , with  $N=1/\alpha$  transducers, the SNR attains 70% of its maximum value, and it takes four times as many transducers to reach 90%, and ten times to reach 95%.

To conclude this section, we have seen that the temporal signal-to-noise ratio (SNR) obtained with one transducer was very well predicted by a “shot noise” model, as long as multiple scattering is not too strong (typically, until  $L \approx 4 l^*$ ). At larger thickness, persistent temporal side lobes appear close to the central peak. We interpret these side lobes as the sign of a correlation between the arrival times due to the crossing of multiple scattering paths, which become more frequent as the sample thickness increases. These temporal side lobes degrade the quality of time-reversal focusing. They result in an increasing of the correlation time  $\tau_c$  of the scattered signal, and consequently in a reduction of the number  $\Delta T/\tau_c$  of information grains available in a time win-

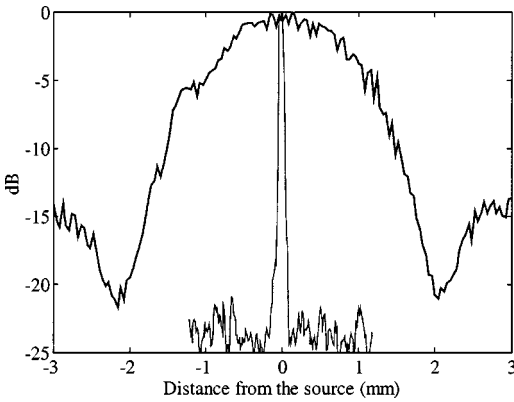


FIG. 14. Directivity pattern of the time-reversed waves around the source position, in water (thin line) and through the rods (thick line), with a 16-element aperture. The sample thickness is  $L=40$  mm. The  $-12$ -dB widths are 33 and 1 mm, respectively. 0.8 vs 22 mm.

dow  $\Delta T$ . However, the shot noise model can be adapted to take this into account, and can predict the SNR correctly even at larger thickness.

When the number of transducers is increased from 1 to  $N$ , the presence of temporal side lobes in all contributions limits the growth of the SNR. This is particularly obvious for thick samples, and at times close to the central peak. In relatively thin samples ( $L \approx 4 l^*$  or less) no temporal side lobes are visible; contributions from different array elements may be correlated, but only over a finite range  $m_{\text{coh}}$ , and consequently the SNR rises as  $\sqrt{N/m_{\text{coh}}}$ . At larger thickness, the temporal side lobes are present on all array elements, inducing a long-range correlation between the contributions from various array elements. This accounts for the saturation of the SNR as  $N$  increases. In the presence of a long-range correlation plateau with height  $\alpha$ , there is no need to have more than  $\sim 4/\alpha$  transducers on the array.

## II. SPATIAL FOCUSING

In a time-reversal experiment, it is also important to check that the recreated pulse is correctly focused around the source. To that end, the piezoelectric element that used to be the source is translated step by step along an axis parallel to the sample ( $x$  axis in Fig. 1), while the same time-reversed signals are continually transmitted through the rods. Thus, the directivity pattern of the time-reversed beam around the source can be measured.

With classical systems, multiple scattering severely degrades the quality of focusing. But with a time-reversal focusing device, experimental results show that the spatial resolution of the system is enhanced by the presence of multiple scattering, compared to the resolution of the array in water. This was demonstrated experimentally in 1995:<sup>5</sup> the resolution of a time-reversal mirror in a multiple scattering medium was found to be one-sixth of the theoretical limit for the mirror's aperture.

A similar result is presented in Fig. 14: a 16-element aperture was used through a 40-mm thick sample. As before, the source and the array were 31 cm apart, and the center frequency was 3.2 MHz. When we compare the results, we find that the spatial resolution is roughly 30 times finer

through a high-order scattering sample than through water! Precisely, the respective widths of the directivity patterns are 0.8 and 22 mm at  $-6$  dB, and 1 and 33 mm at  $-12$  dB.

A physical explanation for this surprising effect is that, due to the presence of the scatterers, high spatial frequencies that would have been lost otherwise are redirected and recorded on the array. In other words, when the time-reversed waves propagate back, the medium acts as a lens that focuses the signal on the source; the angular aperture of that pseudolens is larger than that of the array alone, hence the improvement in resolution.

A more theoretical argument can be developed to account for this result. For simplicity, imagine that the time-reversal operation is performed on one element  $R_i$ , the  $i$ th element on the array. As usual, when the source  $S$  sends a pulse, the detector  $R_i$  records the scattered signal  $h_i(t)$ . The time-reversed signal  $h_i(-t)$  is sent back by  $R_i$  and propagates through the same medium. At the source position, as usual, the recreated signal is

$$h_i(t) \otimes h_i(-t) \quad (22)$$

and is maximum at time  $t=0$ .

If we consider another observing point  $O$ , somewhere around the source, and denote by  $h'_i(t)$  the propagation impulse response from  $R_i$  to  $O$ , the signal recreated outside the source at time  $t=0$  is

$$\int h_i(t) h'_i(t) dt. \quad (23)$$

Now, due to reciprocity, the source  $S$  and the receiver  $R_i$  can be exchanged: in other words,  $h_i(t)$  is also the signal that would be received in  $S$  if the source was in  $R_i$ . Therefore, we can imagine a situation in which  $R_i$  is a source, and the transmitted wave field is observed at two points  $O$  and  $S$  on the other side of the sample. The spatial correlation function of this wave field would be  $\langle h_i(t) h'_i(t) \rangle$ .

Interestingly, Eq. (23) can be viewed as an estimator of this spatial correlation function. In that sense, the spatial resolution of the system (i.e., the  $-6$  dB of the directivity pattern) is simply an estimate of the correlation length of the wave field transmitted through the scattering sample.

This has an important consequence. Indeed, if the resolution of the system essentially depends on correlation properties of the scattered wave field, it should become independent from the array's aperture. This is confirmed by the experimental results. Figure 15 presents the directivity patterns obtained through a 40-mm-thick multiple scattering sample, using either 1 array element or the whole array (122 elements) as a time-reversal mirror. In both cases, the spatial resolution is the same:  $\sim 0.85$  mm at  $-6$  dB. In total contradiction with what happens in a homogeneous medium, enlarging the aperture of the array does not change the spatial resolution. Such an effect was already observed in time-reversal experiments performed in a wave guide,<sup>11</sup> and was interpreted with the image theorem. The situation is very different here since there are neither reflecting interfaces nor virtual images of the array but a random scattering sample.

However, even though the number  $N$  of active array elements does not influence the typical width of the focal spot,

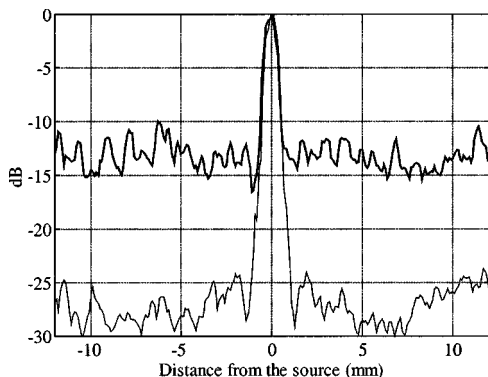


FIG. 15. Directivity pattern of the time-reversed waves around the source position through  $L=40$  mm, with  $N=122$  transducers (thin line) and  $N=1$  transducer (thick line). The  $-6$ -dB resolutions are  $0.84$  and  $0.9$  mm, respectively.

it has a strong impact on the background level of the directivity pattern ( $\sim -12$  dB for  $N=1$ ,  $\sim -28$  dB for  $N=122$ ), as can be seen in Fig. 15.

With an  $N$ -element time-reversal mirror, the signal recreated at the observing point is a sum of  $N$  contributions:

$$r'(t) = \sum_{i=1}^N c'_i(t), \quad (24)$$

where

$$c'_i(t) = h_i(-t) \otimes h'_i(t). \quad (25)$$

Each  $c'_i(t)$  is an estimate of the intercorrelation of signals transmitted from the  $i$ th array element to two points  $S$  (the original source position) and  $O$  (the observing point outside the source). Once again, the amelioration of the background level of the directivity patterns with  $N$  depends on whether the contributions  $c'_i(t)$  are correlated or not. Experimental results (Fig. 16) indicate that the background level decreases until  $N \sim 64$ , then seems to saturate. However, given the saturation value ( $\sim -28$  dB) we cannot conclude whether this

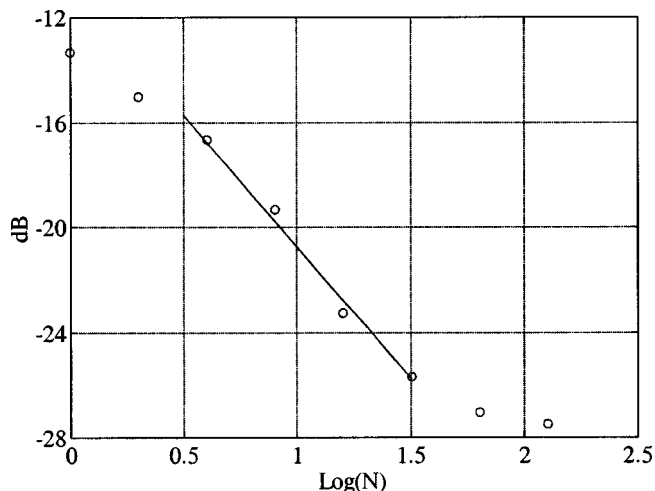


FIG. 16. Background level of the directivity pattern (in dB) versus the number  $N$  of transducers. In the central part of the plot (linear fit), as  $N$  is multiplied by 10, the background decreases by 10 dB. But for large values of  $N$ , the curve seems to saturate around  $-27$  dB.

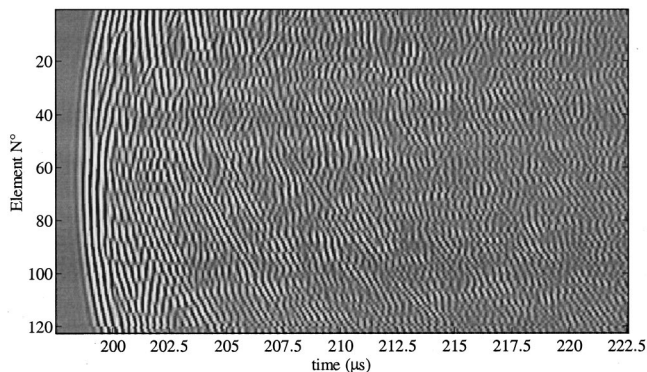


FIG. 17. B-scan of the 122 signals received on the array through a multiple scattering sample ( $L=15$  mm). The ballistic front is clearly visible for this thickness. It arrives first, and is followed by scattered waves.

saturation is due to a long-range correlation between the  $c'_i(t)$  or to classical limitations of the array (spatial and temporal sampling).

We have argued that the spatial resolution of a time-reversal device essentially depends on the correlation length of the scattered wave field: the shorter the coherence length, the finer the resolution. This can be illustrated by a time-reversal experiment through a sample of width  $L=15$  mm. The scattered signals received on the array are represented in Fig. 17 in a B-scan format. In a relatively thin sample (here,  $L \sim 4 l^*$ ) such as this one, the so-called “ballistic” front is clearly visible; it arrives first on each element of the array and represents the remnant, unscattered part of the incoming wave. Therefore, at early times, the signals received on various elements of the array are strongly correlated. After the ballistic front has arrived, the array continues receiving signals corresponding to scattered waves. Late arrivals correspond to longer scattering paths and higher order of scattering.

Figure 18 shows the directivity patterns that were obtained with a 64-element aperture by time-reversing either the early arrivals, or later arrivals. We selected two  $2$ - $\mu$ s time-window, the first one including the ballistic front, the second one beginning  $100$   $\mu$ s later, i.e., containing purely multiple scattering contributions. The  $-6$ -dB resolutions we

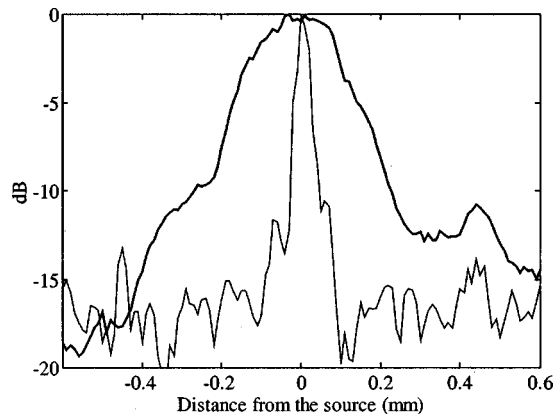


FIG. 18. Directivity patterns obtained after time-reversing a  $2$ - $\mu$ s time window at early times (ballistic front) or at later times (multiple scattering contribution). The spatial resolution is  $3.6$  mm at early times,  $0.5$  mm at later times.

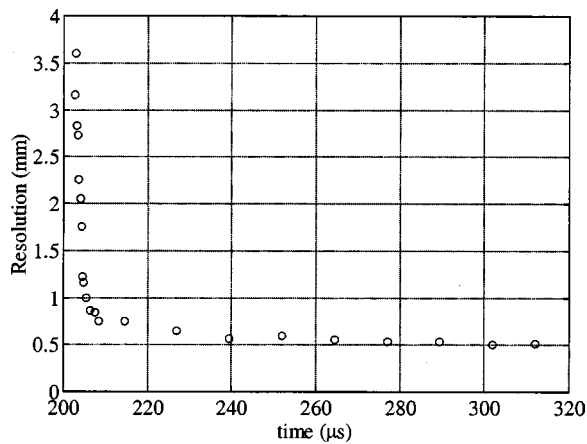


FIG. 19. Spatial resolution versus the beginning of the time-reversal window, for 2- $\mu$ s time-windows through  $L=15$  mm.

obtained are 3.6 mm for early arrivals and only 0.5 mm for late arrivals (it should be noted that the average wavelength is 0.45 mm). It appears that at early times, as in a homogeneous medium, the resolution of the system is limited by the aperture of the array, whereas at later times multiple scattering dominates, and the resolution is essentially limited by the coherence length of the scattered field.

The evolution of the spatial resolution as a function of time is plotted in Fig. 19. Right after the ballistic front, the resolution diminishes very sharply down to  $\sim 0.5$  mm, which seems to be the smallest value that can be attained given the average wavelength (0.45 mm) and the size of our elements (0.38 mm).

To obtain the finest possible resolution, the scattering medium should be arranged in such a way that the correlation length of the scattered field is as small as possible. But a ‘‘spatial white noise’’ is not physically possible: the coherence length of a propagating wave field cannot be arbitrarily small, it is at least half a wavelength. Moreover, due to the propagation in water from the scattering sample to the array, the signal received at one point must be at least partially correlated with its neighbors, even if the scattered field just emerging from the sample could be a true ‘‘spatial white noise.’’ This is based on a fundamental theorem of statistical optics, the Van Cittert–Zernike theorem:<sup>12</sup> it states that the wave field generated by a fully incoherent random source has a typical correlation length  $\lambda z/D$ ,  $D$  denoting the size of the random source and  $z$  the propagation distance. Here the multiple scattering ‘‘halo’’ emerging from the sample can be seen as a random source. As the distance  $z$  between the sample and the receiver increases, the correlation length increases, too. A physical interpretation of the Van Cittert–Zernike theorem is that propagation in water smooths the rugged wave fronts emerging from the sample. So, in theory, one way to improve the resolution would be to put the transducers as close to the sample as possible. In practice, this is difficult though because of back and forth reflections between the transducers and the scattering sample.

Apparently, a satisfying compromise was found in our experiments. The distances between the sample and the source/array were typically between 10 and 20 cm. In the highest-order multiple scattering contributions, resolutions

down to 0.5 mm were attained. The finest resolutions are obtained by time-reversing the last part of the scattered signals. However, since the amplitudes of the transmitted signals decrease with time, the later time contributions have relatively small amplitudes and are more sensitive to noise, which generates a higher background level in the directivity pattern of Fig. 18.

The background level can be easily decreased if we enlarge the time-reversal window; but then we have to include earlier arrivals, for which the resolution is not as fine. A trade-off has to be found: as we showed in Fig. 15, a time-reversal mirror through a high-order multiple scattering is able to have both a fine resolution (0.85 mm, i.e., 2 wavelengths) and a rather flat background level around  $-28$  dB.

### III. CONCLUSION

A transducers array used as a time-reversal mirror is a very efficient and robust device to focus a wave in a inhomogeneous medium. It takes advantage of high-order multiple scattering to actually enhance the resolution of the array. As an example, with a 16-element aperture, the resolution was found to be 30 times finer than in a homogeneous medium. The aim of this paper was to study by how much the quality temporal and spatial focusing varies when the time-reversal window or the aperture of the array are changed.

The quality of temporal focusing depends on the correlation time of the scattered signals. A ‘‘shot noise’’ model was developed, and its predictions were found to be in excellent agreement with experimental results for  $L \sim 4 l^*$ . At larger thickness, the model had to be modified to incorporate additional correlation that we interpreted as a consequence of the increasing number of crossing paths as the sample thickness increases.

As to spatial focusing, the resolution of the system essentially depends on the correlation length of the scattered waves. The more scrambled the waves, the better the resolution. Experimental results showed that with the highest order of multiple scattering, resolutions of the order of the wavelength could be attained.

Finally, the fundamental properties of time-reversal in a random medium lie on the fact that it is a correlator, both in space and in time. The time-reversed waves can be viewed as an estimate of the space and time correlation functions of the waves scattered by a random medium. The quality of time-reversal is the same as that of a statistical estimate: the estimate is better if it is built from a larger number of uncorrelated information, what we termed ‘‘information grains.’’

To obtain the best possible focusing, one would have to design a random medium for which the correlation time and correlation length of the scattered waves could be as small as possible. The solution is not necessarily a very thick medium with very high-order multiple scattering. Indeed, the correlation time was shown to be larger for thicker samples. As to the correlation length, physically, it cannot be smaller than half a wavelength.

So far, the experimental results presented in Figs. 14 and 15 showed that a time-reversal mirror coupled to a high-order multiple scattering medium is able to achieve a fine



resolution (0.85 mm, i.e., 2 wavelengths) and a flat background level (around  $-27$  dB). Designing a random disordered medium to obtain better results with the same array remains a challenge.

- <sup>1</sup>M. Fink, "Time-reversed acoustics," *Phys. Today* 34–40 (March 1997).
- <sup>2</sup>V. Miette, L. Sandrin, F. Wu, and M. Fink, "Optimization of time reversal processing in titanium inspections," *IEEE Ultrasonics Symposium Proceedings*, San Antonio, November 1996, Vol. 1, pp. 643–647.
- <sup>3</sup>W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, "Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror," *J. Acoust. Soc. Am.* **103**, 25–40 (1998).
- <sup>4</sup>M. Tanter, J.-L. Thomas, and M. Fink, "Focusing and steering through absorbing and aberrating layers: Application to ultrasonic propagation through the skull," *J. Acoust. Soc. Am.* **103**, 2792 (A) (1998).
- <sup>5</sup>J.-L. Thomas, F. Wu, and M. Fink, "Time reversal mirror applied to lithotripsy," *Ultrason. Imaging* **18**, 106–121 (1996).

- <sup>6</sup>A. Derode, P. Roux, and M. Fink, "Robust acoustic time reversal with high-order multiple scattering," *Phys. Rev. Lett.* **75**(23), 4206 (1995).
- <sup>7</sup>A. Derode, A. Tourin, and M. Fink, "Ultrasonic pulse compression with one-bit time reversal through high-order multiple scattering," *J. Appl. Phys.* **85**(9), 6343–6352 (1999).
- <sup>8</sup>A. Papoulis, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, New York, 1965), p. 358.
- <sup>9</sup>F. Scheffold, W. Härtl, G. Maret, and E. Matijevic, "Observation of long-range correlations in temporal intensity fluctuations of light," *Phys. Rev. B* **56**, 10942–10952 (1997).
- <sup>10</sup>A. Tourin, A. Derode, P. Roux, B. A. Van Tiggelen, and M. Fink, "Time-dependent coherent backscattering of acoustic waves," *Phys. Rev. Lett.* **79**(19), 3637–3639 (1997).
- <sup>11</sup>P. Roux, B. Roman, and M. Fink, "Time-reversal in an ultrasonic waveguide," *Appl. Phys. Lett.* **70**(14), 1811–1813 (1997).
- <sup>12</sup>J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985).

# Analytical regularization based analysis of a spherical reflector symmetrically illuminated by an acoustic beam

Sergey S. Vinogradov<sup>a)</sup> and Elena D. Vinogradova<sup>b)</sup>

*Department of Mathematics, University of Dundee, Dundee DD1 4HN, Scotland, United Kingdom*

Alexander I. Nosich

*Institute of Radiophysics and Electronics of the National Academy of Sciences,  
Ul. Proskury 12, Kharkov 310085, Ukraine*

Ayhan Altıntaş

*Department of Electrical and Electronics Engineering, Bilkent University, 06533 Ankara, Turkey*

(Received 3 April 1999; revised 16 February 2000; accepted 18 February 2000)

A mathematically accurate and numerically efficient method of analysis of a spherical reflector, fed by a scalar beam produced by a complex source-point feed, is presented. Two cases, soft and hard reflector surface, are considered. In each case the solution of the full-wave integral equation is reduced to dual series equations and then further to a regularized infinite-matrix equation. The latter procedure is based on the analytical inversion of the static part of the problem. Sample numerical results for  $50\text{-}\lambda$  reflectors demonstrate features that escape a high-frequency asymptotic analysis.

© 2000 Acoustical Society of America. [S0001-4966(00)00306-4]

PACS numbers: 43.20.Fn [ANN]

## INTRODUCTION

Most frequently, performance of a reflector antenna is predicted by using asymptotic high-frequency techniques such as Physical Optics (PO), combined with Geometrical Theory of Diffraction (GTD),<sup>1,2</sup> for the off-beam radiation. The Method-of-Moments (MoM) is also used in the integral-equation (IE) analysis of reflectors of small to moderate size in terms of wavelength.<sup>3,4</sup> The merits and limitations of both approaches are well known. In spite of their flexibility, PO or GTD alone is not uniformly accurate with respect to the direction in space, and both fail to characterize smaller reflectors. MoM algorithms for the full-wave IE become computationally expensive for larger reflectors, due to either large matrices or a large time for filling the matrix. Besides, not every MoM approximation scheme is convergent to the exact result as the number of equations is increased, in the sense that the computation error cannot be progressively minimized.

The feed field is normally simulated via a Gaussian beam or a spherical-wave expansion multiplied with an angular window function. Commonly it is neglected by these simulations that such a feed field function does not solve the Helmholtz equation exactly, although the radiated or scattered field is found as a solution of the full-wave integral equation. It has been proposed therefore to use the complex source-point (CSP) beam, or combination of such beams, as a feed field.<sup>5,6</sup> Then, the latter is an exact solution to the Helmholtz equation at every point in the physical observation space. In Refs. 7, 8, this concept was combined with PO and GTD for a characterization of a spherical-wave scalar

beam scattering from a circular aperture. This concept is further developed in Refs. 9, 10, which contain practical and useful results.

In a recent study,<sup>11</sup> it was demonstrated, for a 2D problem, that a very accurate and computationally efficient analysis of reflector antennas can be achieved by using a combination of the CSP method to simulate a beam-like feed field, and an analytical regularization based technique to solve the reflector scattering. In Ref. 11, a circular-cylindrical reflector fed by the CSP-type line sources was considered. First, the IE was discretized into the dual series equations (DSE) in terms of the entire-period angular exponents. Then, the static part of DSE was analytically inverted, using the Riemann–Hilbert Problem solution, resulting in an infinite matrix Fredholm equation of the second kind. A remarkable feature of this equation is that it can be solved numerically with a guaranteed accuracy and small CPU time expenditures.

Our present study is similar to Ref. 11, but deals with a quasi-3D problem of a spherical reflector. Acoustic reflectors of this type have been analyzed in Refs. 12, 13 assuming finite transparency of reflector material and tapered spherical wave illumination. Unlike Refs. 12, 13, we characterize the feed by a scalar CSP beam, assume the reflector to be perfectly hard or soft, and develop a numerically exact solution. It is obtained from the DSE in terms of Legendre polynomials, with the static part inversion based on the Abel integral equation technique.<sup>14,15</sup> This solution is equivalent to a judicious choice of the expansion functions in an MoM procedure, forming a set of the orthogonal eigenfunctions of the static limit of the IE kernel. The usage of a directive CSP field as a feed brings new features, hence we believe that an in-depth physical analysis of two scalar problems of soft and hard reflectors is reasonable. To obtain the results which are applicable to paraboloidal reflectors as well, we restrict the computed examples to “dish” reflectors with a large to mod-

<sup>a)</sup>Visiting Scientist from Kharkov Academy of Civil Engineering, Kharkov 310002, Ukraine.

<sup>b)</sup>Visiting Scientist from Institute of Radiophysics and Electronics NAS, Kharkov 310085, Ukraine.

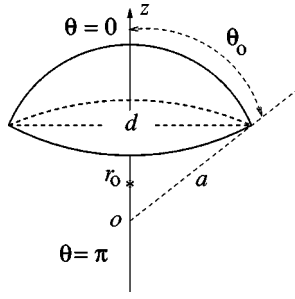


FIG. 1. Geometry of a spherical reflector fed by a CSP feed.

erate  $f/d$  ratio [ $f$  and  $d$  being the Geometrical Optics (GO) focal distance and the dish diameter, respectively].

The remainder of the paper is organized as follows. In Sec. I, we formulate the boundary-value problem for soft and hard reflectors, and derive a rigorous IE for each case. Further we discretize and convert it to the DSE. In Sec. II, we present basic points of the partial inversion of the DSE and reduction to the regularized matrix equation. Section III contains the formulas for far field characteristics. Section IV presents the results of numerical analysis concerning radiation patterns and the directivity of the soft and hard spherical reflectors. Conclusions of the presented work are summarized in Sec. V. Focal shifts in parabolic reflectors. A note should be made that the time dependence is assumed as  $e^{-i\omega t}$  and is omitted.

## I. FORMULATION AND BASIC EQUATIONS

### A. Problem formulation

Consider a zero-thickness, perfectly soft or perfectly hard, spherical reflector of radius  $a$  and angular width  $2\theta_0$ , symmetrically excited by the field of a CSP beam. That is, the feed is located at the point  $(r_s, 0, 0)$  with the radial source coordinate being a complex value:  $r_s = r_0 + ib$ . The geometry of the problem is shown in Fig. 1. The incident scalar wave field is:

$$U^0(r, \theta) = e^{ikR/R}, \quad (1)$$

where  $R = (r^2 - 2rr_s \cos \theta + r_s^2)^{1/2}$ , and  $k = \omega/c$  is the real-valued free-space wave number ( $c$  being the sound propagation velocity). The scattered field  $U^{sc}(r, \theta)$  is the solution of a boundary-value problem for the 3D Helmholtz equation, with the boundary condition of either soft or hard type at the reflector surface  $M$ : ( $r = a, 0 \leq \theta < \theta_0, 0 \leq \varphi \leq 2\pi$ )

$$U^0 + U^{sc}|_M = 0, \quad \left. \frac{\partial(U^0 + U^{sc})}{\partial r} \right|_M = 0. \quad (2)$$

The formulation must also include (i) the edge condition  $U^{sc} \sim O(\rho^{1/2})$ ,  $\partial U^{sc}/\partial r \sim O(\rho^{-1/2})$ , where  $\rho \rightarrow 0$  is the distance from the dish rim, and (ii) the outgoing radiation condition at  $r \rightarrow \infty$ , to ensure the solution uniqueness.<sup>16</sup>

### B. Acoustically soft reflector

Consider first the case of a perfectly soft reflector. We seek the scattered field function as a single-layer potential:

$$U^{sc}(\mathbf{r}) = -\frac{1}{4\pi} \int \int_M j_s(\theta', \varphi') \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} dS', \quad (3)$$

where  $|\mathbf{r}-\mathbf{r}'| = \{r^2 + a^2 - 2ra[\cos \theta' \cos \theta + \sin \theta' \sin \theta \times \cos(\varphi - \varphi')]\}^{1/2}$  is the distance between the observation point and a point at the reflector surface. Note that the unknown density function is related to the jump of the scattered field normal derivative across the reflector as

$$j_s(\theta', \varphi') = \left. \frac{\partial U^{sc}}{\partial r} \right|_{r=a+0} - \left. \frac{\partial U^{sc}}{\partial r} \right|_{r=a-0}. \quad (4)$$

The soft-surface boundary condition Eq. (2) then yields a Fredholm first kind IE as

$$\begin{aligned} \frac{a}{8\pi} \int_0^{\theta_0} \int_0^{2\pi} j_s(\theta', \varphi') \frac{e^{i2ka|\sin(\psi/2)|}}{|\sin(\psi/2)|} \sin \theta' d\theta' d\varphi' \\ = U^0(a, \theta, \varphi), \end{aligned} \quad (5)$$

where  $\psi = \arccos[\cos \theta' \cos \theta + \sin \theta' \sin \theta \cos(\varphi - \varphi')]$ .

Instead of approximating IE Eq. (5) by MoM with sub-domain or  $M$ -domain basis functions, we further discretize it in terms of a complete set of orthogonal functions in the global domain  $0 \leq \theta \leq \pi$ . In our case of a  $\varphi$ -independent solution, such a set is formed by the Legendre polynomials  $P_n(\cos \theta)$  ( $n = 0, 1, 2, \dots$ ). By extending the density function to be identically by zero on the complementary surface ( $r = a, \theta_0 < \theta \leq \pi$ ), as is natural due to Eq. (4), we assume that for all  $\theta'$

$$j_s(\theta') = -\frac{1}{ka^2} \sum_{n=0}^{\infty} x_n^s (2n+1) P_n(\cos \theta'), \quad (6)$$

with the expansion coefficients  $x_n^s$ ,  $n = 0, 1, 2, \dots$  to be found.

Besides, it is known that the free-space Green's function can be expanded as:<sup>15</sup>

$$\begin{aligned} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} &= ik \sum_{m=0}^{\infty} (2 - \delta_{0m}) \cos m(\varphi - \varphi') \\ &\times \sum_{n=m}^{\infty} (2n+1) \frac{(n-m)!}{(n+m)!} \\ &\times \begin{cases} h_n^{(1)}(ka) j_n(kr), & r < a \\ j_n(ka) h_n^{(1)}(kr), & r > a \end{cases} \\ &\times P_n^m(\cos \theta) P_n^m(\cos \theta'), \end{aligned} \quad (7)$$

where  $P_n^m$  are the associated Legendre functions. This yields the following series representation of the kernel function of the IE Eq. (5):

$$\begin{aligned} \frac{e^{i2ka|\sin(\psi/2)|}}{|\sin(\psi/2)|} &= i2ka \sum_{m=0}^{\infty} (2 - \delta_{0m}) \cos m(\varphi - \varphi') \\ &\times \sum_{n=m}^{\infty} (2n+1) \frac{(n-m)!}{(n+m)!} j_n(ka) h_n^{(1)}(ka) \\ &\times P_n^m(\cos \theta) P_n^m(\cos \theta'). \end{aligned} \quad (8)$$

Now the integration in Eq. (5) over  $\theta'$  can be extended from 0 to  $\pi$ . This enables us to use orthogonality of the spherical harmonics

$$\int_0^{2\pi} \cos m(\varphi - \varphi') \int_0^\pi P_s(\cos \theta') P_n(\cos \theta') \sin \theta' d\theta' d\varphi' = \frac{4\pi(s+m)!}{(2n+1)(s-m)!} \delta_{ns} \delta_{m0} \quad (9)$$

when discretizing the IE. Together with the specification that the density function Eq. (6) is zero off the reflector, this brings us to the dual series equations (DSE):

$$\sum_{n=0}^{\infty} x_n^s (2n+1) j_n(ka) h_n^{(1)}(ka) P_n(\cos \theta) = \sum_{n=0}^{\infty} b_n^s P_n(\cos \theta), \quad 0 \leq \theta < \theta_0, \quad (10)$$

$$\sum_{n=0}^{\infty} x_n^s (2n+1) P_n(\cos \theta) = 0, \quad \theta_0 < \theta \leq \pi,$$

where the right-hand side coefficients are determined by the CSP feed field as

$$b_n^s = (2n+1) j_n(kr_s) h_n^{(1)}(ka). \quad (11)$$

Once again, the DSE Eq. (10) can be attacked by a "brute force" numerical solution with a direct MoM scheme. Although the results are generally meaningful, only a few correct digits can be obtained, and there is no possibility of increasing accuracy by taking a greater number of collocation points. That is our motivation for regularizing the DSE Eq. (10), to obtain an algorithm convergent to the exact solution in a pointwise manner.

### C. Acoustically hard reflector

For a perfectly hard spherical reflector fed by a CSP feed, we seek the scattered field function as a double-layer potential:

$$U^{\text{sc}}(\mathbf{r}) = -\frac{1}{4\pi} \int \int_M j_h(\theta', \varphi') \frac{\partial}{\partial r'} \left\{ \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right\} dS', \quad (12)$$

where the  $\varphi$ -independent density function is now related to the jump of the field across the reflector surface:

$$j_h(\theta') = U^{\text{sc}}(a+0, \theta') - U^{\text{sc}}(a-0, \theta'). \quad (13)$$

The hard-surface boundary condition Eq. (2) now yields a hypersingular IE for this function:

$$\frac{a^2}{4\pi} \frac{\partial}{\partial r} \int_0^{\theta_0} \int_0^{2\pi} j_h(\theta', \varphi') \frac{\partial}{\partial r'} \left\{ \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right\} \sin \theta' d\theta' d\varphi' = \frac{\partial}{\partial r} U^0(r, \theta). \quad (14)$$

In order to discretize the IE Eq. (14), we assume that the density function is extended, so that it is identically zero on the rest of the sphere of radius  $a$ ; we expand the density

function in terms of the same complete set of Legendre polynomials as were used in the "soft case":

$$j_h(\theta') = \frac{1}{k^2 a} \sum_{n=0}^{\infty} x_n^h P_n(\cos \theta'). \quad (15)$$

We substitute this series into the IE Eq. (14) and integrate, making use of the orthogonality properties Eq. (9). Moreover, we explicitly enforce the vanishing of  $j_h(\theta')$  off the reflector surface, and so arrive at the DSE as

$$\sum_{n=0}^{\infty} x_n^h j_n'(ka) h_n^{(1)'}(ka) P_n(\cos \theta) = - \sum_{n=0}^{\infty} b_n^h P_n(\cos \theta), \quad 0 \leq \theta < \theta_0, \quad (16)$$

$$\sum_{n=0}^{\infty} x_n^h P_n(\cos \theta) = 0, \quad \theta_0 < \theta \leq \pi,$$

where

$$b_n^h = j_n(kr_s) h_n^{(1)'}(ka). \quad (17)$$

Note that the DSE so obtained are of the same type as in the case of the soft reflector.

## II. PARTIAL INVERSION OF DSE

We shall regularize Eqs. (10) and (16) by performing an analytical inversion of the static part of the DSE. To extract the static parts, use the power series for the spherical Bessel functions,<sup>17</sup> from which it follows that:

$$j_n(ka) = \frac{n!(2ka)^n}{(2n+1)!} \left\{ 1 + O\left(\frac{k^2 a^2}{n}\right) \right\}, \quad (18)$$

$$h_n^{(1)}(ka) = -i \frac{(2n)!}{n!(2ka)^{n+1}} \left\{ 1 + O\left(\frac{k^2 a^2}{n}\right) \right\}.$$

This enables us to show that if  $n \gg ka$ , then

$$(2n+1) j_n(ka) h_n^{(1)}(ka) \sim -i/(ka), \quad (19)$$

$$(2n+1)^{-1} j_n'(ka) h_n^{(1)'}(ka) \sim i/(4k^3 a^3).$$

Based on these estimates, we introduce two coefficient sets as

$$\varepsilon_n^s = 1 - ika(2n+1) j_n(ka) h_n^{(1)}(ka), \quad (20)$$

$$\varepsilon_n^h = 1 + i4(ka)^3 (2n+1)^{-1} j_n'(ka) h_n^{(1)'}(ka).$$

Note that all  $\varepsilon_n^s$  behave as  $O(k^2 a^2 n^{-2})$  for larger  $n$ , or, equivalently, for smaller  $ka$ , while  $\varepsilon_n^h$  behave as  $(2n+1)^{-2}$ . The DSE may be written as

$$\sum_{n=0}^{\infty} x_n^s P_n(\cos \theta) = \sum_{n=0}^{\infty} (\varepsilon_n^s x_n^s - ikab_n^s) P_n(\cos \theta), \quad 0 \leq \theta < \theta_0, \quad (21)$$

$$\sum_{n=0}^{\infty} (2n+1) x_n^s P_n(\cos \theta) = 0, \quad \theta_0 < \theta \leq \pi,$$

and

$$\begin{aligned}
& \sum_{n=0}^{\infty} (2n+1)x_n^h P_n(\cos \theta) \\
&= \sum_{n=0}^{\infty} (2n+1)(x_n^h \varepsilon_n^h + i4k^3 a^3 b_n^h) P_n(\cos \theta), \\
& \quad 0 \leq \theta < \theta_0, \\
& \sum_{n=0}^{\infty} x_n^h P_n(\cos \theta) = 0, \quad \theta_0 < \theta \leq \pi.
\end{aligned} \tag{22}$$

Analytical inversion of the left-hand side the DSE is performed by transforming it to a single function, defined on the complete interval  $[0, \pi]$  of  $\theta$  variation. This is done by reducing each of the functional equations of the DSE to an Abel IE (see Refs. 12, 13), which has a known inversion formula. Here we use the Mehler–Dirichlet formulas for the Legendre polynomials:

$$\begin{aligned}
P_n(\cos \theta) &= \int_0^\theta \frac{\cos(n+1/2)\gamma}{(\cos \gamma - \cos \theta)^{1/2}} d\gamma \\
&= \int_\theta^\pi \frac{\sin(n+1/2)\gamma}{(\cos \theta - \cos \gamma)^{1/2}} d\gamma.
\end{aligned} \tag{23}$$

This enables us to integrate the second equation of Eqs. (21) and the first of Eqs. (22), and reduce each of the DSE to the same function of  $\theta$ , given by its piecewise Fourier expansion on  $[0, \pi]$ :

$$\begin{aligned}
& \sum_{n=0}^{\infty} x_n^s \cos(n+1/2)\theta \\
&= \begin{cases} \sum_{n=0}^{\infty} (x_n^s \varepsilon_n^s + B_n^s) \cos(n+1/2)\theta, & 0 \leq \theta < \theta_0 \\ 0, & \theta_0 < \theta \leq \pi \end{cases},
\end{aligned} \tag{24}$$

$$\begin{aligned}
& \sum_{n=0}^{\infty} x_n^h \sin(n+1/2)\theta \\
&= \begin{cases} \sum_{n=0}^{\infty} (x_n^h \varepsilon_n^h + B_n^h) \sin(n+1/2)\theta, & 0 \leq \theta < \theta_0 \\ 0, & \theta_0 < \theta \leq \pi \end{cases}.
\end{aligned} \tag{25}$$

Here we have denoted

$$B_n^s = -ikab_n^s, \quad B_n^h = i4(ka)^3 b_n^h. \tag{26}$$

Using orthogonality and completeness of the functions  $\cos(n+1/2)\theta$  or  $\sin(n+1/2)\theta$ ,  $n=(0), 1, 2, \dots$  at the interval  $(0, \pi)$ , produces a regularized infinite-matrix equation:

$$x_m^{h,s} = \frac{1}{\pi} \sum_{n=0}^{\infty} (x_n^{h,s} \varepsilon_n^{h,s} + B_n^{h,s}) S_{mn}^{s,h}(\theta_0), \quad m=0, 1, 2, \dots, \tag{27}$$

where

$$S_{mn}^{s,h}(\theta_0) = \frac{\sin(n-m)\theta_0}{n-m} \pm \frac{\sin(n+m+1)\theta_0}{n+m+1} \tag{28}$$

with the upper and lower sign for the soft and hard case, respectively.

Taking account of the large-index behavior of  $\varepsilon_n^s$  or  $\varepsilon_n^h$ , it is easy to verify that the absolute squared norm  $N_A^2 = \sum_{m,n=0}^{\infty} |\varepsilon_n^{s,h}(ka) S_{mn}^{s,h}(\theta_0)|^2 \leq \text{const} \sum_{n=0}^{\infty} |\varepsilon_n^{s,h}|^2$  is finite. This is enough to conclude that the matrix operator of Eq. (27) is of Fredholm second kind type in the space of the square summable sequences  $l^2$ . Besides, the right-hand side of Eq. (27) belongs to  $l^2$ , provided that  $|r_s| < a$ , or, more precisely, if the real-space branch-cut associated with the CSP feed (see Refs. 5–8) does not touch or cross the reflector. Under such a condition, the Fredholm theorems are valid:<sup>16</sup> due to uniqueness, the exact solution of the infinite-matrix equation (27) exists in  $l^2$ . Moreover, it can be shown that the solution satisfies the edge condition: without going into details we point out the connection with the square-root denominators in the integrands of Eq. (23). For computational purposes the most important consequence of the regularization procedure is that, the greater the truncation order of Eq. (27), the closer the numerical solution will be to the exact one. The convergence here is of pointwise-type, not of mean-type, or of some other “weak” form. Note that the matrix elements are remarkably simple, and need no numerical integrations. If  $k=0$ , all the coefficients  $\varepsilon_n^s$  vanish, showing that in the static case Eq. (27) delivers an exact analytical solution. In the case of the hard boundary condition,  $\varepsilon_n^h = (2n+1)^{-2}$  in the limit  $k \rightarrow 0$ . This means that by introducing new coefficients  $\tilde{\varepsilon}_n^h = \varepsilon_n^h - (2n+1)^{-2}$ , one can also obtain an exact analytical solution of the static-counterpart of the hard-surface problem. However, as spherical reflectors are normally used with  $ka \gg 1$ , such a procedure is not necessary in our analysis.

### III. FAR FIELD CHARACTERISTICS

After determining the coefficients  $x_n^{s,h}$  from Eq. (27), one can easily find, with the same guaranteed accuracy, the density function, the far field pattern, the total radiated power, and the directivity (which, in our lossless analysis, is the same as gain). All of these functions and parameters are expressed in terms of series depending on  $x_n^{s,h}$ . For example, the far field pattern is found as

$$\Psi^{s,h}(\theta) = e^{k(b-ir_0)\cos \theta} + \sum_{n=0}^{\infty} (-i)^n w_n^{s,h} y_n^{s,h} P_n(\cos \theta), \tag{29}$$

where we denote  $y_n^s = j_n(ka)x_n^s$ ,  $y_n^h = j_n'(ka)x_n^h$ , and  $w_n^s = 2n+1$ ,  $w_n^h = 1$ .

Due to completeness and orthogonality of the expansion functions on the unit sphere, integration of the time-average far-zone power flux is performed analytically, yielding

$$\begin{aligned}
P_{\text{tot}}^s &= P_0 + P_0 \frac{2kb}{\sinh 2kb} \\
&\times \sum_{n=0}^{\infty} (2n+1) \{ |y_n^s|^2 + 2 \text{Re}[y_n^s j_n(kr_s)] \},
\end{aligned} \tag{30}$$

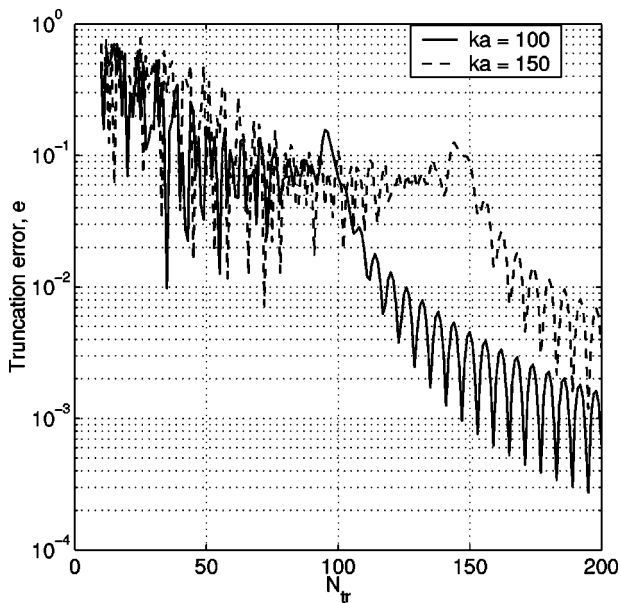


FIG. 2. Normalized computation error as a function of the matrix truncation number, for a hard-surface reflector  $\theta_0 = 30^\circ$ .

$$P_{\text{tot}}^h = P_0 + P_0 \frac{2kb}{\sinh 2kb} \sum_{n=0}^{\infty} \{(2n+1)^{-1} |y_n^h|^2 + 2 \operatorname{Re}[y_n^h j_n(kr_s)]\}. \quad (31)$$

The directivity is

$$D_{\text{tot}}^{s,h} = \frac{4\pi}{k^2 P_{\text{tot}}^{s,h}} \left| e^{ikr_s} + \sum_{n=0}^{\infty} i^n y_n^{s,h} w_n^{s,h} \right|^2. \quad (32)$$

Note that the free-space radiated power and the free-space directivity of the same CSP feed are given, respectively, by the expressions:

$$P_0 = \frac{2\pi \sinh 2kb}{k^2 kb}, \quad D_0 = \frac{2kbe^{2kb}}{\sinh 2kb}. \quad (33)$$

Overall directivity  $D_{\text{tot}}$  should be compared with  $D_0$ .

#### IV. NUMERICAL RESULTS

In principle, the accuracy in solving Eq. (27) is limited only by the digital precision of the computer used, in contrast to the conventional MoM-type numerical approximations (e.g., Refs. 3, 4). For an accuracy in the far field of three digits, and in the near field of two digits, the number of equations to be taken is  $N_{\text{tr}} \geq ka + 20$  independently of the angular width  $2\theta_0$ , and of the feed parameters. This estimate is illustrated by the plots of normalized error of computed density function versus truncation number, presented in Fig. 2 (for the hard-surface case). The error is computed in the maximum-norm sense:

$$e(N) = \frac{\max_{n \leq N} |x_n^{N+1} - x_n^N|}{\max_{n \leq N} |x_n^N|}. \quad (34)$$

However, the error computed in the  $l^2$  sense shows very similar behavior. It should be recalled that the Fredholm nature of Eq. (26) guarantees that  $e(N) \rightarrow 0$  as  $N \rightarrow \infty$ . CPU

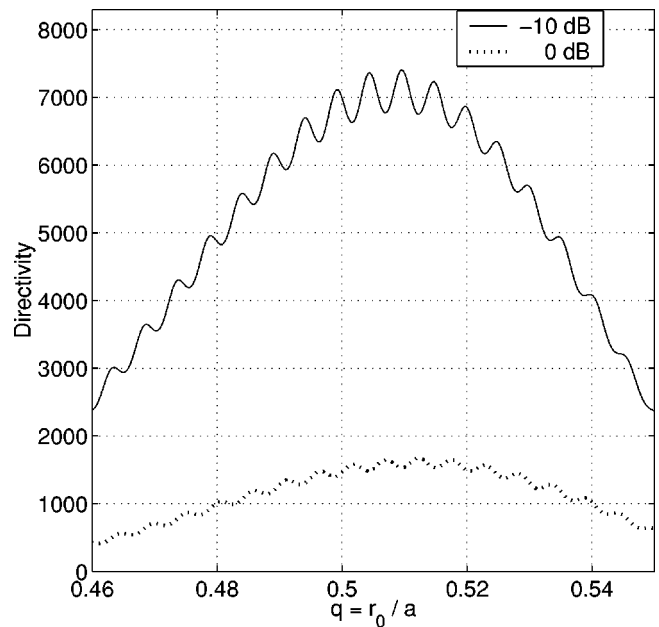


FIG. 3. Directivity of a hard-surface “shallow” reflector as a function of the normalized feed position. Solid curve is for the edge illumination  $-10$  dB, dashed curve is for nondirective source ( $kb=0$ ). Reflector parameters are:  $d=50\lambda$ ,  $f/d=0.97$  ( $ka=606.9$ ,  $\theta_0=15^\circ$ ).

time needed for solving a  $50\text{-}\lambda$  reflector of angular halfwidth  $\theta_0=15^\circ$ , that is  $ka=620$ , with a Pentium 133 computer and Fortran 77 source-code under Windows 3.11 was 2.5 min.

We present some further results on the dependence of overall directivity on the feed location in real space. We examine two  $d=50\lambda$  hard-surface reflectors: a “shallow dish” ( $\theta_0=15^\circ$ , or  $f/d=0.97$ ) and a “deep” one ( $\theta_0=30^\circ$ , or  $f/d=0.5$ ). The spherical reflector is believed to behave as a paraboloidal one, provided that the geometrical deviation between the two surfaces does not exceed  $\lambda/16$  (or even  $\lambda/8$ ).<sup>1</sup> Despite the paper titles, this is why spherical reflectors were considered in Refs. 12, 13. According to Fig. 2 of Ref. 11, this limits the aperture size of our “shallow dish” to the value  $d=53.5\lambda$ . Of course, even deeper or larger reflectors can still be considered (there is no computational difficulty), but spherical aberrations are known to degrade the main beam. This happens in the case of our “deep” reflector. However, such spherical reflectors remain of interest, due to easier manufacturing and mechanical beam steering.

Only an infinite paraboloid generates a plane wave, if the point feed is placed at the GO focus:  $r_0/a=0.5$ . In any finite-size geometry a focal shift occurs, predicted by GO,<sup>18</sup> and studied in Ref. 19 by using PO. However in reality, for a finite  $d$ , the dependence of  $D_{\text{tot}}$  on  $q=r_0/a$  has an even more complex and oscillatory nature. In Fig. 3 is presented the hard “shallow” reflector under two edge illuminations:  $-10$  dB (the same for all  $q$ , solid curve) and 0 dB (omnidirectional source, dashed curve). In the former case, the parameter  $kb$  was slightly varied around the value 8.2 to provide a constant illumination level. One may clearly see that there is not a single, but several positions of the feed providing almost equally good directivity. The broad maximum corresponds to the focal shift predicted by GO:<sup>19</sup> it is between  $r_0/a=0.5$  and  $r_0/a=\sec \theta_0/2$ , that is  $r_0/a=0.518$

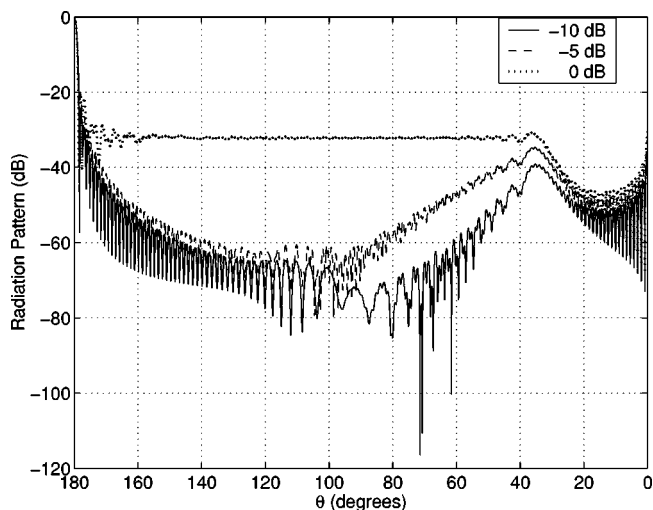


FIG. 4. Normalized far field radiation patterns of the same “shallow” reflector as in Fig. 3, for the feed position  $r_0/a=0.507$ . Edge illuminations are:  $-10$  dB (solid curve),  $-5$  dB (dashed curve), and non-directive source (dotted curve).

here. The period of the smaller oscillations is equal to  $\lambda/2$ . This feature of the near field is clearly not of GO nature, and does not appear to be predicted by high-frequency asymptotic approximations. In Fig. 4, we present total far field radiation patterns, computed for three different edge illuminations. Note that reduction of the edge illumination levels from  $-5$  dB to  $-10$  dB (that is, increasing the beam-width parameter  $kb$ , from 4 to 8.4) mainly affects the far sidelobes, between  $30^\circ$  and  $90^\circ$ . The feed position here corresponds to the optimum.

The next series of results illustrates features of the “deep” reflector. In this case, the spherical shape of the latter has a greater effect. The directivity dependence on the feed position (Fig. 5) shows a broad maximum in the middle

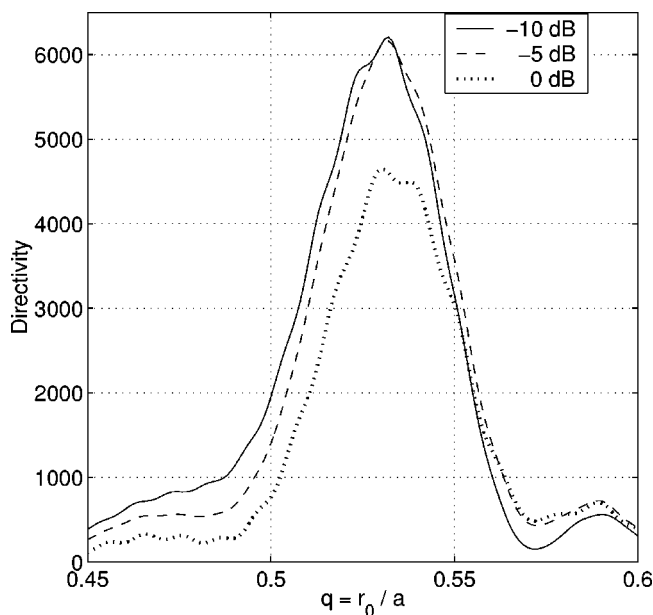


FIG. 5. Directivity of a hard-surface “deep” reflector as a function of the normalized feed position. Edge illuminations are:  $-10$  dB (solid curve),  $-5$  dB (dashed curve), and nondirective source (dotted curve). Reflector parameters are:  $d=50\lambda$ ,  $f/d=0.5$  ( $ka=314.2$ ,  $\theta_0=30^\circ$ ).

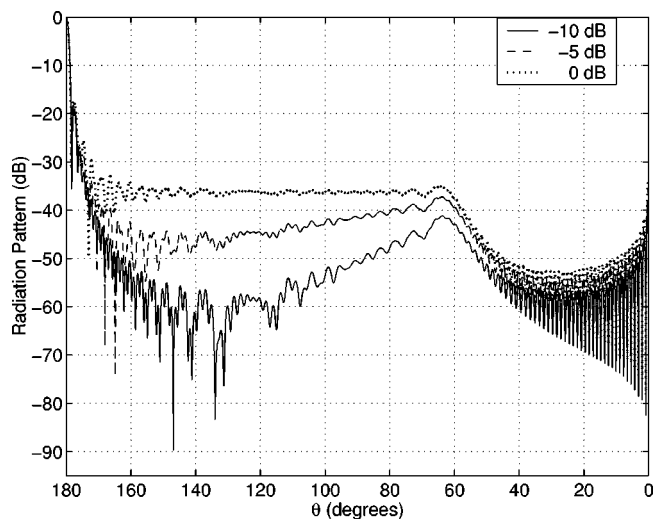


FIG. 6. Normalized far field radiation patterns of the same “deep” reflector as in Fig. 5, for the feed position  $r_0/a=0.526$ . Edge illuminations are:  $-10$  dB (solid curve),  $-5$  dB (dashed curve), and nondirective source (dotted curve).

of GO-predicted interval, between  $r_0/a=0.5$  and  $0.578$ , but no oscillations. Note that the maximum directivity for the omnidirectional feed is double that of the “shallow” reflector case (compare with Fig. 3), due to doubling the area of reflector. However, the directivity in this case is much more critically dependent on the feed position, the GO focus being completely unacceptable. Far field patterns presented in Fig. 6 are more sensitive to the decrease of edge illumination levels from  $-5$  dB to  $-10$  dB (by increasing the beam-width parameter  $kb$ , from 0.8 to 2.1), and show the effect of spherical aberrations, not visible in Fig. 4.

Most of these features are observed for the soft reflectors as well: see Figs. 7 and 8. Similarly for the hard reflector,

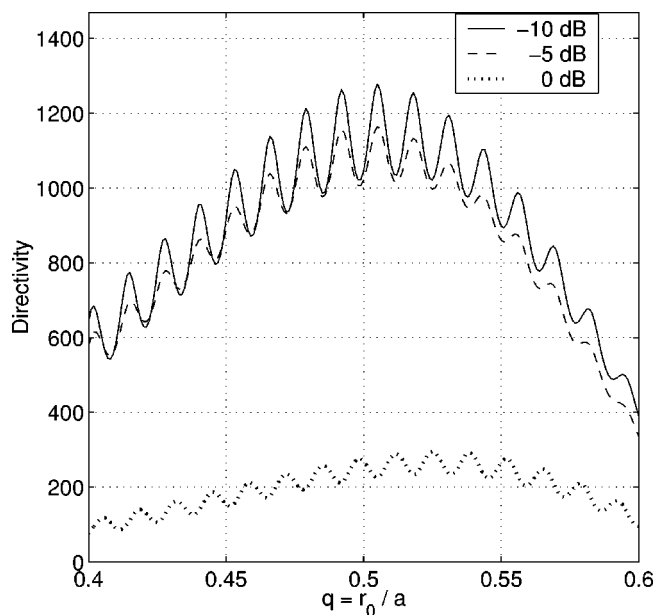


FIG. 7. Directivity of a soft-surface “shallow” reflector as a function of the normalized feed position. Edge illuminations are:  $-10$  dB (solid curve), and nondirective source (dotted curve). Reflector parameters are:  $d=20\lambda$ ,  $f/d=0.97$  ( $ka=242.8$ ,  $\theta_0=15^\circ$ ).

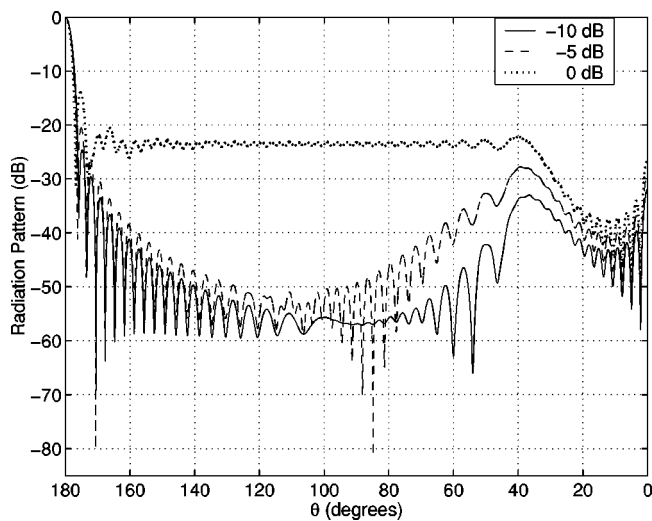


FIG. 8. Normalized far field radiation patterns of the same reflector as in Fig. 7, for the feed position  $r_0/a=0.504$ . Edge illuminations are:  $-10$  dB (solid curve),  $-5$  dB (dashed curve), and nondirective source (dotted curve).

decreasing the edge illumination below  $-5$  dB has a small effect on the overall directivity. As expected, a soft-surface edge produces smaller sidelobes than a hard-surface one.

## V. CONCLUSIONS

We have presented a simple but powerful algorithm for the analysis of spherical reflector antennas fed by a scalar beam. It is based on the exact analytical inversion of the singular part of the corresponding full-wave integral equation. This is achieved through the conversion of the problem to dual series equations, and followed by the Abel integral transform to invert the static part (in the soft case), or the “main” part (in the hard case). The resulting infinite matrix equations are of Fredholm second kind, thus ensuring existence of the exact solution, and the opportunity of obtaining it within machine precision, by taking more and more equations. This approach was previously used in plane-wave scattering analysis; here, it is combined with a complex source-point simulation of a directive feed field employing a function which is an exact solution to the Helmholtz equation. The analytical regularization approach described above can be considered numerically exact because the achievable accuracy is limited only by the computer used, and is uniform with respect to the frequency and the other parameters.

Accurate numerical analysis of the considered problem reveals several interesting features of the wave field that do not seem to be predicted by asymptotic techniques. For example, besides the well-known focal shift toward the reflector, the directivity (as a function of the feed position) displays not a single, but several, local maxima near the GO-predicted shifted focal point. Generally, optimization of antenna geometry depends on the cost functional. Maximum

directivity is a possible choice of cost functional, but minimum sidelobe level is another, or a combination of both is also reasonable. Potential applications of the analysis presented are in the area of hydro-acoustic antenna design.

One possible extension of this analysis is to parabolic or other nonspherical reflectors. Here, an IE, similar to Eq. (5) or (14), will require a modified domain of integration. However, analytical regularization of this IE can be based upon the extraction and inversion of the singular static part of the spherical dish IE operator. Hence, the technique presented above will be at the core of the modified analysis.

## ACKNOWLEDGMENTS

The first two authors are grateful to NATO and the Royal Society for a research fellowship which supported this work, and also to the Department of Mathematics, University of Dundee, for its support for this work.

- <sup>1</sup>J. D. Kraus, *Antennas* (McGraw-Hill, New York, 1988).
- <sup>2</sup>C. Scott, *Modern Methods of Reflector Antenna Analysis and Design* (Artech House, Boston, 1990).
- <sup>3</sup>D. C. Jenn, M. A. Morgan, and R. J. Pogorzelski, “Characteristics of approximate numerical modeling techniques applied to resonance-sized reflectors,” *Electromagnetics* **15**, 41–53 (1995).
- <sup>4</sup>M. R. Barclay and W. V. T. Rusch, “Moment-method analysis of large, axially symmetric reflector antennas using entire-domain functions,” *IEEE Trans. Antennas Propag.* **39**, 491–496 (1991).
- <sup>5</sup>G. A. Deschamps, “Gaussian beam as a bundle of complex rays,” *Electron. Lett.* **7**, 684–685 (1971).
- <sup>6</sup>L. B. Felsen, “Complex-point source solutions of the field equations and their relation to the propagation and scattering of Gaussian beams,” *Symp. Math. Instituta di Alta Matematica* **18**, 39–56 (1975).
- <sup>7</sup>P. A. Belanger and M. Couture, “Boundary diffraction of an inhomogeneous wave,” *J. Opt. Soc. Am.* **73**, 446–450 (1983).
- <sup>8</sup>G. A. Suedan and E. V. Jull, “Scalar beam diffraction by a wide circular aperture,” *J. Opt. Soc. Am. A* **5**, 1629–1634 (1988).
- <sup>9</sup>A. N. Norris, “Complex point-source representation of real point sources and the Gaussian beam summation method,” *J. Opt. Soc. Am. A* **3**, 2005–2010 (1986).
- <sup>10</sup>A. N. Norris and T. B. Hansen, “Exact complex source representations of time-harmonic radiation,” *Wave Motion* **25**, 127–141 (1997).
- <sup>11</sup>T. Oğuzer, A. Altıntaş, and A. I. Nosisch, “Accurate simulation of reflector antennas by the complex source—dual series approach,” *IEEE Trans. Antennas Propag.* **43**, 793–801 (1995).
- <sup>12</sup>Y. N. Feld and A. K. Ansryan, “Diffraction of scalar wave by a parabolic reflector,” *Radio Eng. Electron. Phys.* **22**, 1 (1977).
- <sup>13</sup>I. V. Vovk, “Diffraction of sound waves by a parabolic reflector,” *Sov. Phys. Acoust.* **25**, 287–289 (1979).
- <sup>14</sup>S. S. Vinogradov, “A soft spherical cap in the field of a plane sound wave,” *USSR Comput. Math. Math. Phys.* **18**, 244–249 (1978).
- <sup>15</sup>S. S. Vinogradov, Y. A. Tuchkin, and V. P. Shestopalov, “Summatory equations with kernels in the form of Jacobi polynomials,” *Sov. Phys. Dokl.* **25**, 531–532 (1980).
- <sup>16</sup>D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory* (Wiley-Interscience, New York, 1983).
- <sup>17</sup>M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
- <sup>18</sup>R. C. Spenser and G. Hyde, “Studies of the focal region of a spherical reflector: Geometrical optics,” *IEEE Trans. Antennas Propag.* **AP-16**, 317–324 (1968).
- <sup>19</sup>H. Ling, S.-W. Lee, P. T. C. Lam, and W. V. T. Rusch, “Focal shifts in parabolic reflectors,” *IEEE Trans. Antennas Propag.* **AP-33**, 744–748 (1985).



# Effects of interaction between two bubble scatterers

George Kapodistrias

*Department of Mechanical Engineering and Applied Physics Laboratory, University of Washington, Seattle, Washington 98195*

Peter H. Dahl

*Applied Physics Laboratory, University of Washington, 1013 NE 40th Street, Seattle, Washington 98105*

(Received 19 July 1999; revised 25 January 2000; accepted 6 March 2000)

The backscattering of sound from two regularly arranged bubbles is studied theoretically and experimentally. In well-controlled laboratory experiments a bistatic acoustic system is used to interrogate the scatterers, which are placed on a very fine thread at the same distance  $d$  from the combined beam axis of the set of transmitting and receiving transducers. The radius of each bubble is  $585\ \mu\text{m}$ . The frequency range is 80–140 kHz, and  $d$  is varied so that the variable  $kd$  spans the range 0.2–21, where  $k$  is the acoustic wave number. Scattering calculations are carried out using an exact, closed-form solution derived from the multiple scattering series. Several experiments are performed, and the results are in close agreement with the calculations. It is verified that multiple scattering induces an oscillatory behavior about the exact coherent scattering level, with decreasing amplitude for increasing  $kd$ . For interbubble distance  $2d \approx \lambda/2$  the backscattered radiation is maximized, while for  $2d < \lambda/2$  the radiation is reduced considerably. These and other effects are discussed. © 2000 Acoustical Society of America. [S0001-4966(00)02106-8]

PACS numbers: 43.20.Fn [ANN]

## INTRODUCTION

Understanding the physical processes that develop when sound traverses through and scatters from a dense assemblage of scatterers is of great importance in a number of applications, such as scattering from oceanic bubble clouds,<sup>1</sup> schools of fish,<sup>2</sup> plankton,<sup>3</sup> ultrasonic contrast agents,<sup>4</sup> and blood.<sup>5</sup> In such cases, multiple scattering complicates the inverse scattering problem of extracting information about the scatterers or the scattering medium from the acoustic field at the receiver. In this paper, we address a component of the forward problem pertaining to multiple scattering. The fundamental question we seek to answer is how the backscattered wave is affected by varying the center-to-center distance,  $2d$ , between two bubble scatterers symmetrically arranged about the combined beam axis. Subsequent discussions and analysis will refer to air bubbles in water, although the general theory is not restricted to such.

Consider an assemblage of bubbles with mean interbubble distance  $2d$ , insonified by a plane wave with wavelength  $\lambda$ . For a sparse assemblage of bubbles, say  $2d \gg \lambda$ , that are randomly spaced (as could be the case for some oceanic bubble clouds), the possibility of interaction between the individual bubbles is small, and multiple scattering effects can be neglected (e.g., see Ref. 6). Therefore, the Born approximation is valid, and it can be shown that the total (incoherent) scattered intensity per unit incident intensity at a point in the field assumes a linear dependence between the number of bubbles and their total scattering cross section. However, for regularly spaced (and sparse) assemblages the total intensity at the receiver is the result of coherent scattering, and is therefore equal to the square of the sum of the individual scattered pressures. For dense assemblages, say  $2d \leq \lambda$  (as could be the case for bubble clouds very close to the surface or those associated with surf zones), the theoret-

ical analysis is more complicated, because the interactions between the bubbles play a definite role in the development of the scattered field. These interactions include, among others, multiple scattering.

In multiple scattering the field scattered from each element depends both on the incident wave from the source and on the waves scattered by all other elements in the scattering volume. It is possible to write a series of scattering terms to obtain the acoustic field at the receiver. However, this series is infinite and also will become quite intractable for a large number of scatterers. One approach that simplifies the problem is to perform a partial summation on only the most important contributions to the series (e.g., see Refs. 7 and 8). Another simplification that incorporates multiple scattering effects is to treat the volume occupied by the bubbles as a continuous one with some equivalent density and compressibility. Such a solution, offered in 1945 by Foldy,<sup>9</sup> is the “effective medium” model. This approach is based on averaging the acoustic field scattered by a large number of omnidirectional scatterers governed by a spatial Poisson distribution. Since then other researchers have improved on Foldy’s effective medium model by adding small corrections.<sup>10–12</sup> The effective medium approximation has been successfully compared with experimental data from random assemblages for insonifying frequencies that do not correspond to the resonance frequency of the bubbles. For example, Commander and Prosperetti<sup>1</sup> used the model to analyze five different data sets, and the results showed good agreement between theory and experiment, even for volume fractions up to 10%, a rather remarkable result.

In contrast, a survey of the available literature showed that very little data exist from well-controlled laboratory experiments for which deterministic multiple scattering effects could be studied. One exception we are aware of is the work

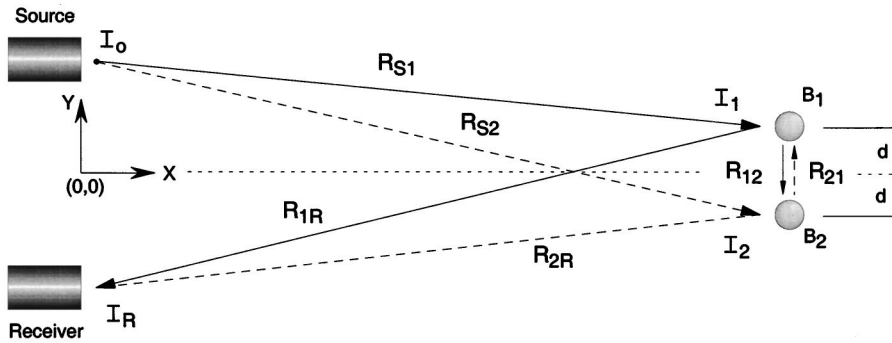


FIG. 1. Geometry of the scattering problem (not to scale).

by Bjørnø and Bjørnø,<sup>8</sup> who used two stainless steel spheres in the laboratory to study the effects of sphere separation and angle of incidence on the backscattering of sound and compared the results with a simple model.

In this paper we derive a solution to a specific scattering problem using the multiple scattering expansion approach and compare it with experimental results. The experimental data consist of measurements of the pressure field backscattered from two nearly identical bubbles symmetrically placed the same distance  $d$  from the combined beam axis of the acoustic transmit and receive beams. The radius  $a$  of each bubble was  $(585 \pm 35) \mu\text{m}$ , with  $ka$  ranging from 0.2 to 0.35, where  $k$  is the acoustic wave number for the surrounding medium. The primary independent variable was  $d$ , and  $kd$  spanned the range from 0.2 to 21. In Sec. I the multiple scattering series is established and a closed form solution is derived, and in Sec. II the frequency response of a single bubble in water is reviewed. The experimental platform and the procedure are described in detail in Sec. III, which includes the consideration of effects brought about by the geometry of the system. The experimental results are compared with the scattering model in Sec. IV, and found to be in good agreement, and a summary of the work is given in Sec. V.

## I. MULTIPLE SCATTERING THEORY

The simplest problem in multiple scattering is one involving just two identical, stationary scatterers symmetrically insonified by a plane wave at normal incidence. Intuition suggests that in such a case scattering will be completely coherent, since no randomness exists in the location of the scatterers or any other parameter. Various methods have been used to investigate this problem theoretically. For example, Twersky<sup>13</sup> obtained a closed form solution for two identical and isotropic scatterers insonified by single-frequency plane waves in terms of the scattering function of one of the scatterers. He went on to develop a set of equations that could be applied to problems involving arbitrary scatterers and angles of incidence. Twersky's work provides an intuitive description of the scattering process, one that can be extended to a variety of problems (for example, see the work by Tolstoy<sup>14,15</sup> on superresonant systems). Gaunard *et al.*<sup>16</sup> performed a partial summation of the multiple scattering series in terms of spherical harmonics and derived a pair of coupled equations that describe acoustic scattering by a pair of rigid and soft spheres for arbitrary angles of incidence. However, for the particular cases modeled the spheres

were in the far field of each other, thereby mitigating strong multiple scattering effects. In this paper the multiple scattering series expansion will be used since for simple problems it is tractable and allows proper accounting of scattering terms of all orders.

The problem under consideration is the scattering of sound by a pair of bubbles (Fig. 1). A bistatic system is utilized to insonify the pair and receive the scattered pressure field in the frequency range 80 to 140 kHz. The bubbles are of the same size, undeformed, and held fixed within the far field of the transducers. The speed of sound in the water is constant ( $c = 1490 \text{ m/s}$ ), and the losses due to attenuation are negligible.

The complex pressure of the acoustic wave incident on bubble  $B_1$  due to the source can be written (dropping the  $e^{-i\omega t}$  term) as

$$p_{S1} = P_0 \frac{e^{ikR_{S1}}}{R_{S1}}, \quad (1)$$

and the pressure incident on bubble  $B_2$  as

$$p_{S2} = P_0 \frac{e^{ikR_{S2}}}{R_{S2}}. \quad (2)$$

Assuming that the pressure field scattered from each bubble propagates in the medium according to spherical spreading, even in the near field of the bubble (see the Appendix for justification of the assumption), the pressure of the acoustic field at the receiver due to bubble  $B_1$  becomes

$$p_{S1R} = p_{S1} f_{B1}(\theta_{SB1R}, R_{1R}) \frac{e^{ikR_{1R}}}{R_{1R}}, \quad (3)$$

and that due to bubble  $B_2$  becomes

$$p_{S2R} = p_{S2} f_{B2}(\theta_{SB2R}, R_{2R}) \frac{e^{ikR_{2R}}}{R_{2R}}. \quad (4)$$

In Eqs. (1)–(4)  $R_{S1}$  and  $R_{S2}$  are the distances from the source to bubbles  $B_1$  and  $B_2$ , respectively,  $R_{1R}$  and  $R_{2R}$  are the distances from the two bubbles to the receiver, and  $f_{B1}$  and  $f_{B2}$  are the complex scattering functions of each bubble. The complex scattering function  $f_B(\theta, r)$  determines the amplitude and phase of the scattered wave at a distance  $r$  from the center of the bubble and angle  $\theta$  with respect to the incident plane wave. The scattering angles associated with the problem are shown in Fig. 2. The scattering function will be discussed in more detail in Sec. II.

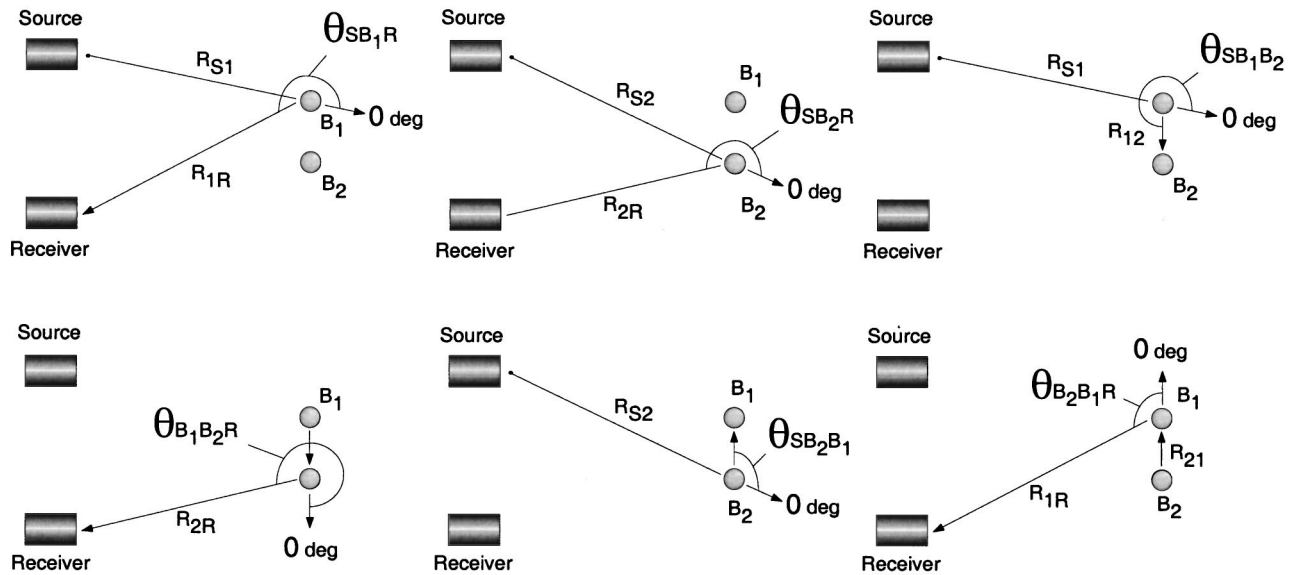


FIG. 2. Scattering angles for the two-bubble problem.

In the following we adopt a shorthand notation for the spherical spreading propagation terms, as  $\mathcal{P}_{\kappa\xi} = e^{ikR_{\kappa\xi}}/R_{\kappa\xi}$ , with  $R_{\kappa\xi}$  assuming the representation of the four distances noted above, namely,  $R_{S1}$ ,  $R_{S2}$ ,  $R_{1R}$  and  $R_{2R}$ . Looking ahead, the bubble-to-bubble distances  $R_{12}$  and  $R_{21}$  will also be used in subsequent calculations, with the corresponding propagation terms expressed as  $\mathcal{P}_{12}$  and  $\mathcal{P}_{21}$ . The distances between the various elements can be calculated with the aid of the position vectors constructed from the known geometry of the system. Substitution of Eqs. (1) and (2) into Eqs. (3) and (4) and summation of the latter will give the complex pressure field at the receiver due to single scattering,

$$P_{\text{SingleScatter}} = P_0 [\mathcal{P}_{S1} f_{B1}(\theta_{SB1R}, R_{1R}) \mathcal{P}_{1R} + \mathcal{P}_{S2} f_{B2}(\theta_{SB2R}, R_{2R}) \mathcal{P}_{2R}]. \quad (5)$$

Note that for two stationary bubbles symmetrically arranged side by side, the intensity is the coherent one, and thus for two identical bubbles (i.e., the case studied herein), the scattering strength due to single scattering is  $20 \log(2) = 6$  dB above the scattering strength of the single bubble.

In deriving the coherent intensity it is assumed that bubbles do not interact, an assumption that is violated when

multiple scattering effects are induced owing to the proximity of the bubbles. To account for the interaction effects between the two bubbles, the multiple scattering series (MSS) will be developed. This will be accomplished with the help of the simplest graphical procedure of the *Feynman diagrams*, a method that was also used by Ye and Ding<sup>11</sup> to apply a correction to Foldy's theory. Adapted from Ref. 17 (see pp. 12–24) the MSS can be graphed as shown in Fig. 3, where each arrow represents a propagation term (describing spherical spreading) and each circle represents scattering from a bubble. The two single scattering events described by Eqs. (3) and (4) are shown as the first two terms of the series. The third and fourth terms represent the one-way scattering events and are written as

$$P_{S12R} = P_{S1} f_{B1}(\theta_{SB1B2}, R_{12}) \mathcal{P}_{12} f_{B2}(\theta_{B1B2R}, R_{2R}) \mathcal{P}_{2R} \quad (6)$$

and

$$P_{S21R} = P_{S2} f_{B2}(\theta_{SB2B1}, R_{21}) \mathcal{P}_{21} f_{B1}(\theta_{B2B1R}, R_{1R}) \mathcal{P}_{1R}. \quad (7)$$

The fifth and sixth terms in the MSS represent the double scattering events, which are expressed as

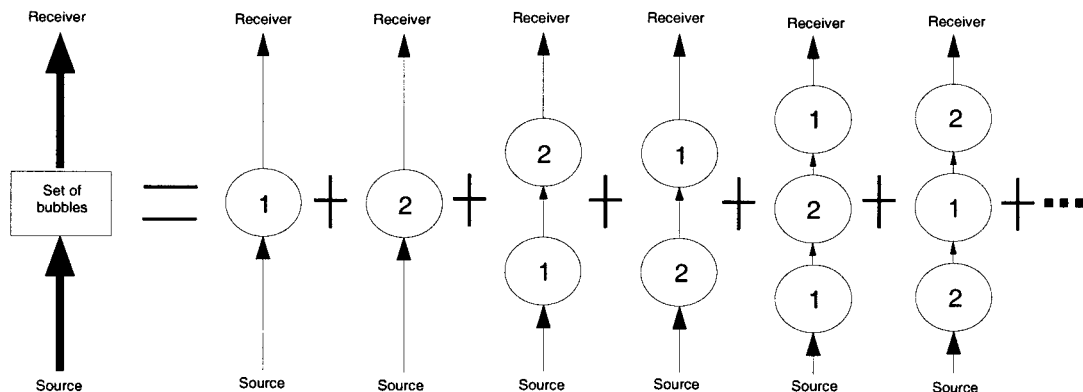


FIG. 3. Diagram of the multiple scattering series for the two bubbles.

$$P_{S121R} = P_{S1f_{B1}}(\theta_{SB_1B_2}, R_{12}) \mathcal{P}_{12f_{B2}}(180^\circ, R_{21}) \times \mathcal{P}_{21f_{B1}}(\theta_{B_2B_1R}, R_{1R}) \mathcal{P}_{1R} \quad (8)$$

and

$$P_{S212R} = P_{S2f_{B2}}(\theta_{SB_2B_1}, R_{21}) \mathcal{P}_{21f_{B1}}(180^\circ, R_{12}) \times \mathcal{P}_{12f_{B2}}(\theta_{B_1B_2R}, R_{2R}) \mathcal{P}_{2R}. \quad (9)$$

For brevity, we let

$$\mathcal{Z} = \mathcal{P}_{21f_{B1}}(180^\circ, R_{12}) \mathcal{P}_{12f_{B2}}(180^\circ, R_{21}). \quad (10)$$

Higher order terms in the MSS can be written simply by multiplying Eqs. (6)–(9) by an appropriate power of  $\mathcal{Z}$ . For example, the triple scattering terms can be written as

$$P_{S1212R} = P_{S12R} \cdot \mathcal{Z}^1 \quad \text{and} \quad P_{S2121R} = P_{S21R} \cdot \mathcal{Z}^1. \quad (11)$$

Summing all terms of the MSS results in the total pressure at the receiver,  $p_T$ , which can be written as

$$p_T = P_0 \cdot \sum_{\kappa=1}^2 \sum_{\xi=1}^2 \mathcal{P}_{S\kappa f_{B\kappa}}(\theta_{SB_\kappa R}, R_{\kappa R}) \mathcal{P}_{\kappa R} + [\mathcal{P}_{S\kappa f_{B\kappa}}(\theta_{SB_\kappa B_\xi}, R_{\kappa\xi}) \mathcal{P}_{\kappa\xi f_{B\xi}}(\theta_{\kappa\xi R}, R_{\xi R}) \mathcal{P}_{\xi R} + \mathcal{P}_{S\kappa f_{B\kappa}}(\theta_{SB_\kappa B_\xi}, R_{\kappa\xi}) \mathcal{P}_{\kappa\xi f_{B\xi}}(180^\circ, R_{\xi\kappa}) \times \mathcal{P}_{\xi\kappa f_{B\kappa}}(\theta_{B_\xi B_\kappa R}, R_{\kappa R}) \mathcal{P}_{\kappa R}] \times \frac{1}{1-\mathcal{Z}}, \quad (12)$$

where  $\kappa \neq \xi$ .

In Eq. (12) the term  $1/(1-\mathcal{Z})$  arises from the fact that the higher order terms that multiply the term in brackets form the series  $(1 + \mathcal{Z} + \mathcal{Z}^2 + \dots)$ , which is geometric. For a similar problem the generalized closed form solution obtained by Twersky<sup>13</sup> is  $2\mathcal{H}(r)\cos(kd\sin\theta)A[1 - A\mathcal{H}(2d)]$ , where  $A$  is the complex scattering function of a monopole,  $r$  is the distance from the bubbles to the receiver,  $\mathcal{H}(r)$  describes the propagation from the bubbles to the receiver,

$\mathcal{H}(2d)$  is the bubble to bubble propagation term and  $\theta$  is the relative bearing of the receiver from the pair of bubbles (see Sec. 1.2 in Ref. 13). Note that for our specific case of back-scattering  $\theta = 180^\circ$ , and  $\mathcal{H}(r) = e^{ikr}/r$ .

Our Eq. (12) is an exact solution for the any two-bubble problem, insofar as the bubble scattering function need not assume a purely monopole form. The bubbles can be of different size while the geometry of the problem can be arbitrary. Provided the scattering function of each bubble is known, and the propagation terms can be determined from the geometry of the system, the response of the pair of bubbles can be determined. Equation (12) will be used in Sec. IV to model the experiment. Importantly, our solution reduces to Twersky's concise expression for the case of scattering from two identical monopoles.

## II. THE RESPONSE OF A SINGLE GAS BUBBLE IN WATER

The general solution for scattering from a fluid sphere was derived by Anderson,<sup>18</sup> who found that the scattered pressure field equals

$$p_B = -p_{\text{inc}} \sum_{m=0}^{\infty} A_m P_m(\cos\theta) h_m^{(1)}(kr). \quad (13)$$

In Eq. (13)  $p_{\text{inc}}$  is the pressure field of the wave incident on the sphere,  $P_m$  is the Legendre polynomial of order  $m$ ,  $\theta$  is the scattering angle,  $h_m^{(1)}$  is the order  $m$  spherical Hankel function of the first kind,  $k$  is the wave number of the surrounding medium,  $r$  is the range from the sphere to a point in the field, and the amplitude coefficient  $A_m$  is given by

$$A_m = \frac{(-i)^m (2m+1)}{1 + iC_m}. \quad (14)$$

For a bubble of radius  $a$  filled with gas of density  $\rho_b$  and speed of sound  $c_b$  immersed in a fluid with density  $\rho_f$  and speed of sound  $c_f$  the coefficient  $C_m$  equals

$$C_m = \frac{\frac{m j_{m-1}(k_b a) - (m+1) j_{m+1}(k_b a)}{j_m(k_b a)} \frac{n_m(ka)}{j_m(ka)} - \frac{m n_{m-1}(ka) - (m+1) n_{m+1}(ka)}{j_m(ka)} \frac{\rho_b c_b}{\rho_f c_f}}{\frac{m j_{m-1}(k_b a) - (m+1) j_{m+1}(k_b a)}{j_m(k_b a)} \frac{j_m(ka)}{j_m(ka)} - \frac{\rho_b c_b}{\rho_f c_f}}, \quad (15)$$

where  $j_m$  is the spherical Bessel function,  $n_m$  is the spherical Neumann function and  $k_b$  is the wave number for the gas in the bubble.

We remark that for bubbles very small compared to the wavelength of the sound ( $ka \ll 1$ ), the solution can be simplified significantly. In this case, the complex scattering function can be written as

$$f_B|_{ka \ll 1} = \frac{a}{(f_{\text{res}}/f)^2 - 1 - i\delta}, \quad (16)$$

where  $f$  is the insonifying frequency,  $\delta$  is the total damping coefficient (see Ref. 19), and  $f_{\text{res}}$  is the resonant frequency of the bubble. Note that Eq. (16) is consistent with the  $e^{-i\omega t}$  dependence. When insonified at the resonance frequency  $f = f_{\text{res}}$ , the amplitude of the bubble wall oscillation is maximized, producing the well-defined peak in the frequency response curve [see Fig. 5(a)]. For an air bubble located at depth  $z$  meters in water the resonant frequency can be approximated as

$$f_{\text{res}} = \frac{3.25\sqrt{1+0.1z}}{a}, \quad (17)$$

where for  $a$  expressed in meters the resonant frequency is obtained in hertz. Note that Eq. (16) is independent of the scattering angle. Derivations and additional comments can be found in standard texts (e.g., Ref. 20). For the present case, the variable  $ka$  ranges from 0.2 to 0.35, which is not sufficiently small to allow the use of Eq. (16), and thus we use Eq. (13) to calculate the response of the bubble. The primary reason is that the insonified bubble oscillates back and forth in response to the forcing of the acoustic wave. The oscillation of the bubble will be out of phase with the oscillation of the surrounding medium, owing to the difference in density, giving rise to a weak dipole term. Consequently, an additional term, provided by Eq. (13), is needed to describe the pressure field scattered by the bubble. Finally, we remark that the general solution only accounts for radiation damping and neither viscous nor thermal damping. However, this does not introduce any errors in our calculations because for frequencies much higher than the resonant one, the total damping is not affected by the viscous and thermal components. Further comments on the differences between the general solution and the low  $ka$  approximation are presented in Sec. IV.

Utilizing Eq. (13), the complex scattering function of the bubble is thus given by

$$f_B(\theta, r) = -\frac{r}{e^{ikr}} \sum_{m=0}^1 A_m P_m(\cos \theta) h_m^{(1)}(kr), \quad (18)$$

and should not be confused with the far field form function,<sup>21</sup> which is a dimensionless representation of the scattered field and is independent of  $r$ . Equation (18) represents only two modes; the first ( $m=0$ ) describes the monopole radiation (breathing mode) of the bubble [thus is analogous to Eq. (16) with differences pertaining to aforementioned damping], while the second ( $m=1$ ) describes the dipole radiation. Higher order modes are suppressed in the calculations because they have negligible effect in the  $ka$  range of interest. For example, if the longitudinal quadrupole mode ( $m=2$ ) is included in Eq. (18), then, for a 585- $\mu\text{m}$  bubble insonified by a 140-kHz acoustic wave, the magnitude difference in the backscattering direction is

$$20 \log \left( \frac{|f_B|_{m=0,1,2}}{|f_B|_{m=0,1}} \right) = 0.014 \text{ dB}. \quad (19)$$

The Legendre polynomial of order zero,  $P_0(\cos \theta)$ , equals 1, while the polynomial  $P_1(\cos \theta)$  equals  $\cos \theta$  (see Ref. 22, p. 333), giving

$$f_B(\theta, r) = -\frac{r}{e^{ikr}} [A_0 h_0^{(1)}(kr) + A_1 \cos \theta h_1^{(1)}(kr)], \quad (20)$$

which will be combined with Eq. (12) to model the experiment.

Numerical results for backscattering from both single bubbles and the pair of bubbles will be reported in terms of target strength (TS). For single bubbles, the square of the

magnitude of  $f_B$  evaluated in the backscattered direction ( $\theta = 180^\circ$ ) gives the backscattering cross section,  $\sigma_{bs}$ , with TS given by

$$\text{TS} = 10 \log \sigma_{bs}, \quad (21)$$

in dB *re* 1 m<sup>2</sup>. For the pair of bubbles, the total scattered pressure at the receiver  $p_T$ , given by Eq. (12), is related to TS using

$$\text{TS} = 10 \log \left( \frac{|p_T|^2}{|p_{\text{inc}}|^2} \cdot R_{BR}^2 \right), \quad (22)$$

where  $p_{\text{inc}}$  is the incident pressure at the bubbles and  $R_{BR}$  is the distance from the bubbles to the receiver. For an exact calculation of the TS, the pressure  $p_{\text{inc}}$  should be substituted by  $p_{S1}$  and  $p_{S2}$  [Eqs. (1) and (2)], while the  $R_{BR}$  should be replaced by the distances  $R_{1R}$  and  $R_{2R}$ .

### III. EXPERIMENTAL APPARATUS AND PROCEDURE

The experiments were performed in a plexiglass tank of dimensions  $1.2 \times 0.45 \times 0.51 \text{ m}^3$  (L $\times$ W $\times$ H) filled with fresh water to a height of 45 cm. Its bottom, front, and back walls were lined with anechoic material to reduce reflections, and all the surfaces of the tank were thoroughly wetted. The two immersion transducers were arranged in a bistatic configuration. The transducers were Panametrics model V1011, with a peak frequency of 104 kHz and  $-6$ -dB points at 70 and 140 kHz. All the experiments were performed at a frequency range of 80 to 140 kHz, henceforth referred to as ‘‘operational frequency range.’’ The transducers were inserted in closed-cell Neoprene tubes to reduce the effect of reflections from the sides of the tank. The source was driven in a pulse mode using a six-cycle tone burst at a very low duty cycle, enabling resolution of the bubble signal. A computer equipped with a data acquisition card (1 MHz sampling rate) and data acquisition software acquired the amplified signal. A schematic of the apparatus is shown in Fig. 4. For each measurement, 20 waveforms were averaged and band-pass filtered, and the root-mean-squared voltage of the signal was calculated and squared.

The system was calibrated using an 18.2-mm radius tungsten–carbide sphere as a reference target. The theoretical analysis of the response of the tungsten–carbide sphere to continuous waves was based on the work by Faran<sup>23</sup> corrected according to a comment by Hickling.<sup>24</sup> The inherent bandwidth associated with a finite pulse was accounted for using a procedure described by Foote.<sup>25</sup>

A fine waxed nylon thread (approximately 150  $\mu\text{m}$  in diameter) held the bubbles in place by virtue of the bubbles adhering to it. Leighton *et al.*<sup>26</sup> studied the resonant-frequency properties of millimeter-sized bubbles also by using a wire to which free floating bubbles adhered. In their case, the wire alone produced a measurable response, which was subtracted from the combined bubble-and-wire response, and good agreement was achieved between optical and acoustic sizing. In our work, the thread alone did not return a signal above the noise floor, which was approximately 20 dB below the level of the signal from the single bubble target.

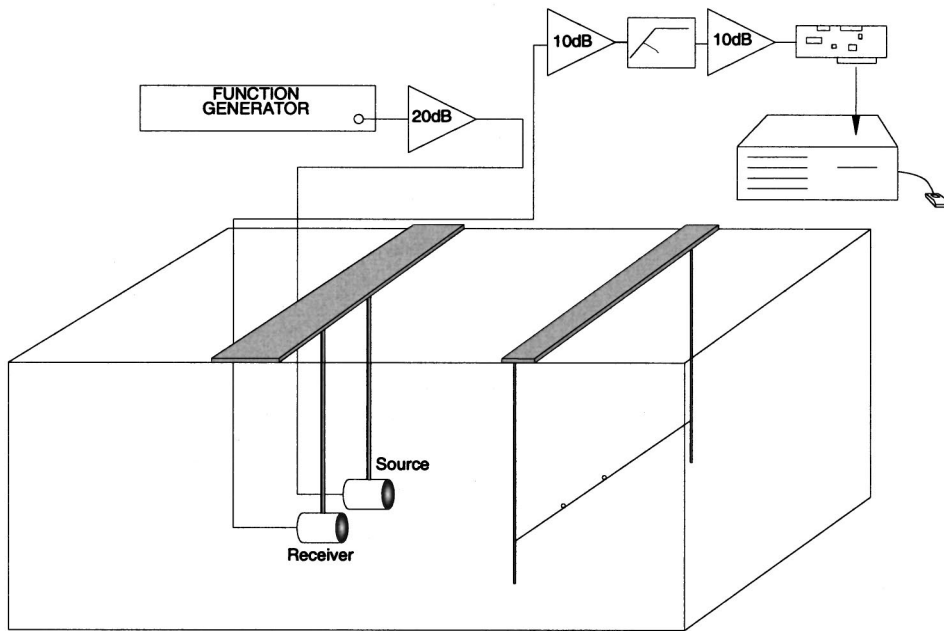


FIG. 4. Experimental apparatus.

The thread was attached to a thin wire frame and positioned on the axis of the combined beam, at a depth of 18 cm. The distance from the transducers to the thread was  $R_0 = 0.58$  m, which puts the targets well into the far field. The Rayleigh distance was 0.5 m at 140 kHz (see later in this article for explanation). The absorption coefficient for fresh water at 100 kHz is about 0.005 dB/m (Ref. 27) at the tank temperature of 20 °C, which for the roundtrip distance of 1.2 m makes attenuation negligible.

Bubbles were created using a 26 gauge hypodermic needle. Since the rate of release of the bubbles was pressure controlled, there was some variation in the size of the bubbles created, an effect that has been investigated by others.<sup>28,29</sup> However, with practice we could produce same-sized bubbles with good consistency. The size of the bubbles was measured with the acoustic system and found to be  $(585 \pm 35)$   $\mu\text{m}$  (see Fig. 5). The size was determined by comparing the TS of a single bubble measured across the

operational frequency range with the equivalent best-fit frequency-response curve obtained by Eq. (18), expressed in TS according to Eq. (21). Since the response is flat in the operational frequency range, the bandwidth of the six-cycle tone burst has no effect in the calculated response which pertains to CW excitation. The size was also checked independently using optical means. In that case a piece of thread was glued on a Petri dish and placed under a microscope, and the same procedure was used to produce and place bubbles on the thread as the one used in the actual experiments. The optically determined size was  $(575 \pm 40)$   $\mu\text{m}$ . This includes the required correction in the size due to the hydrostatic pressure at a depth of 18 cm. The optical system also verified that there was no significant change in the shape of the bubbles (see Fig. 6).

After the bubbles were created, they rose toward the surface and adhered to a thin metal rod. The bubbles, with the aid of the rod, were in turn placed at a predetermined

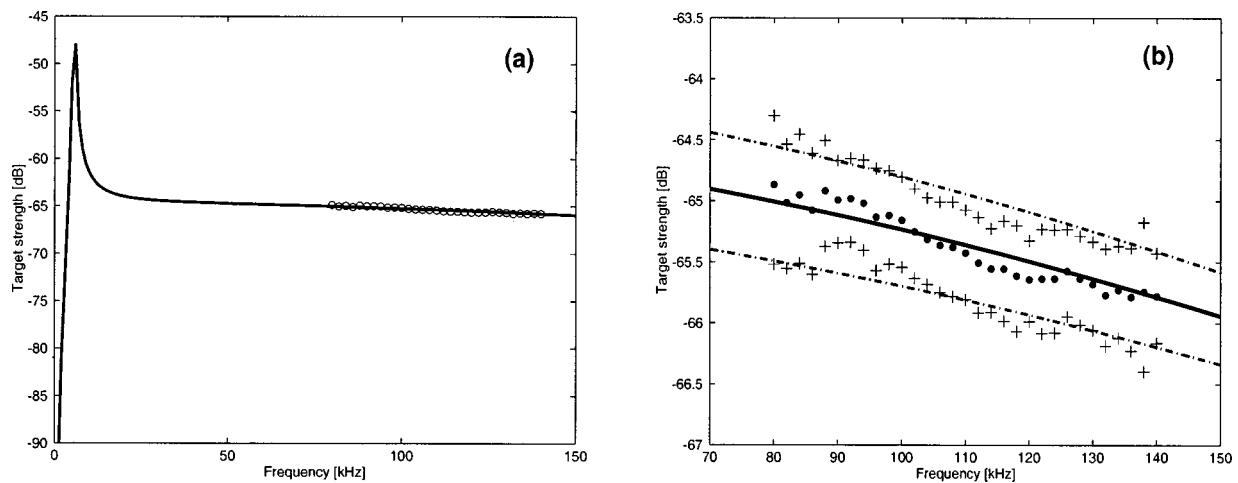


FIG. 5. Frequency response of a single bubble of radius 585  $\mu\text{m}$ , theory versus data. Graph (b) is a detail of graph (a). The middle line is the theoretical response of the 585- $\mu\text{m}$  bubble, while the upper line is for a 620- $\mu\text{m}$  bubble and the lower line for a 550- $\mu\text{m}$  bubble. The dots and crosses are the means and standard deviations of 18 experiments.

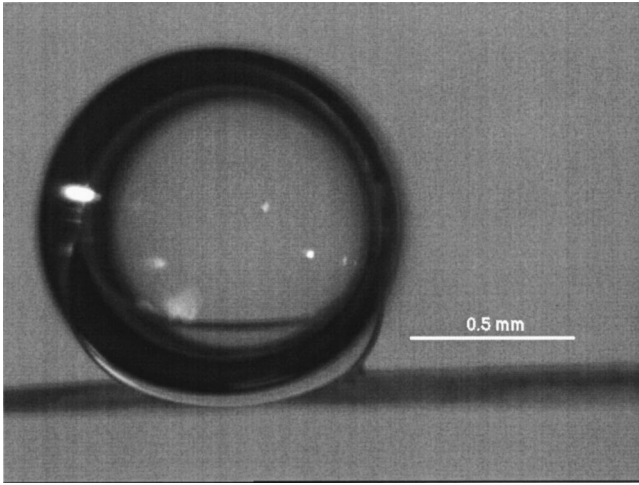


FIG. 6. Single bubble on the fine thread. The radius of bubble is  $(585 \pm 35) \mu\text{m}$ . Note that there is no significant distortion of the shape.

spot on the thread a distance  $d$  from the combined beam axis of the set of transducers.

As soon as the set of bubbles was placed in position, the TS was measured from 80 to 140 kHz in increments of 2 kHz. When the experiment was completed, a new set of bubbles replaced the old one. Each experiment required  $\approx 10$  min. Reduction of the TS due to dissolution of the gas during that time was minimal (about 0.1 dB).

Bubble separation was parametrized using the variable  $kd$ , where  $d$  is the distance from the center of each bubble to the axis of the combined beam. Because of the limited frequency range, it was necessary to perform the experiments for several distances  $d$  in order to span a larger  $kd$  range. The  $-3$ -dB two-way half-angle,  $\phi_{3\text{dB}}/2$ , at 140 kHz was  $3.8^\circ$ , which at  $R_0 = 0.58$  m reduced the maximum usable  $d$  to 38 mm. Note that this is well within the radius of the first Fresnel zone, equal to about 55 mm at 140 kHz. Thus, experimental data were acquired for  $d$  equal to 0.6, 1.5, 2, 2.5, 3, 5, 10, 15, 20, 25, 30, and 35 mm, and thus the variable  $kd$  ranged from 0.2 to 21. The distance  $d$  was measured with an uncertainty of  $\pm 0.5$  mm. At least four experiments were performed at each position to average out the experimental errors associated with the variable size of the bubbles and their positioning on the thread. The mean of the measurements is used in Sec. IV for comparison with the model.

Finally, two effects brought about by the geometry of the system need to be considered. These effects are the directivity pattern of the set of transducers and the directivity pattern of the pair of bubbles. The directivity pattern of the transducers reduces the signal level of targets that are off the axis of the beam. Therefore, it was necessary to correct the data accordingly (since one of the assumptions in the theory was that the targets are on the axis of the beam). It is convenient to obtain the two-way beam pattern of the set of transducers and model it as the beam of a single transducer of some equivalent diameter. Accordingly, transverse profiles of the beam at  $R_0 = 0.58$  m, at various frequencies, were obtained using a standard target (the tungsten-carbide sphere), and a Gaussian beam profile was fitted to the data.

Over the operational frequency range, the measured  $-3$ -dB beamwidth of the two-way beam,  $\phi_{3\text{dB}}$  expressed in degrees, was used to infer an equivalent aperture diameter using  $\phi_{3\text{dB}} \approx 59(\lambda/D)$ , giving  $D \approx 83$  mm. Thus the Rayleigh distance of the transducer pair, defined as  $R_d = \pi D^2/(4\lambda)$ , is  $R_d = 0.5$  m for our highest frequency of 140 kHz. With the beam pattern (out to the  $-3$ -dB level) completely characterized, the correction of the data became feasible.

In addition, since each insonified bubble intercepts and scatters sound, an assemblage of them will behave like an array of sources, and two bubbles in line will behave like a two-element line array. The directivity pattern of the far-field radiation pressure of a two-element broadside line array is<sup>30</sup>

$$H(d, f, \theta) = |e^{-ikd \sin \theta} (1 + e^{ik(2d) \sin \theta})|, \quad (23)$$

where  $d$  is the distance from the center of a bubble to the axis of the combined beam,  $f$  is the insonifying frequency, and  $\theta$  is the scattering angle. When the two bubbles are side by side the pattern is almost omnidirectional; as  $2d$  increases, the main lobe of the array becomes narrower and multiple side lobes appear. This effect was verified experimentally. A hydrophone mounted on a translation stage was used to obtain transverse beam profiles 30 cm from the bubbles. Figure 7 compares the theory and the experimental data.

#### IV. RESULTS AND INTERPRETATION

Considering that the monostatic configuration is a common geometry for many practical applications, we decided to colocate the source and receiver in the model, thus placing both transducers at position  $(x=0, y=0)$  (see Fig. 1), and correct the experimental data utilizing the predetermined directivity patterns of the transducers and the bubbles. The constants used in all of the simulations were (a) speed of sound  $c = 1490$  m/s; (b) mean bubble radius  $a = 585 \mu\text{m}$ ; (c) distance from transducers to set of bubbles  $R_0 = 0.58$  m; and (d) depth of bubbles  $z = 0.18$  m. Based on the known geometry of the system, the propagation terms were established and, with the incorporation of Eqs. (18) and (22), the TS at the receiver was calculated. Figure 8 shows the theoretical curve for backscattering at 110 kHz versus  $kd$ , with straight lines representing the coherent and incoherent scattering from two identical bubbles. The oscillatory behavior is brought about by the interference of the waves scattered between the bubbles. For the specific case shown in Fig. 8 the first maximum occurs at  $kd \approx 1.6$ , with successive maxima occurring every  $T_{kd} \approx 3.2$ . This implies that constructive interference results for

$$2d|_{\max, \alpha} \approx (2\alpha + 1) \frac{\lambda}{2} \quad (\alpha = 0, 1, 2, 3, \dots), \quad (24)$$

while destructive interference results for

$$2d|_{\min, \alpha} \approx \alpha \lambda. \quad (25)$$

Note that the amplitude of the oscillation diminishes with increasing  $kd$ ; for  $kd \rightarrow \infty$  the curve asymptotes to the exact coherent backscattering level. [Simulations showed that Eqs.

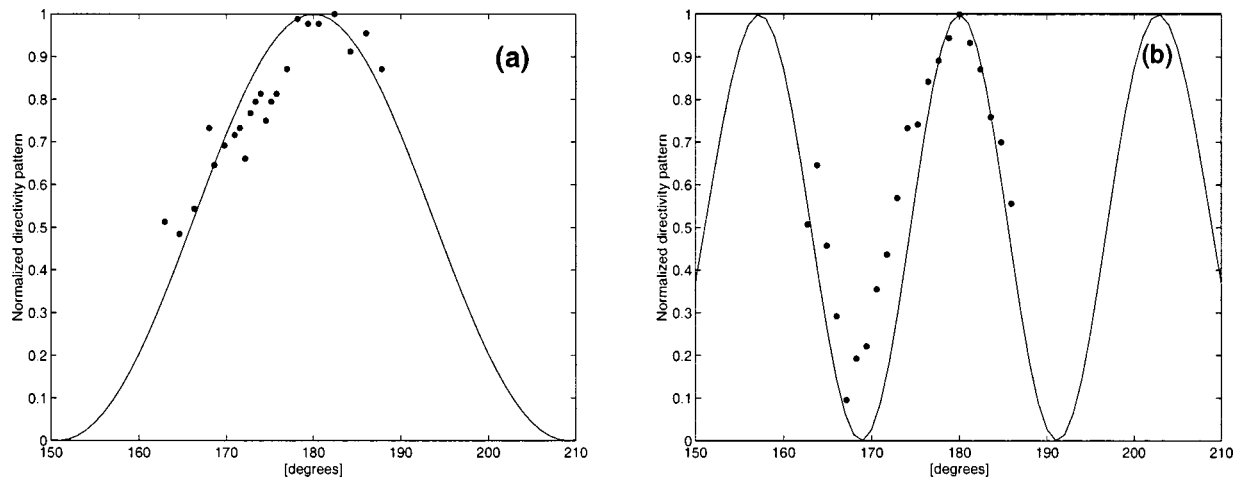


FIG. 7. Theoretical directivity pattern for a two-element line array (solid curve) versus experimental data (dots). Graph (a) shows the pattern for  $2d = 11$  mm and  $f = 140$  kHz. Graph (b) is for  $2d = 31$  mm and  $f = 125$  kHz.

(24) and (25) hold true across the operational frequency range.]

In Fig. 9 several theoretical curves (solid lines) for backscattering at fixed frequencies and variable  $d$  are compared with experimental data (dots). The black line represents backscattering from two  $585\text{-}\mu\text{m}$  bubbles, this radius being our best estimate of the mean size of each bubble. The two gray lines represent calculations pertaining to two bubbles each of radius  $620\ \mu\text{m}$  (upper line) and  $550\ \mu\text{m}$  (lower line). As mentioned in Sec. III,  $35\ \mu\text{m}$  is our best estimate of the standard deviation of the expected bubble radius; thus, these curves represent bounds to the experimental results.

The scattering calculations were carried out for fixed frequency (as shown in Fig. 9) and variable  $d$ . The calculated pressure was then expressed in terms of target strength and plotted against the parameter  $kd$ . The data are the ensemble average of the multiple experiments at each position. For clarity, the standard deviation of each measurement is not shown. It varied between  $\pm 0.2$  to  $\pm 1$  dB, depending on the proximity of the bubble signal to the noise floor (ex-

plained below). To compare the data with the simulations, the corrections mentioned in Sec. III were applied to the data. More specifically, the fitted Gaussian beams (over the operational frequency range) were used to calculate the loss (in dB) due to the bubbles being off axis of the combined beam of the transducers, while Eq. (23) was used to determine the loss (in dB) of the pressure at the receiver due to the pattern of the two-bubble array. The losses calculated were added to the measured TS to approximate the monostatic configuration for which the bubbles are on the axis of the combined beam.

In Fig. 9 the theoretical curves model the response to CW excitation. However, data was acquired using a six-cycle tone burst, which reduces the effects of multiple scattering as the distance between the two bubbles increases. To see how the tone burst influences the maximum  $kd$  that includes multiple scattering effects, we let  $t_{\text{on}}$  be the on time of the tone burst and  $t_T$  be the time it takes sound to travel from one bubble to the neighbouring one. Then, we can write

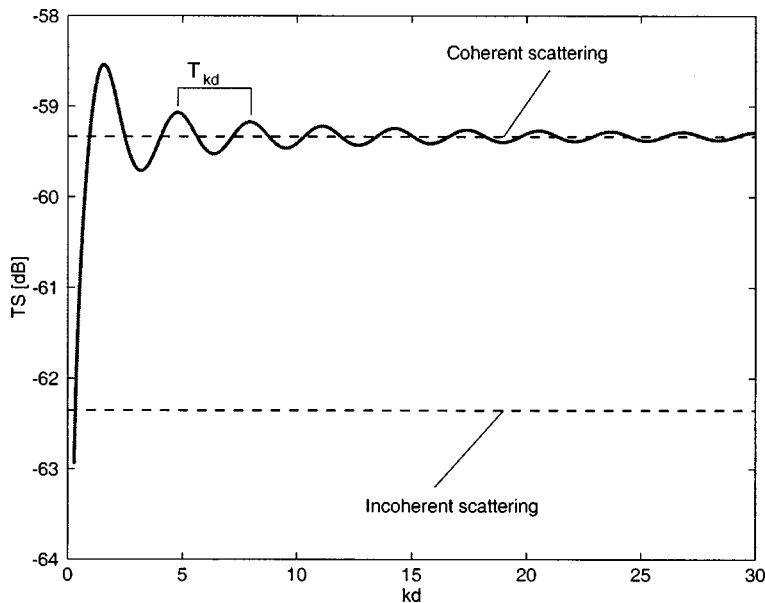


FIG. 8. Theoretical curve for backscattering from two  $585\text{-}\mu\text{m}$  bubbles at a frequency of  $110$  kHz and variable  $d$ , which is the distance from the combined beam axis to the center of each bubble.



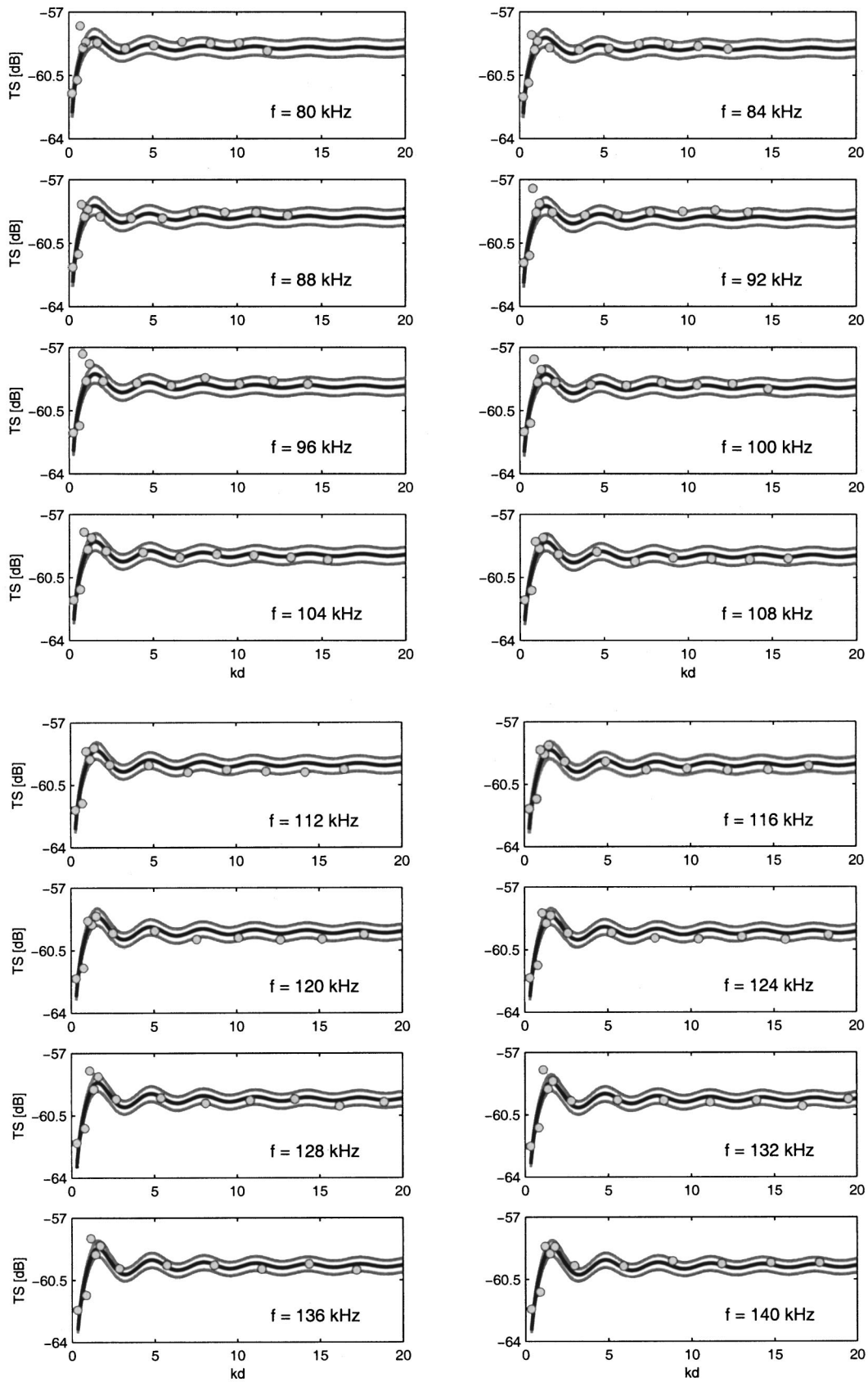


FIG. 9. Theoretical curves for backscattering from two bubbles of radius  $(585 \pm 35) \mu\text{m}$  for fixed frequency (as shown) and variable  $d$  versus experimental data (dots). The black line represents backscattering from two 585- $\mu\text{m}$  bubbles. The data are bounded by the responses of two 620- $\mu\text{m}$  bubbles (upper gray line) and two 550- $\mu\text{m}$  bubbles (lower gray line).

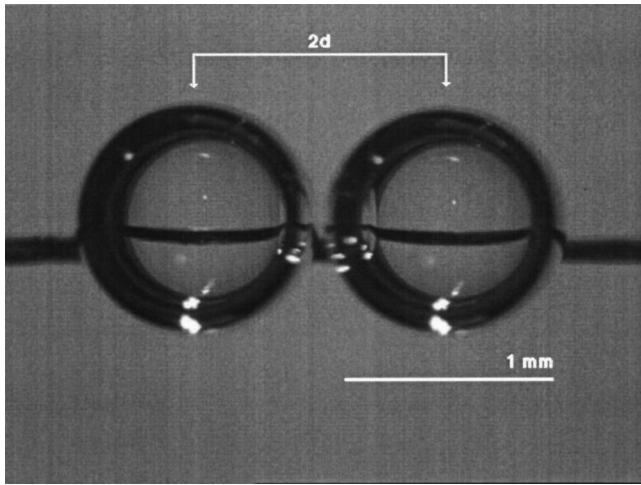


FIG. 10. Photograph of two bubbles separated by center-to-center distance  $2d \approx 1.2$  mm, equivalent to  $kd \approx 0.2$  at 80 kHz.

$$\frac{t_{\text{on}}}{t_T} = \frac{\pi \cdot \text{cycles}}{kd}, \quad (26)$$

where the variable “cycles” represents the number of cycles in the tone burst. Considering that the length of the pulse should be at least two times  $2d$  to observe multiple scattering effects, we deduce that  $kd < 9$ , for a six-cycle tone burst. For higher  $kd$  the absence of significant interaction between the bubble will result in coherent scattering, equal to the TS of a single bubble plus  $20 \log(2)$ .

The experimental data agree reasonably well with the response predicted by the model. For  $kd > 9$  the backscatter can be modeled as purely coherent, since the interaction effects are negligible and the scatter resembles single scattering. The increased scatter of the experimental data for  $kd > 7$  is due to beam width and directivity of the two-bubble array. As the separation between the bubbles increased, they fell off the axis of the combined beam of the transducers and, at the same time, the main lobe of the two-bubble array became narrower. Therefore, the signal received became weaker with increasing bubble separation, and the TS approached the noise floor (which was equivalent to a scatterer with TS equal to  $-80$ -dB, while located on-axis of the beam).

For  $kd < 1.6$  the backscattered radiation level is reduced considerably, presumably due to the effect of the mutual radiation impedance. The pressure radiated from one bubble applies a force to the neighboring one, altering in the process its radiation impedance, which becomes a sum of the mutual radiation impedances. As  $kd \rightarrow 0$  reactance dominates, and the velocity of the fluid particles are out of phase with the pressure and therefore reradiation is less efficient. Note that this is very similar to the problem of interactions between closely packed elements of arrays.<sup>31,32</sup> Interestingly, when  $kd$  falls below 0.3 (as depicted in Fig. 10), the TS approaches that of a  $740$ - $\mu\text{m}$  bubble, i.e., a bubble with a volume equal to twice the volume of a  $585$ - $\mu\text{m}$  bubble. At  $kd \approx 1.6$  the individual waves scattered from the bubbles add in phase, while the bubbles are not sufficiently far from each other that the dipole term is unimportant. The combination of

the increased magnitude and the constructive addition of the phases of the waves results in excess radiation. Note that the transition from reduced scatter to excess scatter to coherent scatter occurs very rapidly.

Finally, it should be mentioned that the incorporation of the complex scattering function for  $ka \ll 1$  in the model (or in Twersky’s concise expression mentioned in Sec. I) produced similar results with less computational strain. However, there were two differences that necessitated the use of the general solution. First, the magnitude of the frequency response for a single bubble calculated with the low  $ka$  approximation results in increased backscatter levels, since  $ka$  is nominally 0.3 for our case. For example, at 140 kHz the TS calculated using Eq. (16) is about 0.65 dB higher than the TS calculated using Eq. (18), an error that carries over in the outcome of Eq. (12). Second, the amplitude of the oscillation (e.g., that seen in Fig. 8), when calculated using the multimodal scattering function in Eq. (12), is slightly higher than this amplitude when calculated using the simplified scattering function. The reason is that the amplitudes of the scattered waves, which depend on scattering angle and distance from the bubble, are correctly accounted for by the combination of Eqs. (18) and (12). Admittedly, these differences are subtle for the geometry studied here (i.e., monostatic configuration and symmetrically arranged bubbles normal to the incident wave). However, we emphasize that our modeling approach is a general one that can be used with almost no modifications for bubbles with high  $ka$ , arbitrary sizes, and scattering geometries. This will be exploited in future work.

## V. SUMMARY

The problem of backscattering of sound from two bubbles was studied theoretically and experimentally in the frequency range 80–140 kHz. An exact, closed-form solution, based on the expansion of the multiple scattering series, was derived and used to model the backscattering response of the two-bubble system. Well-controlled experiments were performed by insonifying two identical bubbles attached to a fine thread, at an angle of  $90^\circ$  with respect to the two-bubble axis. The radius  $a$  of each bubble was  $585 \mu\text{m}$ , which translates to a  $ka$  range of 0.2–0.35, where  $k$  is the wave number of the medium. By varying the separation distance  $d$  between each bubble and the axis of the two-way beam, we spanned the range of  $kd$  from 0.2 to 21. The thread did not register a signal above the noise floor; furthermore, measurement of the size of single bubbles showed that it did not interfere with the response of the bubbles. Comparison of the numerical model and the experimental data showed good agreement. The constructive and destructive interference of the multiply scattered waves induces an oscillation about the exact coherent scattering level, with diminishing amplitude for increasing separation. At  $kd \approx 1.6$  ( $2d \approx \lambda/2$ ) the backscattered radiation was maximized, while for  $kd < 1.6$  ( $2d < \lambda/2$ ) interference effects resulted in reduced pressure levels.

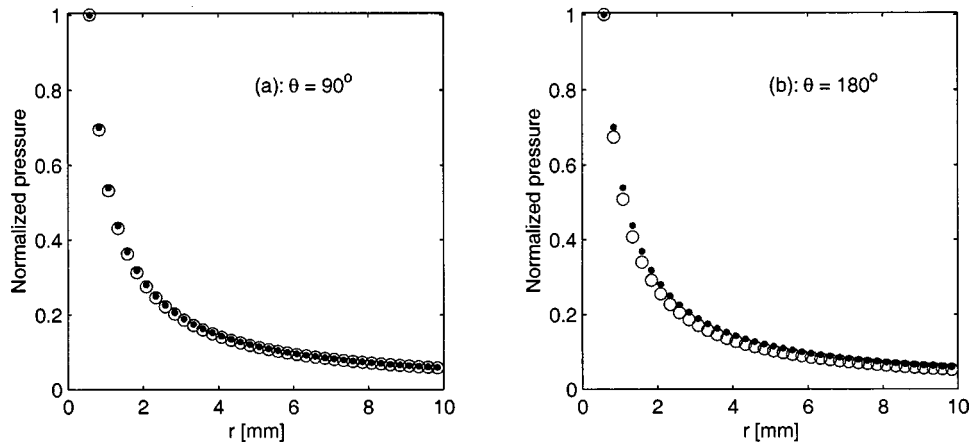


FIG. A1. Comparison between Eq. (A5) (circles) and the spherical spreading approximation  $1/r$  (dots) at  $ha=0.35$ . Each are normalized by their respective values at  $r=585 \mu\text{m}$ .

## ACKNOWLEDGMENTS

The authors wish to thank Dr. Steven Kargl of the Applied Physics Laboratory for the loan of certain instrumentation and for providing valuable feedback. We also appreciate the insightful comments from an anonymous reviewer. This work was funded by the Office of Naval Research, Code 321 OA.

## APPENDIX: SPHERICAL SPREADING IN THE NEAR FIELD OF A BUBBLE

In Sec. II the pressure field scattered by a bubble was written as

$$p_B = -p_{\text{inc}} \sum_{m=0}^{\infty} A_m P_m(\cos \theta) h_m^{(1)}(kr). \quad (\text{A1})$$

It was established that in the  $ka$  range 0.2–0.35 only the modes zero and one are significant, therefore, Eq. (A1) (normalized) can be written as

$$\frac{p_B}{p_{\text{inc}}} = -[A_0 h_0^{(1)}(kr) + A_1 \cos \theta h_1^{(1)}(kr)]. \quad (\text{A2})$$

The spherical Hankel functions of order zero and one can be expressed in terms of the free-field Green's function as

$$h_0^{(1)}(kr) = -\frac{i}{k} \frac{e^{ikr}}{r} \quad (\text{A3})$$

and

$$h_1^{(1)}(kr) = -\frac{i}{k} \frac{e^{ikr}}{r} \left( \frac{1}{i} + \frac{1}{kr} \right). \quad (\text{A4})$$

By substitution of Eqs. (A3) and (A4) in Eq. (A2), the normalized scattered pressure field becomes

$$\frac{p_B}{p_{\text{inc}}} = \frac{i}{k} \frac{e^{ikr}}{r} \left[ A_0 + A_1 \cos \theta \left( \frac{1}{i} + \frac{1}{kr} \right) \right]. \quad (\text{A5})$$

Study of Eq. (A5) reveals that in the far field the monopole term is clearly dominant. For example, for backscattering at 140 kHz

$$20 \log \left( \frac{|A_0 h_0^{(1)}(kr)|}{|A_1 \cos \theta h_1^{(1)}(kr)|} \right) \approx 20 \log \left( \frac{|A_0|}{|-A_1(-i)|} \right) = 18.95 \text{ dB}. \quad (\text{A6})$$

In the near field the weak dipole term becomes important, but the breathing mode of the bubble remains responsible for most of the energy reradiation. Therefore, propagation of the scattered pressure approximates spherical spreading. In Fig. A1 we graph Eq. (A5) (circles) against the normalized spherical spreading approximation  $1/r$  (dots) for two cases: (a)  $\theta=90^\circ$  and (b)  $\theta=180^\circ$ .

In the first case, the dipole term has no effect, since  $\cos 90^\circ = 0$ , therefore the agreement between the actual pressure reduction and spherical spreading is practically perfect. In the second case, the dipole term has maximum effect, since  $\cos 180^\circ = -1$ , which is evident in the difference between the exact and approximate solution. However, the difference is small enough to validate the approximation of spherical spreading in the near field of the bubble.

<sup>1</sup>K. W. Commander and A. Prosperetti, "Linear pressure waves in bubbly liquids: Comparison between theory and experiments," *J. Acoust. Soc. Am.* **85**, 732–746 (1989).

<sup>2</sup>T. K. Stanton, "Multiple scattering with applications to fish-echo processing," *J. Acoust. Soc. Am.* **73**, 1164–1169 (1983).

<sup>3</sup>N. Gorska and Z. Klusek, "Dependence of scattered acoustical signal intensity on the form of distribution of plankton concentration," *J. Acoust. Soc. Am.* **104**, 141–145 (1998).

<sup>4</sup>N. de Jong, F. J. T. Cate, C. T. Lancée, J. R. T. C. Roelandt, and N. Bom, "Principles and recent developments in ultrasound contrast agents," *Ultrasonics* **29**, 324–330 (1991).

<sup>5</sup>K. K. Shung, R. A. Sigelmann, and J. M. Reid, "Scattering of ultrasound by blood," *IEEE Trans. Biomed. Eng.* **BME-23**, 460–467 (1976).

<sup>6</sup>P. M. Morse and U. K. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).

<sup>7</sup>J. Szczucka, "Numerical model of multiple sound scattering from gas bubbles in the sea," *Arch. Acoust.* **20**, 191–205 (1995).

<sup>8</sup>I. Bjørnø and L. Bjørnø, "Numerical modelling of multiple scattering between two elastic particles," in *OCEANS'98 Conference Proceedings* (IEEE/OES, New Jersey, 1998) pp. 598–602.

<sup>9</sup>L. L. Foldy, "The multiple scattering of waves. I. General theory of isotropic scattering by randomly distributed scatterers," *Phys. Rev.* **67**, 107–119 (1945).

<sup>10</sup>A. Sangani, "A pairwise interaction theory for determining the linear acoustic properties of the dilute bubble liquids," *J. Fluid Mech.* **232**, 221–284 (1991).

<sup>11</sup>Z. Ye and L. Ding, "Acoustic dispersion and attenuation relations in bubbly mixture," *J. Acoust. Soc. Am.* **98**, 1629–1636 (1995).

<sup>12</sup>F. S. Henyey, "Corrections to Foldy's effective medium theory for propa-

- gation in bubble clouds and other collections of very small scatterers," J. Acoust. Soc. Am. **105**, 2149–2154 (1999).
- <sup>13</sup>V. Twersky, "Multiple scattering of waves and optical phenomena," J. Opt. Soc. Am. **52**, 145–171 (1962).
- <sup>14</sup>I. Tolstoy, "Superresonant systems of scatterers. I," J. Acoust. Soc. Am. **80**, 282–294 (1986).
- <sup>15</sup>I. Tolstoy, "Erratum: Superresonant systems of scatterers. I," J. Acoust. Soc. Am. **81**, 1987 (1987).
- <sup>16</sup>G. C. Gaunaud, H. Huang, and H. C. Strifors, "Acoustic scattering by a pair of spheres," J. Acoust. Soc. Am. **98**, 495–507 (1995).
- <sup>17</sup>R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem* (Dover, New York, 1992).
- <sup>18</sup>V. C. Anderson, "Sound scattering from a fluid sphere," J. Acoust. Soc. Am. **22**, 426–431 (1950).
- <sup>19</sup>C. Devin, Jr., "Survey of thermal, radiation, and viscous damping of pulsating air bubbles in water," J. Acoust. Soc. Am. **31**, 1654–1667 (1959).
- <sup>20</sup>C. S. Clay and H. Medwin, *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1977).
- <sup>21</sup>L. R. Dragonette, R. H. Vogt, L. Flax, and W. G. Neubauer, "Acoustic reflection from elastic spheres and rigid spheres and spheroids. II. Transient analysis," J. Acoust. Soc. Am. **55**, 1130–1137 (1974).
- <sup>22</sup>I. A. Stegun, "Legendre functions," in *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (U.S. Government Printing Office, Washington, DC, 1972).
- <sup>23</sup>J. J. Faran, Jr., "Sound scattering by solid cylinders and spheres," J. Acoust. Soc. Am. **23**, 405–418 (1951).
- <sup>24</sup>R. Hickling, "Analysis of echoes from a solid elastic sphere in water," J. Acoust. Soc. Am. **34**, 1582–1592 (1962).
- <sup>25</sup>K. G. Foote, "Optimizing copper spheres for precision calibration of hydroacoustic equipment," J. Acoust. Soc. Am. **71**, 742–747 (1982).
- <sup>26</sup>T. G. Leighton, D. G. Ramble, and A. D. Phelps, "The detection of tethered and rising bubbles using multiple acoustic techniques," J. Acoust. Soc. Am. **101**, 2626–2635 (1997).
- <sup>27</sup>F. H. Fisher and V. P. Simmons, "Sound absorption in sea water," J. Acoust. Soc. Am. **62**, 558–564 (1977).
- <sup>28</sup>R. Clift, J. R. Grace, and M. E. Weber, *Bubbles, Drops, and Particles* (Academic, New York, 1978).
- <sup>29</sup>M. S. Longuet-Higgins, B. R. Kerman, and K. Lunde, "The release of air bubbles from an underwater nozzle," J. Fluid Mech. **230**, 365–390 (1991).
- <sup>30</sup>L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982).
- <sup>31</sup>R. L. Pritchard, "Mutual acoustic impedance between radiators in an infinite rigid plane," J. Acoust. Soc. Am. **32**, 730–737 (1960).
- <sup>32</sup>S. Sugimoto, A. Hasegawa, and T. Kikuchi, "Active sound power reduction based on mutual radiation impedance effect between transducers in water," Jpn. J. Appl. Phys. **33**, 3294–3299 (1994).

# Transformation of seismic velocity data to extract porosity and saturation values for rocks

James G. Berryman, Patricia A. Berge, and Brian P. Bonner

*Lawrence Livermore National Laboratory, P.O. Box 808 L-200, Livermore, California 94551-9900*

(Received 5 December 1999; accepted for publication 17 March 2000)

For wave propagation at low frequencies in a porous medium, the Gassmann–Domenico relations are well-established for homogeneous partial saturation by a liquid. They provide the correct relations for seismic velocities in terms of constituent bulk and shear moduli, solid and fluid densities, porosity and saturation. It has not been possible, however, to invert these relations easily to determine porosity and saturation when the seismic velocities are known. Also, the state (or distribution) of saturation, i.e., whether or not liquid and gas are homogeneously mixed in the pore space, is another important variable for reservoir evaluation. A reliable ability to determine the state of saturation from velocity data continues to be problematic. It is shown how transforming compressional and shear wave velocity data to the  $(\rho/\lambda, \mu/\lambda)$ -plane (where  $\lambda$  and  $\mu$  are the Lamé parameters and  $\rho$  is the total density) results in a set of quasi-orthogonal coordinates for porosity and liquid saturation that greatly aids in the interpretation of seismic data for the physical parameters of most interest. A second transformation of the same data then permits isolation of the liquid saturation value, and also provides some direct information about the state of saturation. By thus replotting the data in the  $(\lambda/\mu, \rho/\mu)$ -plane, inferences can be made concerning the degree of patchy (inhomogeneous) versus homogeneous saturation that is present in the region of the medium sampled by the data. Our examples include igneous and sedimentary rocks, as well as man-made porous materials. These results have potential applications in various areas of interest, including petroleum exploration and reservoir characterization, geothermal resource evaluation, environmental restoration monitoring, and geotechnical site characterization. © 2000 Acoustical Society of America. [S0001-4966(00)05306-6]

PACS numbers: 43.20.Jr, 43.20.Gp, 43.60.Pt, 43.30.Ma [DEC]

## INTRODUCTION

In a variety of applied problems, it is important to determine the state of saturation of a porous medium from acoustic or seismic measurements. In the oil and gas industry, it is common to use amplitude-versus-offset (AVO) processing of seismic reflection data to reach conclusions about the presence of gas, oil, and their relative abundances on the opposite sides of a reflecting interface underground (e.g., Ref. 1). For environmental applications, we can expect to be working in the near surface where sensor geometries other than surface reflection surveys become practical. For example, when boreholes are present, it is possible to do cross-well seismic tomography, or borehole sonic logging to determine velocities (e.g., Ref. 2). For AVO processing the data obtained are the seismic impedances  $\rho v_p$  and  $\rho v_s$  (where  $\rho$  is the density, and  $v_p, v_s$  are the seismic compressional and shear wave velocities, respectively), which arise naturally in reflectance measurements. (In this article, we will use the term “velocities” to refer to measured velocities at seismic, sonic, or ultrasonic frequencies, unless otherwise specified.) However, for cross-well applications, we are more likely to have simply velocity data, i.e.,  $v_p$  and  $v_s$  themselves without density information. For well-logging applications, separate measurements of the velocities as well as density are possible. Although a great deal of effort has been expended on AVO analysis, relatively little has been done to invert the simple velocity data for porosity and saturation. It is our purpose to present one method that shows promise for using

velocity data to obtain porosity and saturation estimates. The key physical idea used here is the fact that the Lamé parameter  $\lambda$  and the density  $\rho$  are the two parameters containing information about saturation, while both of these together with shear modulus  $\mu$  contain information about porosity ( $\lambda$  and  $\mu$  are defined in the next section). These facts are well known from earlier work of Gassmann,<sup>3</sup> Domenico,<sup>4</sup> and many others. (It is well-established that even though the Gassmann–Domenico relations are derived for the static case, they have been found to describe behavior measured in the field at sonic and seismic frequencies, and, in some cases, even in laboratory ultrasonic experiments.) The same facts are used explicitly in AVO analysis,<sup>1,5–7</sup> but in ways that are significantly different from those to be described here. A major point of departure is that the present work allows direct information to be obtained about not only the level of the saturation, but also concerning the state of saturation, i.e., whether the liquid and gas present are mixed homogeneously, or are instead physically separated and therefore in a state of patchy saturation.<sup>8–13</sup> Another advantage is that this method uses velocity rather than amplitude information, and therefore may have less uncertainty and may also require less data processing for some types of field experiments.

One of the main points of the analysis to be presented is the purposeful avoidance of the well-known complications that arise at high frequencies, due in large part to velocity dispersion and attenuation.<sup>14–23</sup> Our point of view is that seismic data (as well as most sonic, and some ultrasonic

data) do not suffer from contamination by the frequency-dependent effects to the same degree typically seen for high-frequency laboratory measurements. By restricting our range of frequencies to those most useful in the field, we anticipate a significant simplification of the analysis and therefore an improvement in our ability to provide both simple and robust interpretations of field data. In Sec. III, we also provide a means of identifying data in need of correction for dispersion effects.

We introduce the basic physical ideas in the next section. Then we present two new methods of displaying the velocity data. One method is used to sort data points into sets that have similar physical attributes, such as porosity. Then, the second method is used to identify both the level of saturation and the type of saturation, whether homogeneous, patchy, or a combination of the two. We show a subset of the large set of data we have examined that confirms these conclusions empirically. We then provide some discussion of the results and what we foresee as possible future applications of the ideas. Finally, we summarize the accomplishments of the article in the concluding section.

## I. ELASTIC AND POROELASTIC WAVE PROPAGATION

For isotropic elastic materials there are two bulk elastic wave speeds,<sup>24</sup> compressional  $v_p = \sqrt{(\lambda + 2\mu)/\rho}$  and shear  $v_s = \sqrt{\mu/\rho}$ . Here  $\rho$  is the overall density, and the Lamé parameters  $\lambda$  and  $\mu$  are the constants that appear in Hooke's law relating stress to strain in an isotropic material. The constant  $\mu$  gives the dependence of shear stress on shear strain in the same direction. The constant  $\lambda$  gives the dependence of any compressional/tensional uniaxial stress on extensional/dilatational uniaxial strains in the two orthogonal directions, i.e.,  $\lambda$  is the off-diagonal component of the isotropic stiffness tensor in the absence of shearing effects. For a porous system with porosity  $\phi$  (void volume fraction) in the range  $0 < \phi < 1$ , the overall density of the rock or sediment is just the volume weighted density given by

$$\rho = (1 - \phi)\rho_s + \phi[S\rho_l + (1 - S)\rho_g], \quad (1)$$

where  $\rho_s$ ,  $\rho_l$ ,  $\rho_g$  are the densities of the constituent solid, liquid and gas, respectively. Here  $S$  is the liquid saturation, i.e., the fraction of liquid-filled void space in the range  $0 \leq S \leq 1$  (see Ref. 4). When liquid and gas are distributed uniformly in all pores and cracks, Gassmann's equations say that, for quasistatic isotropic elasticity and low-frequency wave propagation, the shear modulus  $\mu$  will be mechanically independent of the properties of any fluids present in the pores, while the overall bulk modulus  $K \equiv \lambda + \frac{2}{3}\mu$  of the rock or sediment including the fluid depends in a known way on porosity and elastic properties of the fluid and dry rock or sediment.<sup>3</sup> Thus, in the Gassmann model, the Lamé parameter  $\lambda$  is elastically dependent on fluid properties, while  $\mu$  is not. The density  $\rho$  also depends on saturation, as shown in Eq. (1). At low liquid saturations, the bulk modulus of the fluid mixture is dominated by the gas, and therefore the effect of the liquid on  $\lambda$  is negligible until the porous medium approaches full saturation. This means that both velocities  $v_p$

and  $v_s$  will decrease with increasing fluid saturation<sup>4</sup> due to the "density effect," wherein the only quantity changing is the density, which increases in the denominators of both  $v_p^2$  and  $v_s^2$ . As the medium approaches full saturation, the shear velocity continues its downward trend, while the compressional velocity suddenly (over a very narrow range of saturation values) shoots up to its full saturation value. A well-known example of this behavior was provided by Murphy.<sup>25</sup> Figure 1 shows how plots of these data for sandstones will appear in several choices of display, with Fig. 1(a) being one of the more common choices. This is the expected (ideal Gassmann–Domenico) behavior of partially saturated porous media. The Gassmann–Domenico relations hold for frequencies low enough (sonic and below) that the solid frame and fluid will move in phase, in response to applied stress or displacement. The fluid pressure must be (at least approximately) uniform throughout the porous medium, from which assumption follows the homogeneous saturation requirement.

## II. PREDICTIONS OF THE THEORY AND EXAMPLES

### A. Gassmann–Domenico relations

Gassmann's equations<sup>3</sup> for fluid substitution state that

$$K = K_{dr} + \frac{\alpha^2}{(\alpha - \phi)/K_m + \phi/K_f} \quad \text{and} \quad \mu = \mu_{dr}, \quad (2)$$

where  $K_m$  is the bulk modulus of the single solid mineral and  $K_{dr}$  and  $\mu_{dr}$  are the bulk and shear moduli of the drained porous frame. The special combination of moduli defined by  $\alpha = 1 - K_{dr}/K_m$  is the Biot–Willis<sup>26</sup> parameter. The porosity is  $\phi$ , while  $K$  and  $\mu$  are the effective bulk and shear moduli of the undrained porous medium that is saturated with a fluid mixture having bulk modulus  $K_f$ . For partial saturation conditions with homogeneous mixing of liquid and gas, so that all pores contain the same relative proportions of liquid and gas, Domenico<sup>4</sup> among others shows that

$$1/K_f = S/K_l + (1 - S)/K_g. \quad (3)$$

The saturation level of liquid is  $S$  lying in the range  $0 \leq S \leq 1$ . The bulk moduli are  $K_l$  for the liquid, and  $K_g$  for the gas. When  $S$  is small, (3) shows that  $K_f \approx K_g$ , since  $K_g \ll K_l$ . As  $S \rightarrow 1$ ,  $K_f$  remains close to  $K_g$  until  $S$  closely approaches unity. Then,  $K_f$  changes rapidly (over a small range of saturations) from  $K_g$  to  $K_l$ . (Note that the value of  $K_l$  may be several orders of magnitude larger than  $K_g$ , as in the case of water and air—2.25 GPa and  $1.45 \times 10^{-4}$  GPa, respectively.)

Since  $\mu$  has no mechanical dependence on the fluid saturation, it is clear that all the fluid dependence of  $K = \lambda + (2/3)\mu$  in (2) resides within the Lamé parameter  $\lambda$ . Other recent work<sup>27</sup> on layered elastic media indicates that  $\lambda$  should be considered as an important independent variable for analysis of wave velocities, and Gassmann's results provide some confirmation of this deduction (and furthermore provided a great deal of the motivation for the present line of research). The parameters  $K$  ( $= \lambda + (2/3)\mu$ ) and  $K_{dr}$  ( $= \lambda_{dr} + (2/3)\mu_{dr}$ ) can be replaced in (2) by  $\lambda$  and  $\lambda_{dr}$  without changing the validity of the equation. Thus, like  $K$ , for

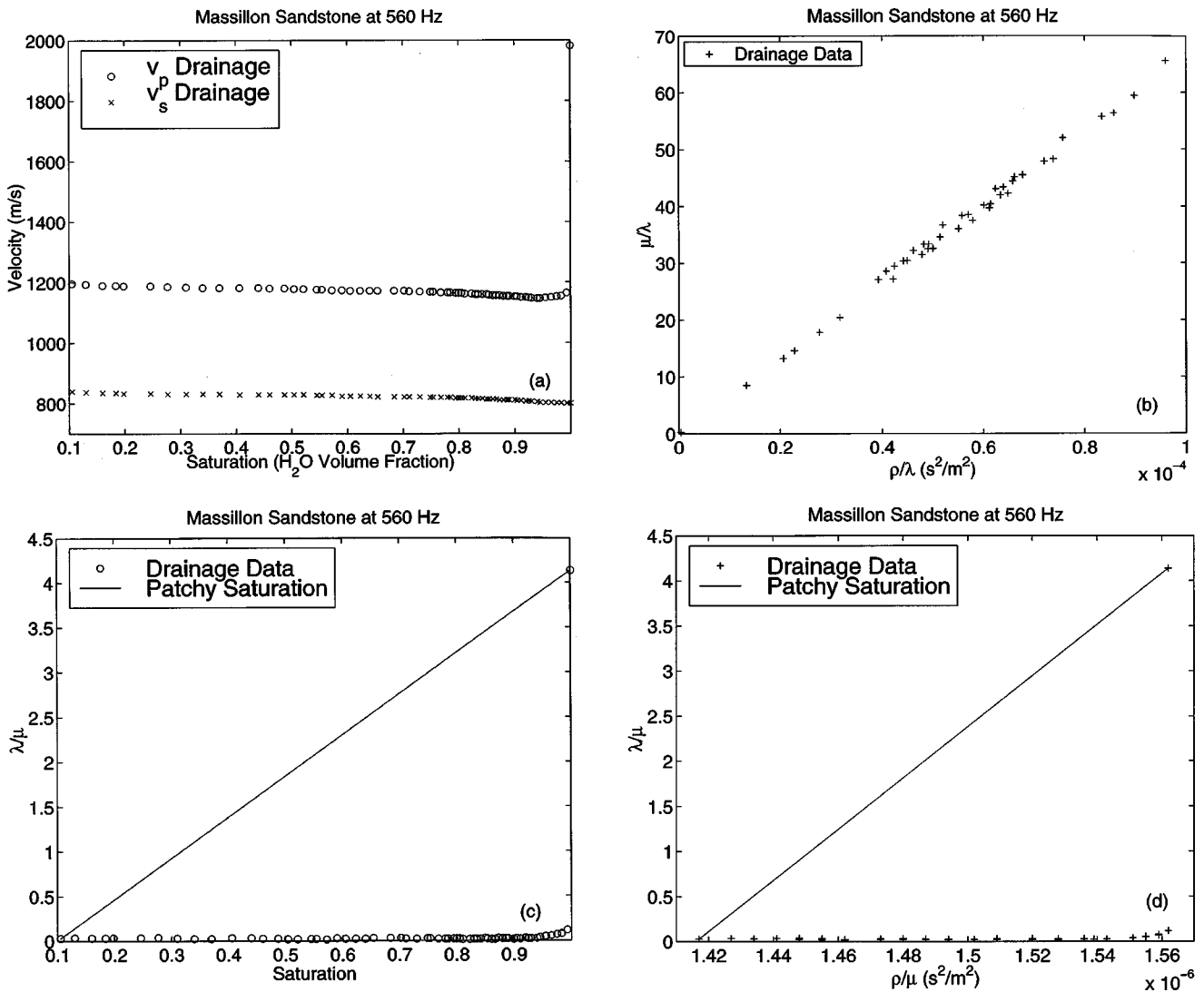


FIG. 1. Various methods of plotting 560-Hz Massillon sandstone data of Murphy (Ref. 25): (a) Compressional and shear wave velocities as a function of saturation, (b) transform to  $(\rho/\lambda, \mu/\lambda)$ -plane, (c)  $\lambda/\mu$  versus saturation, and (d) transform to  $(\lambda/\mu, \rho/\mu)$ -plane. All of these behaviors are anticipated by the Gassmann–Domenico relations for homogeneously mixed fluid in the pores.

increasing saturation values  $\lambda$  will be almost constant until the porous medium closely approaches full saturation.

Now the first problem that arises with field data is that we usually do not know the reason why data collected at two different locations in the earth differ. It could be that the differences are all due to the saturation differences we are concentrating on in this paper. Or it could be that they are due entirely or only partly to differences in the porous solids that contain the fluids. In fact, solid differences easily can mask any fluid differences because the range of detectable solid mechanical behavior is so much greater than that of the fluids (especially when fractures are present).

It is essential to remove such differences due to solid heterogeneity. A related issue concerns differences arising due to porosity changes throughout a system of otherwise homogeneous solids. One way of doing this would be to sort our data into sets having similar porous solid matrix. For simplicity and because of the types of laboratory data sets available, we will use porosity here as our material discriminant.

Considering our three main parameters,  $\lambda$ ,  $\mu$ , and  $\rho$ , we

see that all three depend on porosity, but only  $\lambda$  and  $\rho$  depend on saturation. Using formulas (1)–(3), we can take partial derivatives of each of these expressions first with respect to  $\phi$  while holding  $S$  constant, and then with respect to  $S$  while holding  $\phi$  constant. For now, we are only interested in trends rather than the exact values, and these are displayed in Table I. The trend for  $\partial\lambda/\partial S|_{\phi} > 0$  requires the additional reminder that, although this term is always positive, its value is often so small that it may be treated as zero except in the small range of values close to  $S=1$ . Also, using Hashin–

TABLE I. Monotonicity properties of the Lamé parameters  $\lambda$  and  $\mu$  and the density  $\rho$  as the porosity  $\phi$  and liquid saturation  $S$  vary.

	Lamé $\lambda$	Lamé $\mu$	Density $\rho$
$\Delta\phi$	$\frac{\partial\lambda}{\partial\phi} _S = ?$	$\frac{\partial\mu}{\partial\phi} _S < 0$	$\frac{\partial\rho}{\partial\phi} _S < 0$
$\Delta S$	$\frac{\partial\lambda}{\partial S} _{\phi} > 0$ (or $\approx 0$ )	$\frac{\partial\mu}{\partial S} _{\phi} = 0$	$\frac{\partial\rho}{\partial S} _{\phi} > 0$

Shtrikman bounds<sup>28</sup> as a guide, it turns out that it is not possible to make a general statement about the sign of  $\partial\lambda/\partial\phi|_S$ , since the result depends on the particular material constants. (Related differences of sign are also observed in the data we show later in this article; thus, this ambiguity is definitely real and observable.)

Assuming that the primary variables are  $\lambda$ ,  $\mu$ , and  $\rho$  (further justification of this choice of primary variables is provided later in the paper), then the two pieces of velocity data we have can be used to construct the following three ratios:

$$\frac{\mu}{\lambda} = \frac{v_s^2}{v_p^2 - 2v_s^2}, \quad (4)$$

$$\frac{\rho}{\lambda} = \frac{1}{v_p^2 - 2v_s^2}, \quad (5)$$

and

$$\frac{\rho}{\mu} = \frac{1}{v_s^2}. \quad (6)$$

We will consider first of all what happens to these ratios for homogeneous mixing of fluids, and then consider the simpler case of ideal patchy saturation, where some pores in the partially saturated medium are completely filled with liquid and others are completely dry (or filled with gas).

## B. Homogeneous saturation

For homogeneous saturation, as  $S$  varies while porosity remains fixed, the ratio  $\mu/\lambda$  does not change significantly until  $S \rightarrow 1$ . At that point,  $\lambda$  increases dramatically and  $\mu/\lambda$  therefore decreases dramatically. Similarly, as  $S \rightarrow 1$ , the only changes in  $\rho/\lambda$  over most of the dynamic range of  $S$  are in  $\rho$ , which increases linearly with  $S$ . Then, when  $\rho$  is almost at its maximum value,  $\lambda$  increases dramatically, causing the  $\rho/\lambda$  ratio to decrease dramatically. Thus,  $\rho/\lambda$  does not change monotonically with  $S$ , but first increases a little and then decreases a lot. These two ratios may be conveniently compared by plotting data from various rocks and man-made porous media examples in the  $(\rho/\lambda, \mu/\lambda)$ -plane [see Figs. 1(b) and 2]. We see that, when data are collected at approximately equal intervals in  $S$ , the low saturation points will all cluster together with nearly constant  $\mu/\lambda$  and small increases in  $\rho/\lambda$ , but the final steps as  $S \rightarrow 1$  lead to major decreases in both ratios. The resulting plots appear as nearly straight lines in this plane, with drained samples plotting to the upper right and fully saturated samples plotting to the lower left in each of the examples shown in Fig. 2. The remaining ratio  $\rho/\mu$  has the simplest behavior, since  $\rho$  increases monotonically in  $S$ , and  $\mu$  does not change. So  $\rho/\mu$  is a monotonically increasing function of  $S$ , and therefore can be considered a useful proxy of the saturation variable  $S$ . [Compare Fig. 1(c) and (d), and see Fig. 3.]

Figure 2(a) includes the same sandstone data from Fig. 1, along with other sandstone data. Similar data for five limestone samples<sup>13</sup> are plotted in Fig. 2(b). The straight line correlation of the data in the sandstone display is clearly

reconfirmed by the limestone data. Numerous other examples of the correlation have been observed. [Fully dry and fully saturated examples are shown here for some of these examples in Fig. 2(c) and (d),<sup>29,30</sup> for which partial saturation data were unavailable.] No examples of appropriate data for partially saturated samples have exhibited major deviations from this behavior, although an extensive survey of available data sets has been performed for materials including limestones,<sup>13</sup> sandstones,<sup>10,25</sup> granites,<sup>30</sup> unconsolidated sands, and some artificial materials such as ceramics and glass beads.<sup>29</sup> This straight line correlation is a very robust feature of partial saturation data. The mathematical trick that brings about this behavior will now be explained.

Consider the behavior as  $\phi$  increases for fixed  $S$ . Two of the parameters ( $\mu$  and  $\rho$ ) decrease as  $\phi$  increases, but at different rates, while the third ( $\lambda$ ) can have arbitrary variation. [Recall<sup>31</sup> that rigorous bounds on the parameters are  $0 \leq K < \infty$ ,  $0 \leq \mu < \infty$ ,  $0 < \rho < \infty$ , and  $-(2/3)\mu \leq \lambda \leq \infty$ .] To understand the behavior on these plots in Fig. 1 as  $\phi$  changes, it will prove convenient to consider polar coordinates  $(r, \theta)$ , defined by

$$r^2 = w^4 \left( \frac{\rho}{\lambda} \right)^2 + \left( \frac{\mu}{\lambda} \right)^2 \quad (7)$$

and

$$\tan \theta = \frac{\mu}{\rho w^2}, \quad (8)$$

where  $w$  is an arbitrary scale factor with dimensions of velocity (chosen so that  $r$  is a dimensionless radial coordinate for plots like those in Fig. 1). Now, if in addition we choose  $w$  to be sufficiently large so that  $v_s/w \ll 1$  for typical values of  $v_s$  in our data sets, then, using standard perturbation expansions, we have

$$r = \frac{\rho}{\lambda} w^2 \left( 1 + \frac{v_s^4}{w^4} \right)^{1/2} \approx \frac{\rho}{\lambda} w^2 \left( 1 + \frac{v_s^4}{2w^4} \right) \quad (9)$$

and

$$\theta = \tan^{-1} \left( \frac{v_s^2}{w^2} \right) \approx \frac{v_s^2}{w^2}. \quad (10)$$

Thus, the angle  $\theta$  is well approximated by the ratio in (10), which depends only on the shear velocity  $v_s$ . We know the shear velocity is a rather weak function of saturation [e.g., Fig. 1(a)], but a much stronger function of porosity (see, for example, Ref. 29). So we see that the angle in these plots is most strongly correlated with changes in the porosity. In contrast, the radial position  $r$  is principally dependent on the ratio  $\rho/\lambda$ , which we have already shown to be a strong function of the saturation  $S$ , especially in the region close to full liquid saturation. This analysis shows why the plots in Figs. 1(b) and 2 look the way they do and also why we might be inclined to call these quasi-orthogonal (polar) plots of saturation and porosity. Because of the function these plots play in our analysis, we will call them the ‘‘data-sorting’’ plots.

In contrast, the plots in Fig. 3 contain information about fluid spatial distribution, as will be discussed at greater



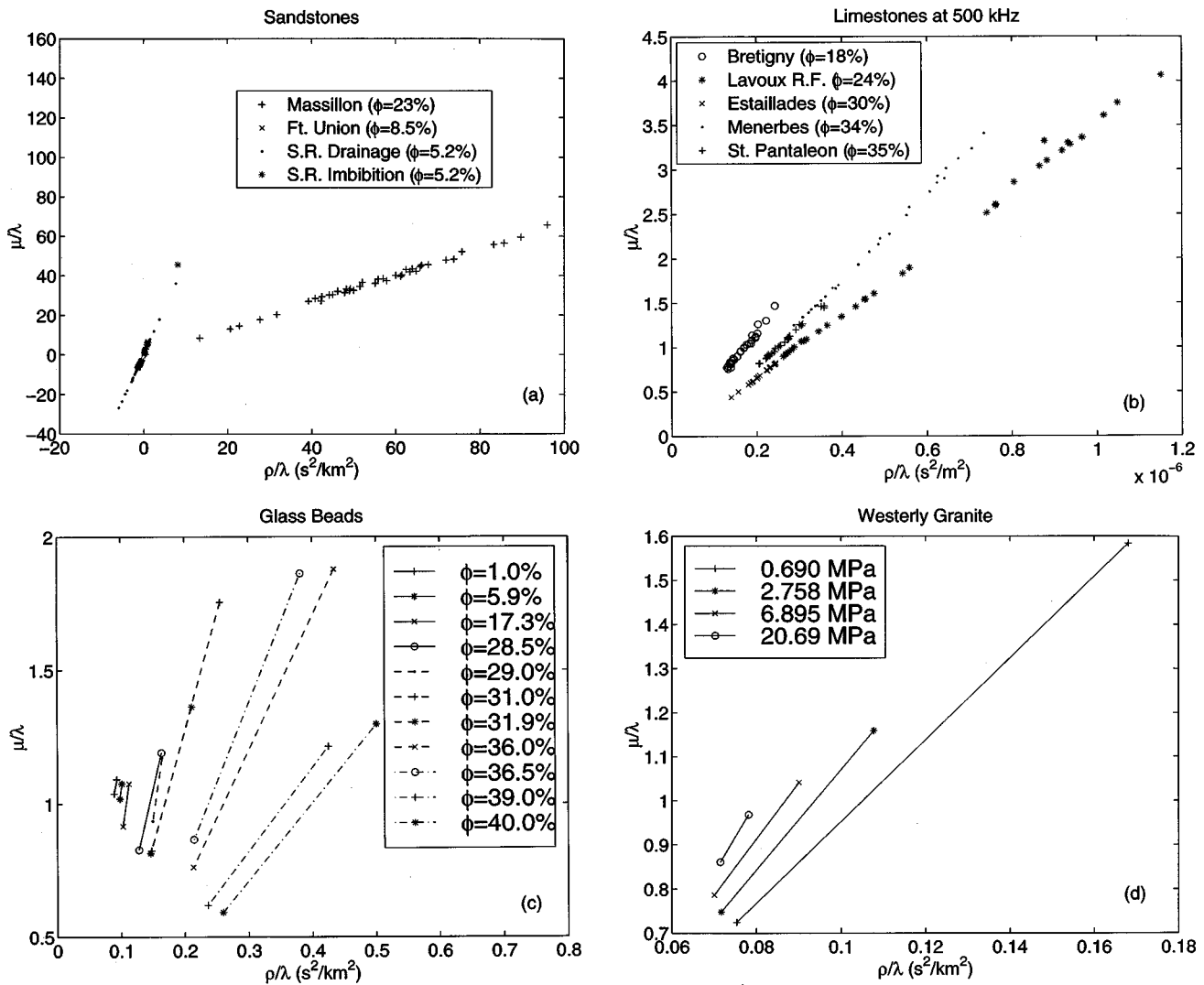


FIG. 2. Examples of the correlation of slopes with porosity in the data-sorting plots: (a) three Spirit River (S.R.) sandstone (Ref. 10) and Massillon and Ft. Union sandstones (Ref. 25), (b) five limestones (Ref. 13), (c) 11 fused glass-bead samples (Ref. 29), (d) Westerly granite (Ref. 30) at four pressures. The observed trend is that high porosity samples generally have lower slopes than lower porosities on these plots, although there are a few exceptions as discussed in the text. These trends are easily understood since the slopes are determined approximately by the average value of  $v_s^2$  for each material, which is a decreasing function of porosity  $\phi$ .

length later in this paper. The bulk modulus  $K_f$  contains the only  $S$  dependence in (2). Thus, for porous materials satisfying Gassmann's homogeneous fluid conditions and for low enough frequencies, the theory predicts that, if we use velocity data in a two-dimensional plot with one axis being the saturation  $S$  and the other being the ratio  $\lambda/\mu = (v_p/v_s)^2 - 2$ , then the results will lie along an essentially straight (horizontal) line until the saturation reaches  $S \approx 1$  (around 95% or higher), where the curve formed by the data will quickly rise to the value determined by the velocities at full liquid saturation. On such a plot, the drained data appear in the lower left while the fully saturated data appear in the upper right. This behavior is illustrated in Fig. 3(a) for Espeil limestone. The behavior of the other plots in Fig. 3 will be described below.

Before leaving this discussion of homogeneous saturation, we should note that there is one laboratory saturation technique for which it is known—from direct observations<sup>13</sup> using x-ray imaging—that very homogeneous liquid–gas

mixtures will generally be produced. This method is called “depressurization.” When such data are available (see Fig. 3), we expect they will always behave according to the Gassmann–Domenico predictions. In contrast, the more common approach which produces drainage data is less predictable, since the manner and rate of drainage depend strongly on details of particular samples—especially on surface energies that control capillarity and on permeability magnitude and distribution. Thus, the drainage technique can produce homogeneous saturation, or patchy saturation, or anything in between.

### C. Patchy saturation

The preceding analysis centered on homogeneous saturation of porous media. On the other hand, consider a porous medium containing gas and liquid mixed in a heterogeneous manner, so that patches of the medium hold only gas while other patches hold only liquid in the pores. Then, the theory

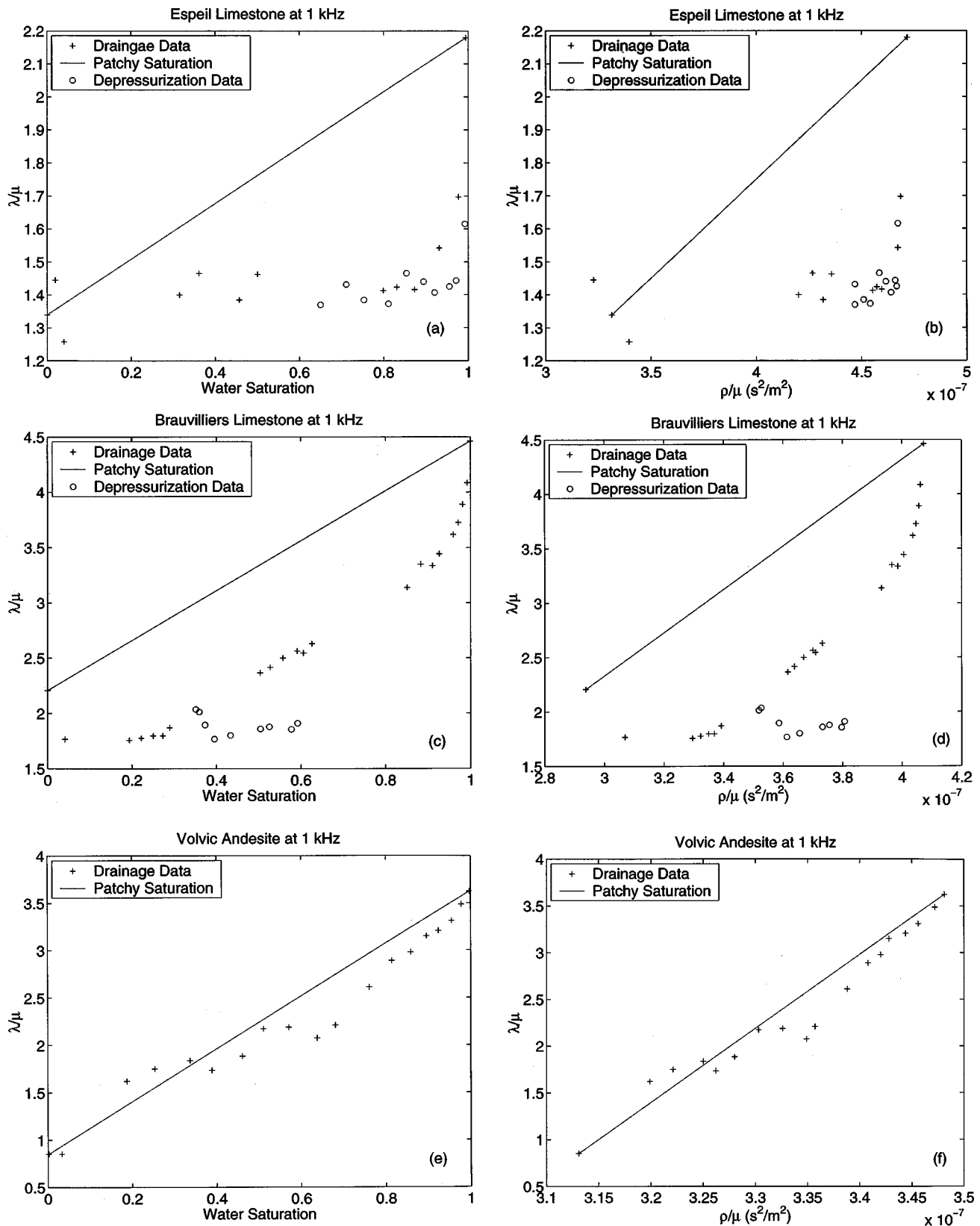


FIG. 3. Lamé parameter ratio  $\lambda/\mu$  plotted versus (a) saturation and (b)  $\rho/\mu$  for Espelil limestone, (c) saturation and (d)  $\rho/\mu$  for Brauvilliers limestone, and (e) saturation and (f)  $\rho/\mu$  for Volvic andesite. All extensional and shear wave measurements (Refs. 13, 34 and 35) were made at 1 kHz. Note that (a) and (b) indicate homogeneous mixing of liquid and gas, and (c) and (d) show an intermediate state of mixing for the drainage data, but more homogeneous mixing for the depressurization data. The plots on the right are saturation-proxy plots, having essentially the same behavior as the plots on the left but requiring only velocity data.

predicts that, depending to some extent on the spatial distribution of the patches, the results will deviate overall from Gassmann's results (although Gassmann's results will hold locally in each individual patch). If we consider the most extreme cases of spatial distribution possible, which are laminated regions of alternating liquid saturation and gas saturation, then the effective bulk modulus will be determined by an average of the two extreme values of (2):  $K|_{S=0} = K_{dr}$  and  $K|_{S=1}$ . Using saturation as the weighting factor, the harmonic mean and the mean are the two well-known extremes of behavior.<sup>32</sup> Of these two, the one that differs most from (2) for  $0 < S < 1$  is the mean. And, because of  $K$ 's linear dependence on both  $\lambda$  and  $\mu$ , and  $\mu$ 's independence of  $S$ , we therefore have

$$\lambda_{\text{patchy}}(S) = (1 - S)\lambda_{dr} + S\lambda|_{S=1}. \quad (11)$$

So, on our plot in the  $(\lambda/\mu, \rho/\mu)$ -plane, the results for the mean will again lie along a straight line, but now the line goes directly from the unsaturated value ( $S=0$ ) to the fully saturated value ( $S=1$ ) [e.g., Fig. 3(e)]. The two straight lines described [the one given by (11) and the horizontal one discussed in the preceding paragraph for saturations up to about 95%] are rigorous results of the theory, and form two sides of a triangle that will contain all data for partially saturated systems, regardless of the type of saturation present. The third side of this triangle provides a rigorous bound on the behavior as full saturation is approached (it just corresponds to the physical requirement that  $S \leq 1$ , so values with  $S > 1$  have no physical significance). In general, heterogeneous fluid distribution can produce points anywhere within the resulting triangle, but not outside the triangle (within normal experimental error).

A brief presentation of some examples (Fig. 3) will now follow a reminder of an important and well-known caveat.

### D. Caveats for chemical effects

Some deviations from these conclusions can be expected at the lowest saturations. Chemical effects, which have not been accounted for in the mechanical analysis, can and often do lead to the situation that dry and drained (nearly dry or room dry) samples have somewhat different properties.<sup>33</sup> These differences are larger than can be explained by mechanical analyses alone. [For example, see Fig. 3(a) and (b). Take special note of the three lowest saturation values in these figures.] We discuss this point at greater length in Sec. III.

### E. Some illustrative examples

Figure 3 shows three examples of the results obtained with plots in the  $(S, \lambda/\mu)$ -plane and in the  $(\rho/\mu, \lambda/\mu)$ -plane (using  $\rho/\mu$  as a proxy for  $S$ ) for two limestones and one andesite from laboratory data of Carodet *et al.*<sup>13,34,35</sup> In Fig. 3, the true saturation data are used in the figures on the left and the proxy for saturation ( $\rho/\mu$ ) is used on the right. We therefore call the right-hand diagrams "saturation-proxy" plots. Using the interpretations arising for our analysis of Gassmann–Domenico partial-saturation theory, we see that Fig. 3(a) and (b) indicates homogeneous mixing of liq-

uid and gas, while Fig. 3(e) and (f) indicates extremely patchy mixing, and Fig. 3(c) and (d) shows an intermediate state of mixing for the drainage data, but more homogeneous mixing (as expected) for the depressurization data. The Espeil limestone was observed to be the most dispersive of all those rocks considered in the data sets of Carodet *et al.*<sup>13,34,35</sup> So, this case is a very stringent test of the method. In fact, if we were to plot the corresponding data for Espeil limestone at 500 kHz, we would not find such simple and easily interpreted behavior on these plots. Our explanation for this difference between the 500- and 1-kHz results for Espeil limestone is that the dispersion introduces effects not accounted for by the simple Gassmann–Domenico theory, and that there is then no reason to think that our method should work for such high frequencies as 500 kHz. We have found other examples where it does work for frequencies higher than one might expect the method to be valid. The point is that, if we restrict the range of frequencies considered to 1 kHz or less, the method appears to work quite well on most (and perhaps all) samples. (But, at higher frequencies, the solid and fluid can move out of phase and other relations developed by Biot<sup>14–16</sup> and others<sup>17–23</sup> apply.)

## III. DISCUSSION

### A. Rocks containing more than one mineral

The analysis presented here has been limited for simplicity to the case of single mineral porous rocks. In fact, the main parts of the analysis do not change in any significant way if the rock has multiple constituents. The well-known result of Brown and Korrington<sup>36</sup> states that

$$K = K_{dr} + \frac{\alpha^2}{\alpha/K_s - \phi/K_\phi + \phi/K_f} \quad \text{and} \quad \mu = \mu_{dr}, \quad (12)$$

where  $K_s$  is theunjacketed bulk modulus of the composite solid frame,  $K_\phi$  is theunjacketed pore modulus of the composite solid frame, and  $\alpha = 1 - K_{dr}/K_s$  is the appropriate Biot–Willis<sup>26</sup> parameter for this situation. The remaining parameters have the same significance as in (2). The functional dependence of  $K_{sat}$  on the saturation  $S$  is clearly the same in both formulas. If we were trying to infer properties of the solid from these formulas, then of course (12) would be more difficult to interpret. But for our present purposes, we are only trying to infer porosity, saturation values, and saturation state. For these physical parameters, the analysis goes through without change.

### B. On uniqueness of $\lambda$ -diagrams

Since the possible linear combinations of the elastic bulk and shear moduli ( $K$  and  $\mu$ ) are infinite, it is natural to ask why (or if) the choice  $\lambda = K - (2/3)\mu$  is special. Is there perhaps some other combination of these constants that works as well or even better than the choice made here? There are some rather esoteric reasons based on recent work<sup>27</sup> in the analysis of layered anisotropic elastic media that lead us to believe that the choice  $\lambda$  is indeed special, but we will not try to describe these reasons here. Instead we will point out some general features of the two types of plots that make it clear that this choice is generally good, even though

others might be equally good or even better in special circumstances. First, in the diagram using the  $(\rho/\mu, \lambda/\mu)$ -plane, it is easy to see that *any* plot of data using linear combinations of the form  $(\rho/\mu, (\lambda + c\mu)/\mu)$ , where  $c$  is any real constant, will have precisely the same information and the display will be identical except for a translation of the values along the ordinate by the constant value  $c$ . Thus, for example, taking  $c = (2/3)$ , plots of  $(\rho/\mu, K/\mu)$  will have exactly the same interpretational value as those presented here. But, if we now reconsider the data-sorting plot (e.g., Fig. 2) for each of these choices, we need to analyze plots of the form  $(\rho/(\lambda + c\mu), \mu/(\lambda + c\mu))$ . Is there an optimum choice of the parameter  $c$  that makes the plots as straight as possible whenever the only variable is the fluid saturation? It is not hard to see that the class of best choices always lies in the middle of the range of values of  $\lambda/\mu$  taken by the data. So setting  $-c = (1/2)(\min(\lambda/\mu) + \max(\lambda/\mu))$  will always guarantee that there are very large positive and negative values of  $\mu/(\lambda + c\mu)$ , and therefore that these data fall reliably (if somewhat approximately) along a straight line. But the minimum value of  $\lambda/\mu$  has an absolute minimum of  $-(2/3)$ , based on the physical requirement of positivity of  $K$ . So  $c < (2/3)$  is a physical requirement, and since  $\max \lambda/\mu \approx + (2/3)$  is a fairly typical value for porous rocks, it is expected that an optimum value of  $c \leq 0$  will generally be obtained using this criterion. Thus, plots based on bulk modulus  $K$  instead of  $\lambda$  will not be as effective in producing the quasi-orthogonality of porosity and saturation that we have obtained in the data-sorting style of plotting. We conclude that the choice  $\lambda$  is not unique (some other choices might be as good for special data sets), but it is nevertheless an especially simple choice and is also expected to be quite good for most real data.

### C. Transforming straight lines to straight lines

One important feature concerning connections between the points in the two planes  $(\rho/\lambda, \mu/\lambda)$  and  $(\rho/\mu, \lambda/\mu)$  is the fact that (with only a few exceptions that will be noted) straight lines in one plane transform into straight lines in the other. For example, points satisfying

$$\frac{\lambda}{\mu} = A + B \frac{\rho}{\mu} \quad (13)$$

in the  $(\rho/\mu, \lambda/\mu)$ -plane (where  $A$  and  $B$  are constant intercept and slope, respectively) then satisfy

$$\frac{\mu}{\lambda} = A^{-1} - A^{-1} B \frac{\rho}{\lambda} \quad (14)$$

in the  $(\rho/\lambda, \mu/\lambda)$ -plane. So long as  $A \neq 0$  in (13), the straight line in (13) transforms into the straight line in (14). This observation is very important because the straight line in (11) corresponds to a straight line in the saturation-proxy plot in the  $(\rho/\mu, \lambda/\mu)$ -plane. But this line transforms into a straight line in the data-sorting plot in the  $(\rho/\lambda, \mu/\lambda)$ -plane. In fact, the apparent straight line along which the data align themselves in these plots is just this transformed patchy saturation line.

When  $A = 0$  in (13) (which seems to happen rarely if ever in the real data examples, but needs to be considered in general), the resulting transformed line will just be one of constant  $\rho/\lambda = B^{-1}$ , which is a vertical line on the  $(\rho/\lambda, \mu/\lambda)$ -plane. The more interesting special case is when  $B = 0$ , in which situation  $\lambda/\mu = A$  or  $\mu/\lambda = A^{-1}$ . But this case includes that of Gassmann–Domenico for homogeneous mixing of the fluids at low to moderate saturation values. For  $B = 0$ , on both planes we have horizontal straight lines, but their lengths can differ significantly on the two displays.

## D. Interpreting the data point locations

### 1. Data points inside the triangle

The triangle described in Sec. III C provides rigorous bounds on mechanical properties of porous media. For plots in the  $(\rho/\mu, \lambda/\mu)$ -plane such as those included in Figs. 1(d) and 3(b), (d), and (f), some data points lie between the ideal patchy saturation line and the Gassmann ideal lower bound. The relative position of the data points may contain information about the fluid distribution. Consider the case of a core sample that is nearly saturated, above 90% for example. If the weight of the core is used to determine the saturation but the core contains a few gas bubbles, the background saturation will be underestimated and the bubbles themselves represent patches. This is an example of a material having a few isolated patches contained in an otherwise homogeneous partially saturated background. Such data would plot above but close to the Gassmann curve. In an analogous case for field seismic data, the background saturation may be known from measurements made at lower frequencies or in a nearby region, and it may be possible to use such information to determine the relative volume of patches. For data lying in the middle (i.e., between the bounding curves), some assumptions about fluid distribution could be made and then various estimates about patchy volumes could be applied to different models such as the Hashin–Shtrikman bounds<sup>28</sup> or effective medium theories. Exploration of these issues will be the subject of future work.

### 2. Data points outside the triangle

The sides of the triangle described above set rigorous boundaries for effects associated with homogeneous saturation and patchy saturation at low frequencies or for situations in which frequency-dependent dispersion can be neglected. However, when the data do not in fact satisfy these assumptions of the theory, plotting the data this way provides an opportunity to observe and interpret deviations from the behavior predicted by the theory. For example, data which plot above the patchy saturation line represent excessively stiff rock. One possible cause of systematically high stiffness values is frequency-dependent dispersion.<sup>14–23</sup> Chemical effects, not taken into account in the analysis, might also cause measurements to deviate systematically from predicted behavior. For example, adhesive effects associated with chemical reactions between pore fluid and solid constituents might cause systematically high values. Another consequence of rock–water interactions is softening of intragranular cements. In this case, data for susceptible rocks would system-

atically plot below the Gassmann line at low saturations. Direct indications from elastic data of rock–water interactions (e.g., see Ref. 33) may lead to new methods of determining other rock properties controlled by chemical effects, such as the tensile strength.

#### IV. CONCLUSIONS

We have shown that seismic/sonic velocity data can be transformed to polar coordinates that have quasi-orthogonal dependence on saturation and porosity. This observation is based on the Gassmann–Domenico relations, which are known to be valid at low frequencies. The transformation loses its effectiveness at high frequencies whenever dispersion becomes significant, because then Biot theory and/or other effects play important roles in determining the velocities. So the simple relations between  $v_p$ ,  $v_s$ , and  $\lambda$ ,  $\mu$ ,  $\rho$ , and  $S$  break down at high frequencies. Our results are, nevertheless, quite encouraging because the predicted relationships seem to work in many cases up to frequencies of 1 kHz, and in a few special cases to still higher frequencies. These results present a straightforward method for obtaining porosity, saturation, and some information about spatial distribution of fluid (i.e., patchy versus homogeneous) in porous rocks and sediments, from compressional and shear wave velocity data alone. These results have potential applications in various areas of interest, including petroleum exploration and reservoir characterization, geothermal resource evaluation, environmental restoration monitoring, and geotechnical site characterization. The methods may also provide physical insight suggesting new approaches to AVO data analysis.

#### ACKNOWLEDGMENTS

We thank Bill Murphy and Rosemarie Knight for providing access to their unpublished data files. We thank Norman H. Sleep for his insight clarifying the significance of our sorting method for plotting seismic data. This work was performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-ENG-48 and supported specifically by the Geosciences Research Program of the DOE Office of Energy Research within the Office of Basic Energy Sciences, Division of Engineering and Geosciences, and by the Environmental Management Science Program of the Office of Environmental Management and the Office of Energy Research.

<sup>1</sup>J. P. Castagna and M. M. Backus, *Offset-Dependent Reflectivity—Theory and Practice of AVO Analysis* (Society of Exploration Geophysicists, Tulsa, OK, 1993).  
<sup>2</sup>J. M. Harris, R. C. Nolen-Hoeksema, R. T. Langan, M. Van Schaack, S. K. Lazaratos, and J. W. Rector III, “High-resolution crosswell imaging of a west Texas carbonate reservoir: Part 1—Project summary and interpretation,” *Geophysics* **60**, 667–681 (1995).  
<sup>3</sup>F. Gassmann, “Über die elastizität poröser medien,” *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* **96**, 1–23 (1951); J. G. Berryman, “Origin of Gassmann’s equations,” *Geophysics* **64**, 1627–1629 (1999).  
<sup>4</sup>S. N. Domenico, “Effect of water saturation on seismic reflectivity of sand reservoirs encased in shale,” *Geophysics* **39**, 759–769 (1974).  
<sup>5</sup>W. J. Ostrander, “Plane-wave reflection coefficients for gas sands at non-normal angles of incidence,” *Geophysics* **49**, 1637–1648 (1984).  
<sup>6</sup>J. P. Castagna, M. L. Batzle, and R. L. Eastwood, “Relationship between

compressional-wave and shear-wave velocities in clastic silicate rocks,” *Geophysics* **50**, 571–581 (1985).  
<sup>7</sup>D. J. Foster, R. G. Keys, and D. P. Schmitt, “Detecting subsurface hydrocarbons with elastic wavefields,” in *Inverse Problems in Wave Propagation*, edited by G. Chavent, G. Papanicolaou, P. Sacks, and W. Symes (Springer, New York, 1997), pp. 195–218.  
<sup>8</sup>J. G. Berryman, L. Thigpen, and R. C. Y. Chin, “Bulk elastic wave propagation in partially saturated porous solids,” *J. Acoust. Soc. Am.* **84**, 360–373 (1988).  
<sup>9</sup>A. L. Endres and R. Knight, “The effect of microscopic fluid distribution on elastic wave velocities,” *Log Anal.* **30**, 437–444 (1989).  
<sup>10</sup>R. Knight and R. Nolen-Hoeksema, “A laboratory study of the dependence of elastic wave velocities on pore scale fluid distribution,” *Geophys. Res. Lett.* **17**, 1529–1532 (1990).  
<sup>11</sup>G. Mavko and R. Nolen-Hoeksema, “Estimating seismic velocities at ultrasonic frequencies in partially saturated rocks,” *Geophysics* **59**, 252–258 (1994).  
<sup>12</sup>J. Dvorkin and A. Nur, “Acoustic signatures of patchy saturation,” *Int. J. Solids Struct.* **35**, 4803–4810 (1998).  
<sup>13</sup>T. Cadoret, G. Mavko, and B. Zinszner, “Fluid distribution effect on sonic attenuation in partially saturated limestones,” *Geophysics* **63**, 154–160 (1998).  
<sup>14</sup>M. A. Biot, “Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range,” *J. Acoust. Soc. Am.* **28**, 168–178 (1956).  
<sup>15</sup>M. A. Biot, “Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range,” *J. Acoust. Soc. Am.* **28**, 179–1791 (1956).  
<sup>16</sup>M. A. Biot, “Mechanics of deformation and acoustic propagation in porous media,” *J. Appl. Phys.* **33**, 1482–1498 (1962).  
<sup>17</sup>R. J. O’Connell and B. Budiansky, “Viscoelastic properties of fluid-saturated cracked solids,” *J. Geophys. Res.* **82**, 5719–5736 (1977).  
<sup>18</sup>G. M. Mavko and A. Nur, “The effect of nonelliptical cracks on the compressibility of rocks,” *Geophysics* **83**, 4459–4468 (1978).  
<sup>19</sup>J. G. Berryman, “Elastic wave propagation in fluid-saturated porous media,” *J. Acoust. Soc. Am.* **69**, 416–424 (1981).  
<sup>20</sup>C. McCann and D. M. McCann, “A theory of compressional wave attenuation in noncohesive sediments,” *Geophysics* **50**, 1311–1317 (1985).  
<sup>21</sup>D. L. Johnson, J. Koplik, and R. Dashen, “Theory of dynamic permeability and tortuosity in fluid-saturated porous-media,” *J. Fluid Mech.* **176**, 379–402 (1987).  
<sup>22</sup>A. N. Norris, “Low-frequency dispersion and attenuation in partially saturated rocks,” *J. Acoust. Soc. Am.* **94**, 359–370 (1993).  
<sup>23</sup>A. I. Best and C. McCann, “Seismic attenuation and pore-fluid viscosity in clay-rich reservoir sandstones,” *Geophysics* **60**, 1386–1397 (1995).  
<sup>24</sup>K. Aki and P. G. Richards, *Quantitative Seismology: Theory and Methods* (Freeman, New York, 1980), Vols. I and II.  
<sup>25</sup>William F. Murphy III, “Acoustic measures of partial gas saturation in tight sandstones,” *J. Geophys. Res.* **89**, 11549–11559 (1984).  
<sup>26</sup>M. A. Biot and D. G. Willis, “The elastic coefficients of the theory of consolidation,” *J. Appl. Mech.* **24**, 594–601 (1957).  
<sup>27</sup>J. G. Berryman, V. Y. Grechka, and P. A. Berge, “Analysis of Thomsen parameters for finely layered VTI media,” *Geophys. Prospect.* **47**, 959–978 (1999).  
<sup>28</sup>Z. Hashin and S. Shtrikman, “A variational approach to the theory of elastic behavior of polycrystals,” *J. Mech. Phys. Solids* **10**, 343–352 (1962).  
<sup>29</sup>P. A. Berge, B. P. Bonner, and J. G. Berryman, “Ultrasonic velocity-porosity relationships for sandstone analogs made from fused glass beads,” *Geophysics* **60**, 108–119 (1995).  
<sup>30</sup>A. Nur and G. Simmons, “The effect of saturation on velocity in low porosity rocks,” *Earth Planet. Sci. Lett.* **7**, 183–193 (1969).  
<sup>31</sup>T. Bourbié, O. Coussy, and B. Zinszner, *Acoustics of Porous Media* (Gulf, Houston, TX, 1987), p. 56.  
<sup>32</sup>R. Hill, “The elastic behavior of crystalline aggregate,” *Proc. Phys. Soc., London, Sect. A* **65**, 349–354 (1952).  
<sup>33</sup>B. P. Bonner, D. J. Hart, P. A. Berge, and C. M. Aracne, “Influence of chemistry on physical properties: Ultrasonic velocities in mixtures of sand and swelling clay (abstract),” LLNL report UCRL-JC-128306abs, *Eos, Transactions of the American Geophysical Union* **78**, Fall Meeting Supplement, F679 (1997).  
<sup>34</sup>T. Cadoret, “Effet de la saturation eau/gaz sur les propriétés acoustiques

des roches. Étude aux fréquences sonores et ultrasonores," Ph.D. dissertation, Université de Paris VII, Paris, France, 1993.

<sup>35</sup>T. Cadoret, D. Marion, and B. Zinszner, "Influence of frequency and fluid distribution on elastic wave velocities in partially saturated limestones," J.

Geophys. Res. **100**, 9789–9803 (1995).

<sup>36</sup>R. J. S. Brown and J. Korrington, "On the dependence of the elastic properties of a porous rock on the compressibility of the pore fluid," Geophysics **40**, 608–616 (1975).

# Analysis of multiscale scattering and poroelastic attenuation in a real sedimentary rock sequence

Chris L. Hackert and Jorge O. Parra

*Division of Instrumentation and Space Research, Southwest Research Institute, P. O. Drawer 28510, San Antonio, Texas 78228*

(Received 20 July 1999; accepted for publication 2 March 2000)

Compressional waves in heterogeneous permeable media experience attenuation from both scattering and induced pore scale flow of the viscous saturating fluid. For a real, finely sampled sedimentary sequence consisting of 255 layers and covering 30 meters of depth, elastic and poroelastic computer models are applied to investigate the relative importance of scattering and fluid-flow attenuation. The computer models incorporate the known porosity, permeability, and elastic properties of the sand/shale sequence in a binary medium, plane layered structure. The modeled elastic scattering attenuation is well described by stochastic medium theory if two-length scale statistics are applied to reflect the relative thickness of the shale layers when compared to the sand layers. Under the poroelastic Biot/squirt flow model, fluid-flow attenuation from the moderate permeability ( $10^{-14}$  m<sup>2</sup>) sands may be separated in the frequency domain from the attenuation due to the low permeability ( $5 \times 10^{-17}$  m<sup>2</sup>) shale layers. Based on these models, the overall attenuation is well approximated by the sum of the scattering attenuation from stochastic medium theory and the volume weighted average of the attenuations of the sequence member rocks. These results suggest that a high permeability network of sediments or fractures in a lower permeability host rock may have a distinct separable attenuation signature, even if the overall volume of high permeability material is low. Depending on the viscosity of the saturating fluid, the magnitude of the flow-based attenuation can dominate or be dominated by the scattering attenuation at typical sonic logging frequencies ( $\sim 10$  kHz). © 2000 Acoustical Society of America. [S0001-4966(00)03506-2]

PACS numbers: 43.20.Jr, 43.20.Gp, 91.30.Fn [DEC]

## INTRODUCTION

Recent advances in understanding the influence of pore-scale flow on seismic wave propagation (Kazi-Azoual, 1988; Parra, 1991; Dvorkin and Nur, 1993; Parra, 1997) have led to the possibility of using measurements of seismic wave attenuation and velocity dispersion to estimate matrix and fluid properties. In many practical cases of interest, however, the material is inhomogeneous and variations exist in the local permeability, density, and compressibility. These variations, resulting in impedance heterogeneity, introduce a scattering-based dispersion and attenuation for propagating waves which exist in addition to the fluid-based and anelastic dispersion and attenuation. Better understanding of the dispersion mechanisms operating between seismic and ultrasonic frequencies will also lead to better scaling from core measurements to *in situ* seismic data. This can be critical in time lapse seismic applications in which oil reservoirs are monitored for small changes in velocity, e.g., Wang *et al.* (1998).

The theory of poroelasticity is developed from recognition that a compressional wave propagating through a porous medium forces the pore spaces to alternately expand and contract. This local change in pore volume and pore pressure may cause the saturating fluid to flow from contracting pores to expanding pores in a manner such that the fluid is not at rest relative to the solid. The internal motion of a viscous fluid thus extracts energy from the passing wave and also affects the rigidity of the porous volume as a whole, which in turn alters the wave velocity. Biot (1956, 1962) presented a theory describing the frequency dependent behavior of attenuation and dispersion for fluid motion parallel to the di-

rection of wave propagation. In contrast, the squirt flow mechanism accounts for fluid motion perpendicular to the direction of wave propagation. Bourbie *et al.* (1987) suggested that the squirt mechanism is a dominant contributor to experimentally observed attenuation in sedimentary rocks.

One theory describing the squirt mechanism was constructed by Dvorkin and Nur (1993), who presented a unified analysis of the Biot and squirt flow (BISQ) mechanism. The BISQ formulation relies on a heuristic parameter (the squirt flow length), and may not be applicable to all rock types. Nevertheless, the BISQ mechanism has been demonstrated to be quantitatively consistent with measurements of attenuation in sandstone samples as a function of permeability and saturation (Dvorkin *et al.*, 1994). An important observation is that the critical frequency for Biot flow is inversely proportional to  $K/\mu$ , while the squirt flow critical frequency is directly proportional to  $K/\mu$ . Here,  $\mu$  is the fluid viscosity, and  $K$  the medium permeability. The Biot mechanism characteristic frequency goes to zero in the limit of an infinitely permeable medium, while the squirt flow characteristic frequency goes to infinity in the same limit. For this reason, Biot's mechanism is most often observed in highly permeable media (such as unconsolidated beach sands), while the squirt flow mechanism is more often associated with less permeable media (such as sedimentary rocks).

Akbar *et al.* (1994) performed a detailed analysis on the effects of fluid saturation on velocity and attenuation in various types of pore structure. Considering both macro- and microscopic patchy saturation, they developed models based on pore and conduit compressibility. Akbar *et al.* (1994)

compared the results of their theory to a fairly extensive array of experimental data, including effects of confining pressure, fluid saturation, and frequency of investigation. As part of their development, they presented an equation by which the frequency of an observed flow-based attenuation peak can be related to the squirt flow length scale and permeability of the medium.

Sams *et al.* (1997) used a suite of measurements from the well-characterized Imperial College (UK) borehole test site to examine wave propagation over frequency changes of four orders of magnitude. By combining measurements from vertical seismic profiles, crosswell seismic, sonic well logs, and core samples, these researchers were able to put together an estimate of the frequency dependent attenuation and dispersion at the test site. They found a peak in attenuation in the sonic frequency range, slightly higher than 10 kHz. This result is used with basic squirt flow models to estimate the crack aspect ratio in test site rocks, a parameter potentially related to field permeability.

While fluid motion may dominate the attenuation for wave propagation in porous media over certain frequencies, it is certainly not the only mechanism which reduces wave amplitude. Scattering, in the form of reflections from variations in the medium's impedance, has long been recognized as a significant contributor to attenuation. O'Doherty and Anstey (1971) provided one of the first attempts to quantify statistically the frequency dependence of scattering from bed boundaries in seismic methods. They showed that scattering attenuation should be proportional to the power spectrum of the reflection coefficient series. Later workers gave this result a sound mathematical basis (e.g., Burridge *et al.*, 1988; Shapiro and Zien, 1993), with applications generally assuming a single dominant length scale for medium transition (e.g., Kerner and Harris, 1994). Hermann (1998) suggested that sedimentary rocks might be considered to have a fractal multiscale, with reflecting surfaces at every point and a varying range of dominating scales.

Gurevich *et al.* (1997) provided a study considering the effects of scattering and Biot flow. Instead of considering the local squirt flow, however, they considered the global or interlayer flow. The interlayer flow phenomenon is (in some sense) a globalization of the pore scale squirt flow in which a passing pressure wave squeezes fluid out of a thin porous layer into surrounding, more compliant, material. Because of the relatively large interlayer length scales, this effect is predominantly confined to low frequency waves in moderate to high permeability material, while the Biot effect is generally confined to high frequency waves in high permeability material. The pore scale squirt flow mechanism, in contrast, is considered the most likely cause of observed attenuation peaks in moderate permeability rocks (see Bourbie *et al.*, 1987) at the intermediate frequencies (1–100 kHz). Nevertheless, the work of Gurevich *et al.* (1997) demonstrated that flow-based and scattering-based attenuations are separable in that they can be treated independently at first, and later added to produce the total attenuation.

Shapiro and Müller (1999) demonstrated that heterogeneity in permeability is also important in determining the dispersion and attenuation due to the interlayer flow. As the

permeability in the medium changes, the frequencies most strongly affected by the local flow change as well. The overall attenuation of the medium as a whole is thus controlled by those sections in which the dissipation is maximized. For this reason, the effective permeability of a layered formation for large-scale vertical fluid flow may be different from the effective permeability of a formation for micro-scale Biot and global flow induced by vertically propagating seismic waves. The first type is given generally by  $K_{\text{eff}} = \langle K^{-1} \rangle^{-1}$ , while the second may be approximately given (according to Shapiro and Müller) by  $K_{\text{eff}} = \langle K \rangle$  for normal distributions of permeability. Shapiro and Müller (1999) presented an example in which the difference between the two effective permeabilities is more than two orders of magnitude.

In this paper, we first examine the role of scattering on two distinct length scales in the light of stochastic medium theory and a real world sedimentary sequence. We show that inclusion of multiple length scales in typical stochastic medium wave theory is straightforward if the transitions associated with each length scale are statistically independent. Verification of the theoretical development is accomplished using an elastic plane wave program and a known real sedimentary sequence. Finally, we consider the interaction of scattering and the squirt flow attenuation mechanism in heterogeneous permeable media.

## I. PERTURBATION THEORY

This section outlines the results of a one-dimensional stochastic medium wave propagation theory. We follow the development of Parra *et al.* (1999), but similar results would be obtained using the development of Shapiro *et al.* (1994). Consider a one-dimensional medium with density,  $\rho$ , and elastic constant,  $\alpha$ . If all motion in this medium is constrained to the  $x$ -direction, then the unperturbed wave propagation velocity is given by  $\alpha_0 = \rho_0 V_0^2$ . Let the properties of this medium have a random variation in such a way that

$$\rho(x) = \rho_0 e^{\sigma_R R(x)}, \quad \alpha(x) = \alpha_0 e^{\sigma_A A(x)}, \quad (1)$$

where  $\sigma_R$  and  $\sigma_A$  are small parameters, and  $R(x)$  and  $A(x)$  are stationary Gaussian random fields with zero mean and unit variance. These two fields have a covariance  $\chi$ , and each has a spatial distribution governed by the spectral density  $S(k)$ . The spectral density is the Fourier transform of the autocorrelation function  $E(a) = \langle R(x)R(x+a) \rangle = \langle A(x)A(x+a) \rangle$ , where brackets  $\langle \rangle$  denote expectation value over all  $x$ .

Many sedimentary sequences can be considered to be layered mediums with layer thicknesses given by a Poisson distribution (Velzeboer, 1981). With this type of spatial distribution, the autocorrelation function  $E$  is

$$E(x) = e^{-x/l}, \quad (2)$$

and its Fourier transform, the spectral density, is

$$S(k) = \frac{l}{\pi} \frac{1}{1+k^2 l^2}, \quad (3)$$

where  $l$  is the layer transition length scale. Note this means that the average layer thickness is  $2l$ .

The inhomogeneous one-dimensional wave equation is



$$\rho(x) \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x} \left( \alpha(x) \frac{\partial u}{\partial x} \right), \quad (4)$$

where  $u$  is displacement. We will treat the small parameters  $\sigma_R$  and  $\sigma_A$  as perturbations to the linear wave equation. Taking an ensemble average to remove the zero mean fluctuations, and transforming to the frequency wave number domain, the phase velocity of the second order solution to the perturbed wave equation is given in Parra *et al.* (1999) as

$$V = V_0 \left[ 1 - \frac{1}{4} (\sigma_R^2 + \sigma_A^2) (C(k_0) + 1) - \frac{1}{2} \sigma_R \sigma_A \chi C(k_0) \right], \quad (5)$$

where

$$C(k_0) = -\pi k (2k_0 l + j) S(2k_0), \quad (6)$$

and  $k_0$  is  $\omega/V_0$ . By separating out the imaginary part of the wave number, the attenuation due to scattering effects is

$$\frac{1}{Q} = 2 |\text{Im}(C(k_0))| \left[ \frac{1}{4} (\sigma_R^2 + \sigma_A^2) + \frac{1}{2} \chi \sigma_R \sigma_A \right]. \quad (7)$$

The frequency dependent quality factor  $Q$  describes wave attenuation in a form proportional to  $\exp(-\omega x/2VQ)$ , where  $V$  is the phase velocity and  $x$  is the propagation distance (Aki and Richards, 1980).

This attenuation is maximized for  $k_0 l = 1/2$ , corresponding to a wavelength equal to  $4\pi$  times the length scale of the random structure or  $2\pi$  times the average layer thickness. This is in good experimental agreement with Rio *et al.* (1996), who found the scattering effect to be centered on a wavelength-to-layer thickness ratio of seven. The expressions also have theoretical justification since the low and high frequency limiting velocities are consistent with the known low and high frequency velocities of layered media (Backus, 1962; Marion *et al.*, 1994).

## II. MULTIPLE LENGTH SCALE RESULTS

### A. Incorporation into perturbation theory

Acoustic measurements in rocks have been taken over frequency ranges from a few Hz (in long range seismic experiments) to MHz (in lab scale experiments on core samples). Similarly, length scales which can induce scattering effects can range from sedimentary layers hundreds of meters thick to sub-millimeter pore scale variability. Analysis of well logs by White *et al.* (1990) indicates that it is not uncommon for the autocorrelation function of velocity well logs to be controlled by more than one length scale. In this section, we show that it is not difficult to include the multiple length scale scattering effect in the perturbation theories of Parra *et al.* (1999) or Shapiro *et al.* (1994).

The effect of the length scale occurs only in the spectral density, which is the Fourier transform of the autocorrelation function,  $E$ . To include the effect of multiple length scales, we need only to determine the autocorrelation function of a multiple length scale medium. For simplicity, we choose two length scales. For ease of evaluation, we will assume that the variation on each length scale is independent of the other.

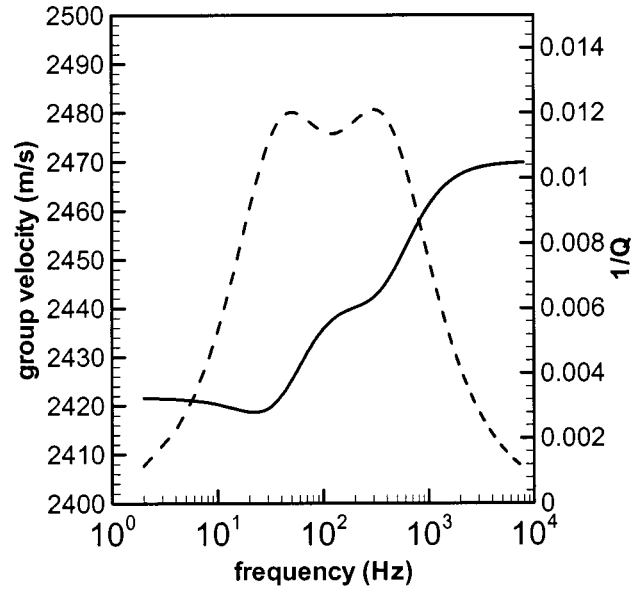


FIG. 1. Theoretical scattering-based group velocity dispersion (solid line) and attenuation (dashed line) for a random medium having variability on two length scales separated by one order of magnitude (50 cm and 5 m). The assumed variability is 20% in each of the density and compressional modulus.

More complicated forms of the autocorrelation function may be included in the perturbation theory, but will often require numerical instead of analytical integration.

Let  $R_1(x)$  and  $R_2(x)$  be independent, zero mean, unit variance, normally distributed random fields with exponential autocorrelation functions  $E$  of different length scales. If  $\sigma_1 R_1(x) + \sigma_2 R_2(x)$  describes the fluctuations in a property like density, then the combined fluctuation has zero mean and variance  $\sigma_R^2 = \sigma_1^2 + \sigma_2^2$  (Kalbfleisch, 1985). If  $R_1(x)$  and  $R_2(x)$  are independent and uncorrelated, then the autocorrelation function of the combined fluctuation is the sum of the individual autocorrelation functions, or

$$\sigma_R^2 E_R(x) = \sigma_1^2 E_1(x) + \sigma_2^2 E_2(x) = \sigma_1^2 e^{-x/l_1} + \sigma_2^2 e^{-x/l_2}. \quad (8)$$

The variability and autocorrelation function used in the perturbation theory will thus be  $\sigma_R^2$  and  $E_R$ , respectively. Similar expressions may be found for the variations in stiffness.

This function (8) is a linear combination of functions of the form Eq. (2), and as such, the resulting spectral density  $S(k)$  is a linear combination of functions of the form Eq. (3). This spectral density may be used with the theory of Shapiro *et al.* (1994) or Eqs. (5)–(7) above to produce the scattering effect of two length scales in the phase velocity and attenuation. From Fig. 1, it may be seen that the effects of each length scale may be observed separately in the attenuation if the lengths are at least an order of magnitude apart. Given that here only linear transforms are made of  $E$ , it is apparent that the effects of scattering from different length scales are additive, as long as the variations at each length scale are uncorrelated with the other length scale.

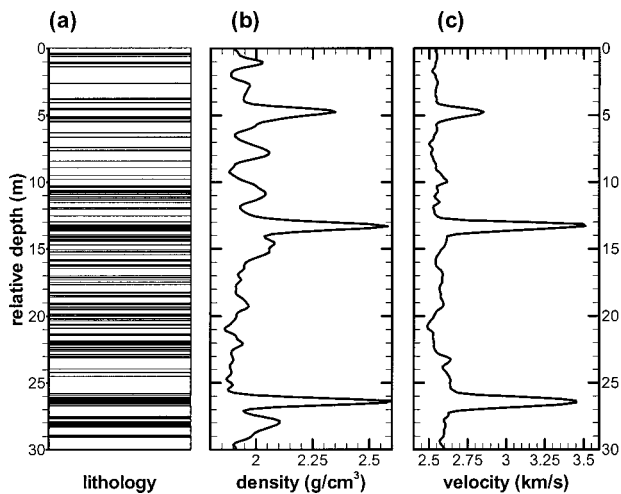


FIG. 2. (a) Actual shale/sand/carbonate sequence for an oil reservoir near Bakersfield, California. Black areas represent sands and carbonates while white areas represent shales. This 30-m interval contains 255 separate layers, based on visual inspection of cores recovered from the wellbore. The interval contains 11% sands and carbonates by volume. Density (b) and velocity (c) logs for the same interval as (a) are also presented. These logs do not represent the true structure of the medium because they have insufficient resolving power.

### B. Computational investigation of a real sequence

The Buena Vista Hills field near Bakersfield, California, includes some oil producing rock formations with very fine structure. The Antelope shale in this field is a high porosity ( $\sim 30\%$ ) but low permeability (50 microdarcy, or  $5 \times 10^{-17} \text{ m}^2$ ) rock. Intercalated with the shale are hundreds of thin (1 mm to 1 m) layers of sand and fractured carbonates. These thin beds have much higher permeability ( $\sim 10$  millidarcy, or  $10^{-14} \text{ m}^2$ ) than the host shale, and are believed to significantly contribute to the oil production from the formation. For this reason, it is important to be able to identify and characterize this fine structure, which mostly falls below the resolution of typical sonic and density logging tools. Given the high permeability contrast between the shales and sands, it may be possible to use dispersion and attenuation measurements to identify the thin bed density.

Figure 2 shows the lithology of a typical section of the Antelope formation, along with the corresponding velocity and density logs. There are 255 alternating shale, sand, and carbonate layers in this 30-m-thick region. The lithology has been determined here to approximately 3 mm resolution by visual inspection of rock cores removed during the drilling of the well. The logging tools typically used to obtain the velocity and density information in the well sample over a coarser interval and essentially average the rock properties over a vertical distance of 30 to 60 cm. Standard well logging, therefore, does not capture much information about these thin beds, and core removal and analysis are often prohibitively expensive and time consuming. The lithology of this section is used to construct a plane layered poroelastic model which is then used as input to a plane wave propagation program (Parra *et al.*, 1997). This program optionally includes the effects of the BISQ mechanism, and so the influence of the variable permeability between layers can be taken into account.

TABLE I. Properties of Buena Vista Hills sediments for the layered model. The volume weighted average is based on the actual lithology prevalence of 11% sand by volume. The squirt flow length is estimated as 20 times the pore throat radius (based on mercury porosimetry measurements) for each lithology. The uniaxial modulus is  $\alpha = \rho V_p^2$ .

Property	Shale	Sand	Weighted Avg.	Std. deviation
Uniaxial modulus (GPa)	11.3	16.5	11.9	14%
Density ( $\text{g/cm}^3$ )	1.90	2.20	1.93	5%
Permeability ( $\times 10^{-15} \text{ m}^2$ )	0.05	10		
Porosity (%)	27	27		
Squirt flow length (mm)	0.01	1		
Solid grain bulk modulus (GPa)	26.8	37.9		

While the two carbonate layers (at relative depths of 13 and 26 m) provide the most striking contrast in the density and velocity logs of Fig. 2, this is somewhat misleading because they actually comprise only about 1% of the formation volume. For simplicity, the carbonates will be replaced by sands in the layered model. (It is not significantly more difficult to create a three-medium model, but because of the small volume of carbonates there is little advantage gained.) Since the actual velocities and densities are not resolved by the well logs, the shale and sands are assigned constant properties given in Table I. The autocorrelation function  $E$  of the layered model is used to provide input to the random medium theory and is shown in Fig. 3 along with two possible fits. Because the shale layers are generally significantly thicker than the sand layers, the autocorrelation function has the double exponential (two length scales) form discussed previously. The best least squares fit using a single exponential function [Eq. (2)] is found to have a correlation length of 4.5 cm. A far better fit is found using a double length scale function [Eq. (8)] with correlation lengths of 1.2 cm (66% weight) and 18 cm (34% weight). Here, the smaller length scale corresponds to the sand layer half-thickness, while the

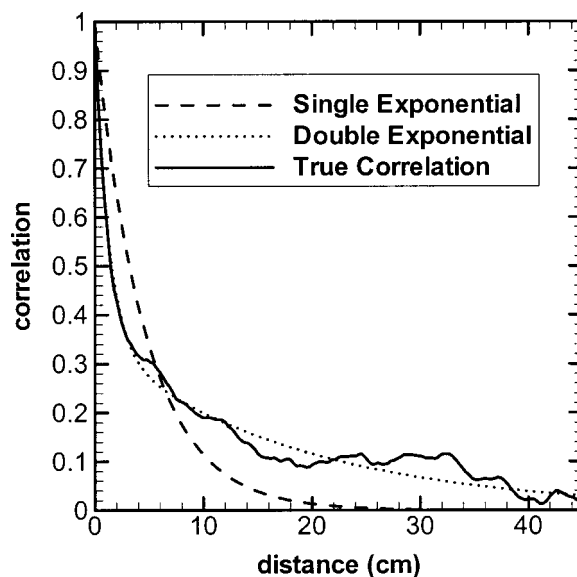


FIG. 3. Autocorrelation function of the lithological sequence in the study interval. The solid line is the actual autocorrelation function. The dashed line is a single length scale fit, while the dotted line is a fit using two length scales.

larger length scale corresponds to the shale layer half-thickness. Interestingly, these two length scales are more than an order of magnitude apart, implying that there will be some visible multiscale scattering effects.

### 1. Elastic behavior

This prediction of visible multiscale scattering is confirmed by applying an elastic plane wave computer program (Parra *et al.*, 1997) to the 255-layer lithologic sequence shown in Fig. 2. This program, operating in the frequency domain, applies a normally incident plane wave to a stack of planar sand/shale layers with properties given in Table I. This program handles the piecewise continuous media by enforcing continuity of displacement and stress across each interface. Within each layer, up- and down-going waves are allowed to propagate with the usual uniform medium plane wave equations. For a given frequency, then, the relative phases and amplitudes within each layer are the solution of a large matrix equation. Real rocks, however, are naturally attenuating to some degree, regardless of scattering or fluid saturation. In the absence of additional information, researchers typically assume or infer a frequency independent  $Q$  value for the rock (Aki and Richards, 1980). This introduces what we will define as intrinsic attenuation to separate it from scattering attenuation. Intrinsic attenuation is a property of the rock and its saturating fluid, while scattering attenuation is a result of interfaces between rock types. The elastic plane wave program can incorporate fixed  $Q$  attenuation by assigning complex moduli for each layer. The imaginary part of the resulting wave numbers acts to exponentially reduce the wave amplitude.

By including no effects of intrinsic attenuation in the layered model, the frequency dependent scattering attenuation may be simply determined by measuring the transmission coefficient on the far side of the layered model. The observed attenuation may be transformed into a  $Q$  value by taking into account the distance traversed and the frequency of measurement. The resulting scattering attenuation is shown in Fig. 4, along with the scattering attenuation prediction of the stochastic medium model. Agreement is quite good, although by necessity the actual ‘realization’ of the random medium causes fluctuations around the predicted mean value. Further numerical experiments show that the magnitude of the fluctuation can be reduced by extending the propagation length in the random medium, increasing the number of layers traversed. This follows the ergodic hypothesis that as the size of a particular random sample increases, the resulting behavior will regress to the ensemble average properties. For this sequence, the frequencies of peak scattering attenuation (1 to 12 kHz) overlap the frequencies used by borehole sonic logging tools (2 to 20 kHz).

It is of practical interest to note that even though there are more than 250 reflecting surfaces in the 30-m study section, the actual attenuation due to scattering is quite low. Peak attenuations correspond to  $Q$  values on the order of 200–400. This is partly because the sands comprise only 11% of the study section, and the actual statistical variance of the medium properties is far lower than the contrast be-

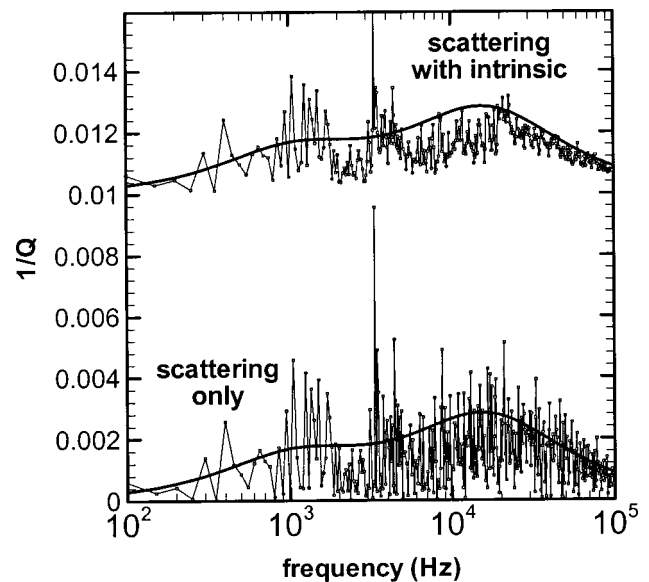


FIG. 4. Attenuation of compressional waves in the study interval. Lower set of curves is scattering attenuation only. Upper set includes an arbitrary intrinsic attenuation of  $Q=100$  in both sand and shale. Heavy line denotes prediction of stochastic medium theory. Light line with symbols is computed with elastic/anelastic plane wave model using the observed layer sequence.

tween sand and shale might at first indicate. Scattering, however, is unlikely to be the only operating attenuating mechanism in a sedimentary sequence.

Adding an arbitrary degree of intrinsic attenuation to the medium in the form of a frequency independent  $Q$  of 100 for all layers provides a background level of attenuation ( $1/Q=0.01$ ). (This background attenuation could represent the imperfect elasticity of consolidated sediments.) As demonstrated by Gurevich *et al.* (1997), the scattering attenuation adds to the intrinsic background, providing up to a 30% increase, and possibly creating a measurable fluctuation. The stochastic medium theory slightly overpredicts the attenuation at high frequencies. This may be due to the high skew (89% shale) and kurtosis (binary medium assumption) in the lithology statistics.

### 2. Poroelastic behavior

Finally, we examine the effects of including the poroelastic BISQ mechanism into the 255-layer model. The poroelastic plane wave program (Parra *et al.*, 1997) operates in a similar manner to the elastic plane wave program, but instead uses the BISQ equations to compute the frequency-wave-number relations for fast and slow  $P$  waves in each layer. The elastic boundary conditions are extended to require continuity of normal fluid displacement across the layer interfaces, which has recently been demonstrated to be rigorously correct for poroelastic media by Gurevich and Schoenberg (1999). The saturating fluid is first assumed to be oil, with viscosity  $\mu=1$  g/cm/sec. In Fig. 5, the dashed and dotted lines show the attenuation due to BISQ flow in unbounded sands and shales, respectively, assuming the medium properties from Table I. The peak attenuation in the sand is at about 2 kHz, covering roughly the same range of

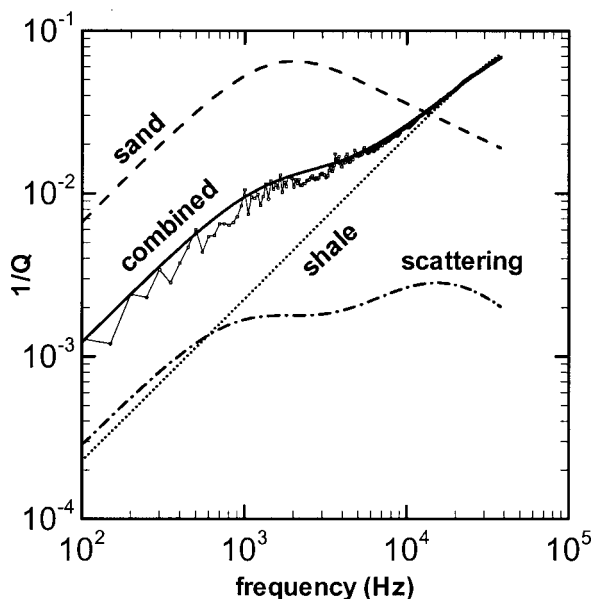


FIG. 5. Attenuation of compressional waves including flow effects for 100% oil saturation. Dashed line: BISQ attenuation in an unbounded sand body having properties as in Table I. Dotted line: BISQ attenuation in an unbounded shale body having properties as in Table I. Dash-dot line: scattering attenuation from stochastic medium theory. Heavy solid line: averaged fluid attenuation plus scattering attenuation for study interval. Light solid line with symbols: attenuation from poroelastic plane wave program.

frequencies as the scattering peak attenuation. The magnitude of the BISQ attenuation is significantly higher than the scattering only attenuation examined earlier, showing that fluid-based attenuation has the potential to entirely dominate scattering attenuation. The BISQ attenuation in the shale exists at higher frequencies due to the low permeability and short squirt flow length. The scattering-only attenuation is reproduced in Fig. 5 as a dot-dash line, for comparison of magnitudes. The actual attenuation, given by the poroelastic plane wave program, is shown by the light line with symbols.

The formations of the Buena Vista Hills reservoir are in fact not completely saturated with oil. The saturations determined from core measurements are in the general range of 80% water and 20% oil. This presents a physical situation which is difficult to model, since water is a wetting fluid and oil is a nonwetting fluid. In practice, this means that the water probably occupies the pore space adjacent to the solid phase, while the oil exists as dispersed droplets at the centers of the larger pores. While the mechanical response of this uneven two-fluid saturation is unknown, we can examine the effects of complete water saturation, and these are shown in Fig. 6. While oil and water have roughly similar densities and compressibilities, the viscosity of water is much less: only 0.01 g/cm/sec. This reduced viscosity moves the squirt flow effects to significantly higher frequencies. The shale body is now essentially nonattenuating over this frequency range, while the sand body has its attenuation peak at about 200 kHz. This leaves the scattering attenuation as the dominant attenuation mechanism in the sonic frequency range of 1–20 kHz, at least in the absence of other attenuating effects.

For both the oil-saturated and the water-saturated cases,

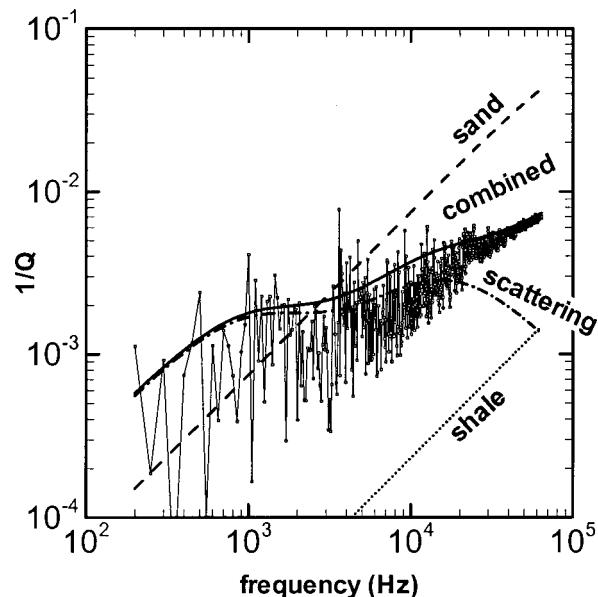


FIG. 6. Attenuation of compressional waves including flow effects for 100% water saturation. Dashed line: BISQ attenuation in an unbounded sand body having properties as in Table I. Dotted line: BISQ attenuation in an unbounded shale body having properties as in Table I. Dash-dot line: scattering attenuation from stochastic medium theory. Heavy solid line: averaged fluid attenuation plus scattering attenuation for study interval. Light solid line with symbols: attenuation from poroelastic plane wave program.

the overall attenuation is well approximated by the sum of three terms: the scattering attenuation, 11% of the sand body attenuation, and 89% of the shale body attenuation. In other words (assuming the validity of the BISQ model), the attenuation response of a poroelastic medium over this frequency range is approximately given by the volume weighted average of the individual (unbounded) lithology responses, plus the scattering attenuation,

$$Q^{-1}(\omega) = Q_{\text{scat}}^{-1}(\omega) + \sum_i v_i Q_i^{-1}(\omega). \quad (9)$$

Here,  $Q^{-1}$  is the observed attenuation,  $Q_{\text{scat}}^{-1}$  is the scattering attenuation,  $v_i$  is the volume fraction of lithology component  $i$ , and  $Q_i^{-1}$  is the intrinsic attenuation of component  $i$ . The result of this formula is plotted as the heavy solid line in Figs. 5 and 6. This is entirely compatible with the observations of Shapiro and Müller (1999) using the Biot mechanism. Shapiro and Müller (1999) in fact suggested a similar equation for attenuation due to layers with a normal distribution of permeability. Physically, then, the local-flow attenuation for waves propagating normal to the layering in a heterogeneous formation is dependent on the local pore-space properties of individual formation members and the quantity of each rock type within the formation.

### III. CONCLUSIONS

A real sedimentary lithological sequence has been shown to have more than one dominant length scale. An elastic plane wave model is used to demonstrate the ability of stochastic medium theory to successfully address scattering on two length scales. While the scattering attenuation for the particular sequence examined is fairly low, it is important to recall that the scattering attenuation is proportional to the

variance of the medium modulus and density. This variance could be increased dramatically by examining a sequence with a more even split between sand and shale (i.e., 50% sand, 50% shale instead of 11% sand, 89% shale), or by examining a sequence with greater contrast between components (i.e., shale and limestone). In any case, the acoustic scattering for the Buena Vista Hills formation has its strongest effect in the frequencies typically employed by sonic logging tools. Sequences with thicker laminations will have scattering effects at lower frequencies.

Localized intrinsic attenuation, such as the BISQ mechanism, adds directly to the scattering attenuation. The very different rock types present in the Buena Vista Hills reservoir lead to quite different BISQ mechanism behavior and spread the effects of pore flow attenuation over a large range of frequencies (two orders of magnitude). The characteristic frequencies of BISQ flow attenuation are controlled by the local properties of each formation member, and not by the average properties of the formation as a whole. Thus the influence of the dissimilar sand and shale layers can be observed separately in the frequency dependent attenuation profile of the test section. This will complicate efforts to determine formation permeability by attenuation measurements as they may not be a well defined peak attenuation in highly heterogeneous media. This is of particular import in fractured media, where a network of discrete, localized, high permeability fractures may comprise a significant fraction of the medium permeability. The viscosity of the saturating fluid is also shown to have a dramatic effect on the BISQ attenuation, and more work is needed to understand the flow response of saturation with two immiscible fluids to compressional waves.

## ACKNOWLEDGMENTS

This work was performed with support from the U.S. Department of Energy (DOE), under Contract No. DE-AC26-99BC15203. The assistance of Mr. Purna Halder (DOE) and permission granted by DOE to publish this work is gratefully acknowledged. Lithological information on the Buena Vista Hills field was obtained with help from Hughbert Collier of Collier Consulting, and Mike Morea and Tom Zalan of Chevron. Comments from Jack Dvorkin and Boris Gurevich helped to improve this manuscript.

Akbar, N., Mavko, G., Nur, A., and Dvorkin, J. (1994). "Seismic signatures of reservoir transport properties and pore fluid distribution," *Geophysics* **59**, 1222–1236.

Aki, K., and Richards, P. G. (1980). *Quantitative Seismology: Theory and Methods* (W. H. Freeman, New York), Vol. I, pp. 177–180.

Backus, G. E. (1962). "Long wave anisotropy produced by horizontal layering," *J. Geophys. Res.* **66**, 4427–4440.

Biot, M. A. (1956). "Theory of propagation of elastic waves in fluid-saturated porous solid: I. Low-frequency range, II. Higher frequency range," *J. Acoust. Soc. Am.* **28**, 168–191.

Biot, M. A. (1962). "Generalized theory of acoustic propagation in porous dissipative media," *J. Acoust. Soc. Am.* **34**, 1254–1264.

Bourbie, T., Coussy, O., and Zinszner, B. (1987). *Acoustics of Porous Media* (Editions Technip, Paris), pp. 220–228.

Burrige, R., Papanicolaou, G. S., and White, B. S. (1988). "One-dimensional wave propagation in a highly discontinuous medium," *Wave Motion* **10**, 19–44.

Dvorkin, J., and Nur, A. (1993). "Dynamic poroelasticity: A unified model with the squirt and Biot mechanisms," *Geophysics* **58**, 524–533.

Dvorkin, J., Nolen-Hoeksema, R., and Nur, A. (1994). "The squirt-flow mechanism: Macroscopic description," *Geophysics* **59**, 428–438.

Gurevich, B., and Schoenberg, M. (1999). "Interface conditions for Biot's equations of poroelasticity," *J. Acoust. Soc. Am.* **105**, 2585–2589.

Gurevich, B., Zyryanov, V. B., and Lopatnikov, S. L. (1997). "Seismic attenuation in finely layered porous rocks: Effects of fluid flow and scattering," *Geophysics* **62**, 319–324.

Hermann, F. J. (1998). "Evidence of scaling for acoustic waves in multi-scale media and its possible implications," 68th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, paper ST9.2.

Kalbfleisch, J. G. (1985). *Probability and Statistical Inference*, 2nd ed. (Springer-Verlag, New York), Vol. 1, p. 184.

Kazi-Azoual, M. (1988). "Green's function in an infinite transversely isotropic saturated poroelastic medium," *J. Acoust. Soc. Am.* **84**, 1883–1889.

Kerner, C., and Harris, P. (1994). "Scattering attenuation in sediments modeled by ARMA processes—Validation of simple  $Q$  models," *Geophysics* **59**, 1813–1826.

Marion, D., Mukerji, T., and Mavko, G. (1994). "Scale effects on velocity dispersion: From ray to effective medium theories in stratified media," *Geophysics* **59**, 374–378.

O'Doherty, R. F., and Anstey, N. A. (1971). "Reflections on amplitudes," *Geophys. Prospect.* **19**, 430–458.

Parra, J. O. (1991). "Analysis of elastic wave propagation in stratified fluid-filled porous media for interwell seismic applications," *J. Acoust. Soc. Am.* **90**, 2557–2575.

Parra, J. O. (1997). "The transversely isotropic wave equation including the Biot and the squirt flow mechanisms: Theory and application," *Geophysics* **62**, 309–318.

Parra, J. O., Collier, H. A., Datta-Gupta, A., Owen, T., Peddibhotla, S., Xu, P.-C., and Zook, B. (1997). "Characterization of fractured reservoirs using static and dynamic data: From sonic and 3D seismic to permeability distribution," Report DOE/DC91008 (DE96001284), National Petroleum Technology Office, Department of Energy, Tulsa, Oklahoma.

Parra, J. O., Hackert, C. L., Ababou, R., and Sablik, M. J. (1999). "Dispersion and attenuation of acoustic waves in randomly heterogeneous media," *J. Appl. Geophys.* **42**, 99–115.

Rio, P., Mukerji, T., Mavko, G., and Marion, D. (1996). "Velocity dispersion and upscaling in a laboratory simulated VSP," *Geophysics* **61**, 584–593.

Sams, M. S., Neep, J. P., Worthington, M. H., and King, M. S. (1997). "The measurement of velocity dispersion and frequency-dependent intrinsic attenuation in sedimentary rocks," *Geophysics* **62**, 1456–1464.

Shapiro, S. A., and Zien, H. (1993). "The O'Doherty-Anstey formula and localization of seismic waves," *Geophysics* **58**, 736–740.

Shapiro, S. A., Zien, H., and Hubral, P. (1994). "A generalized O'Doherty-Anstey formula for waves in finely layered media," *Geophysics* **59**, 1750–1762.

Shapiro, S. A., and Müller, T. M. (1999). "Seismic signatures of permeability in heterogeneous porous media," *Geophysics* **64**, 99–103.

Wang, Z., Cates, M. E., and Langan, R. T. (1998). "Seismic monitoring of a CO<sub>2</sub> flood in a carbonate reservoir: A rock physics study," *Geophysics* **63**, 1604–1617.

White, B., Sheng, P., and Nair, B. (1990). "Localization and backscattering spectrum of seismic waves in stratified lithology," *Geophysics* **55**, 1158–1165.

Velzeboer, C. J. (1981). "The theoretical seismic reflection response of sedimentary sequences," *Geophysics* **46**, 843–853.

# Model equation for strongly focused finite-amplitude sound beams

Tomoo Kamakura<sup>a)</sup>

*The University of Electro-Communications, 1-5-1, Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan*

Tsuneo Ishiwata

*Oki Electric Industry Co., Ltd., 4-10-3, Shibaura, Minato-ku, Tokyo 108-8551, Japan*

Kazuhisa Matsuda

*Koganei Technical High School, 6-8-9, Honmachi, Koganei-shi, Tokyo 184-8581, Japan*

(Received 1 March 1999; revised 6 March 2000; accepted 7 March 2000)

A model equation that describes the propagation of sound beams in a fluid is developed using the oblate spheroidal coordinate system. This spheroidal beam equation (SBE) is a parabolic equation and has a specific application to a theoretical prediction on focused, high-frequency beams from a circular aperture. The aperture angle does not have to be small. The theoretical background is basically along the same analytical lines as the composite method (CM) reported previously [B. Ystad and J. Berntsen, *Acustica* **82**, 698–706 (1996)]. Numerical examples are displayed for the amplitudes of sound pressure along and across the beam axis when sinusoidal waves are radiated from the source with uniform amplitude distribution. The primitive approach to linear field analysis is readily extended to the case where harmonic generation in finite-amplitude sound beams becomes significant due to the inherent nonlinearity of the medium. The theory provides the propagation and beam pattern profiles that differ from the CM solution for each harmonic component. © 2000 Acoustical Society of America. [S0001-4966(00)03706-1]

PACS numbers: 43.25.Cb [MAB]

## INTRODUCTION

The purpose of this article is to present a new model equation that describes well the nonlinear propagation of ultrasound beams from a concave focusing source with a circular aperture. Since sound energy concentration is sufficiently high at the focus even if the source pressure level is low, waveform distortion is inevitably generated due to the inherent nonlinearity of the medium. In particular, the nonlinearity depends directly on wave amplitude, so harmonic generation in the beam becomes active and significant near the focus.

As regards theoretical analyses of nonlinearly focusing beams, quite a number of research reports have been published so far. A brief review of the various theories is found in the report of Naze Tjøtta and Tjøtta.<sup>1</sup> As far as the authors know, almost all the works begin with the Khokhlov–Zabolotskaya–Kuznetsov (KZK) equation<sup>2</sup> that is extensively used as a model equation for describing the combined effects of diffraction, absorption, and nonlinearity in directional beams. The KZK equation is derived under the paraxial approximation. Therefore, its applicability to focused beams is limited to small opening angles of aperture. In fact, Tjøtta *et al.* predict that the upper limit of the applicability is roughly  $16^\circ$  for the half-aperture angle  $\alpha_0$ .<sup>3</sup>

Several theories differing from the paraxial approach have been presented and treat the nonlinear behavior of beams in wide-angle propagation. Assuming weak nonlinearity, Tsuchiya and Kagawa<sup>4</sup> numerically solve the loss-less

Kuznetsov's wave equation<sup>5</sup> by a finite-element method, and obtain the fundamental and second harmonic components along and across the beam axis for  $\alpha_0 \leq 70^\circ$ . Yano gives numerical solutions of governing equations such as the continuity equation for a perfect gas, taking account of sound energy losses at the shock front.<sup>6</sup> Nonlinearity does not necessarily have to be weak. For sinusoidal excitation at the source, Yano depicts the spatial evolution of the waveform and the distribution of acoustic streaming on the axis for  $\alpha_0 = 30^\circ$ . Unfortunately, both theoretical approaches are limited to cases in which the focal length is equal to several wavelengths, so the focusing effect is relatively weak, although their theories are substantially applicable to focused beams from a high-frequency, strongly curved source. Christopher and Parker provide a new algorithm of computer simulation for nonlinear and diffractive beams from unfocused and focused source.<sup>7</sup> The main concept of their theory is to use the spatial domain point-spread function to remove the effect of the parabolic approximation, and to introduce an operator-splitting method to individually calculate the terms of diffraction, absorption, and nonlinearity. The source model adopted to a focused field is a geometrical focal length of 10 cm and a radius of 1 cm, yielding a small aperture angle of  $5.7^\circ$ . Later, Christopher applies their computational model to an analysis of shock waves produced by an extracorporeal lithotripter.<sup>8</sup> Using the operator-splitting method, Tavakkoli *et al.* develop a time-domain numerical model that differs from existing theories in that their model principally uses full diffraction formulation.<sup>9</sup> The numerical solutions for the time waveforms of pulses emitted from a spherical, highly focused transducer with a focal length of

<sup>a)</sup>Electronic mail: kamakura@ee.uec.ac.jp

190 mm and an aperture radius of 86 mm is in good agreement with experiments. Unfortunately, their comparison is restricted to the range from the source to the focus. Apart from computer simulation, Ostrovsky and Sutin introduce a step-by-step analytical technique to assess nonlinearity and diffraction contribution to a waveform evolution with propagation.<sup>10</sup> Their concise expressions are mostly focused on the description of maximum values of pressure and intensity.

Utilizing easy extension of the parabolic approximation to nonlinear field analysis, Ystad and Berntsen<sup>11</sup> propose a composite method (CM) that describes the field in terms of two beam equations. One of them is the mixed model equation (MME)<sup>1</sup> and is accurate close to the source; the other is the KZK equation and is accurate near the focus. Both equations have to be appropriately connected along a curve in an azimuth plane to uniquely determine the solution. They also give some numerical examples for on-axis profiles and beam patterns of the first few harmonic pressures. The source parameters assigned are  $kd=421$  and  $\alpha_0=26.75^\circ$ , where  $k$  is the wave number and  $d$  is the focal length.

The present theory is, in principle, along the same analytical lines as the CM. A crucially different point is the use of an oblate spheroidal coordinate system,<sup>12</sup> which is like rectangular coordinates near the focus and like spherical coordinates near the source, to successfully account for beam focusing and spreading.<sup>13</sup> Since the effect of diffraction is the most important factor in accurately describing beam propagation, the formulation of our model equation is first pursued in a linear field. Next, numerical computation is performed to solve the equation via a finite difference scheme. Axial pressure curves and beam patterns are then compared with other theories to validate the model equation. In the final stage of discussion, the theory is extended to nonlinear propagation of strongly focused beams. Harmonic generation in the beams is quantitatively evaluated based on the same source conditions as those of Ystad and Berntsen.<sup>11</sup> Some interesting features and comments on the model equation are then presented.

## I. LINEAR FIELD

Figure 1 shows the oblate spheroidal coordinate system,<sup>12</sup> in which the variables  $(\sigma, \eta, \varphi)$  are related to the rectangular coordinates  $(x, y, z)$  by

$$\begin{aligned} x &= b\sqrt{(1+\sigma^2)(1-\eta^2)}\cos\varphi, \\ y &= b\sqrt{(1+\sigma^2)(1-\eta^2)}\sin\varphi, \\ z &= b\sigma\eta, \end{aligned} \quad (1)$$

with

$$-\infty < \sigma < \infty, \quad 0 \leq \eta \leq 1, \quad 0 \leq \varphi < 2\pi, \quad (2)$$

where  $2b$  is the interfocal length. In terms of  $\sigma$ ,  $\eta$ , and  $\varphi$ , the scalar wave equation

$$\nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} = 0, \quad (3)$$

is transformed into

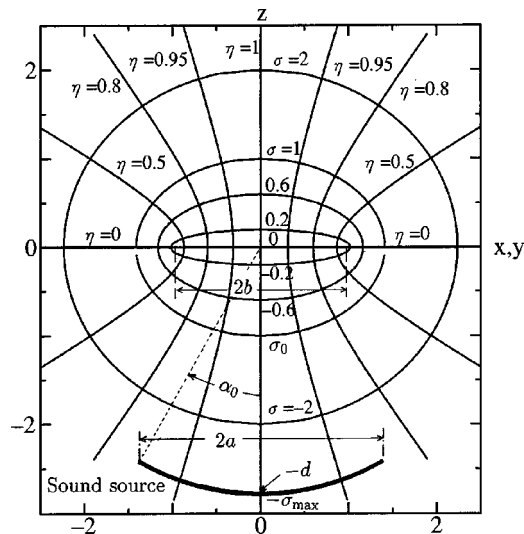


FIG. 1. Oblate spheroidal coordinate system.

$$\begin{aligned} \frac{\partial}{\partial \sigma} \left[ (1 + \sigma^2) \frac{\partial p}{\partial \sigma} \right] + \frac{\partial}{\partial \eta} \left[ (1 - \eta^2) \frac{\partial p}{\partial \eta} \right] \\ - \frac{b^2(\sigma^2 + \eta^2)}{c_0^2} \frac{\partial^2 p}{\partial t^2} = 0, \end{aligned} \quad (4)$$

where  $p$  is the sound pressure,  $c_0$  is the sound speed,  $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ , and  $t$  is time. In Eq. (4), the sound field is assumed to be axisymmetric, so the equation does not contain the variable  $\varphi$ . For simplicity, it is also assumed that there is no sound absorption.

Figure 2 shows the configuration of a concave sound source with a circular aperture. The notation  $d$  is the geometrical focal length,  $a$  is the aperture radius, and  $\alpha_0$  is the half-aperture angle. The aperture angle is related to  $a$  and  $d$  as follows:

$$\sin \alpha_0 = \frac{a}{d}. \quad (5)$$

We hypothetically divide the focused field into two regions,  $\sigma < \sigma_0 < 0$  and  $\sigma_0 \leq \sigma$ , where  $\sigma = \sigma_0 (= \text{const} < 0)$  denotes a specific transition location. The former region is close to the source where spherically converging waves are predominant and the spherical coordinates are preferable to the rectangular ones for field analysis; the latter region is near the focus where the paraxial or planar parabolic ap-

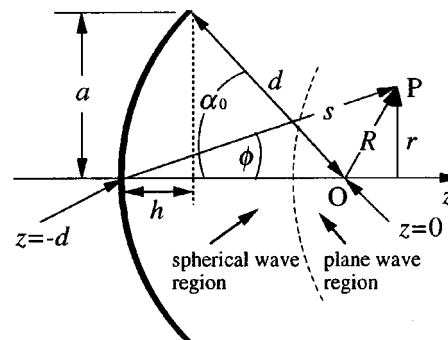


FIG. 2. Focusing source geometry.

proximation is useful. For each region, different retarded time is introduced to observe progressive waves in a frame moving with  $c_0$ .

### A. Spherical wave expansion ( $\sigma < \sigma_0 < 0$ )

Near the source the spherically retarded time,

$$t_s = t + \frac{R}{c_0}, \quad (6)$$

is chosen, where  $R$  is the distance between an observation point  $P$  and the origin  $O$ , and is given by

$$R = \sqrt{x^2 + y^2 + z^2} = b\sqrt{\sigma^2 + 1 - \eta^2}. \quad (7)$$

Substituting Eqs. (6) and (7) into Eq. (4), we have

$$\begin{aligned} & \frac{\partial}{\partial \sigma} \left[ (1 + \sigma^2) \frac{\partial p}{\partial \sigma} \right] + \frac{\partial}{\partial \eta} \left[ (1 - \eta^2) \frac{\partial p}{\partial \eta} \right] \\ & + \frac{2b}{c_0 \sqrt{\sigma^2 + 1 - \eta^2}} \left[ \sigma(1 + \sigma^2) \frac{\partial^2 p}{\partial t_s \partial \sigma} \right. \\ & \left. - \eta(1 - \eta^2) \frac{\partial^2 p}{\partial t_s \partial \eta} + (\sigma^2 + \eta^2) \frac{\partial p}{\partial t_s} \right] = 0. \end{aligned} \quad (8)$$

The first term on the left-hand side can be discarded when the wave travels a distance of many wavelengths without pronounced changes in its shape. In actual focusing systems, the focal length is much larger than the wavelength  $\lambda$ ; i.e., the condition  $d \gg \lambda$  or  $kd \gg 1$ , where  $k = 2\pi/\lambda$  is the wave number, is always satisfied. We thus reduce Eq. (8) to a parabolic beam equation

$$\begin{aligned} & \frac{\partial^2 \bar{p}}{\partial \tau_s \partial \sigma} + \frac{1}{2} \frac{\sin 2\theta}{\sigma(1 + \sigma^2)} \frac{\partial^2 \bar{p}}{\partial \tau_s \partial \theta} + \frac{\varepsilon \sqrt{\sigma^2 + \sin^2 \theta}}{\sigma(1 + \sigma^2)} \\ & \times \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right) + \frac{\sigma^2 + \cos^2 \theta}{\sigma(1 + \sigma^2)} \frac{\partial \bar{p}}{\partial \tau_s} = 0, \end{aligned} \quad (9)$$

where the new variables

$$\bar{p} = \frac{p}{p_0}, \quad \tau_s = \omega t_s, \quad (10)$$

are employed to rewrite the equation in a dimensionless form.  $\varepsilon = 1/2kb$  in Eq. (9) is generally a small quantity compared with unity. Also,  $p_0$  is the amplitude of sound pressure on the source face, and  $\omega$  is the angular frequency of ultrasound. The angle variable  $\theta$  is related to  $\eta$  as in Eq. (11),

$$\eta = \cos \theta. \quad (11)$$

The degenerate surface  $\theta = 0$  in the coordinate system is the beam axis, and the surface  $\theta = \pi/2 (= 90^\circ)$  is the focal plane.

Now imagine that the observation point approaches the focal plane. When  $\sigma$  tends toward zero, the second to fourth terms in Eq. (9) increase in magnitude and particularly have singularity at the focus. This means that the reduced beam equation of Eq. (9) is inaccurate close to the focal plane. The same problem occurs in the MME and the spherical coordinate version of the nonlinear progressive wave equation.<sup>11,14</sup> In order to accurately describe the field around the focus,

another model equation that is expanded in the planar wave may be expected to be more amenable to analysis owing to the validity of the paraxial approximation.

### B. Planar wave expansion ( $\sigma \geq \sigma_0, \sigma_0 < 0$ )

In place of  $t_s$ , the retarded time pertinent to a planar wave is introduced,

$$t_p = t - \frac{z}{c_0} = t - \frac{b\sigma\eta}{c_0}. \quad (12)$$

Substituting Eq. (12) into Eq. (4) and neglecting the term  $(\partial/\partial\sigma)(1 + \sigma^2)(\partial p/\partial\sigma)$  for the same consideration as in the derivation of Eq. (9) or due to the slowness of waveform change with respect to  $\sigma$ , we obtain<sup>13</sup>

$$\begin{aligned} & \cos \theta \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} - \frac{\sigma}{1 + \sigma^2} \sin \theta \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \theta} \\ & - \frac{\varepsilon}{1 + \sigma^2} \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right) = 0, \end{aligned} \quad (13)$$

where all variables except for  $\tau_p (= \omega t_p)$  have the same meanings as those in Eq. (9). In a region near the axis, Eq. (13) is approximated as

$$\frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} \approx \frac{\varepsilon}{1 + \sigma^2} \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right), \quad (14)$$

because  $\cos \theta \approx 1$  and  $\sin \theta \approx 0$ . From rewriting the first term on the left-hand side of Eq. (13) as

$$\begin{aligned} \cos \theta \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} &= \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} - (1 - \cos \theta) \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} \\ &\approx \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} - \frac{\varepsilon(1 - \cos \theta)}{1 + \sigma^2} \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right), \end{aligned} \quad (15)$$

an alternative version for the planar beam equation follows:

$$\begin{aligned} & \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} - \frac{\sigma}{1 + \sigma^2} \sin \theta \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \theta} - \frac{\varepsilon(2 - \cos \theta)}{1 + \sigma^2} \\ & \times \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right) = 0. \end{aligned} \quad (16)$$

We thus have derived two beam equations that are convenient for field evaluation in regions near the source and near the focus. Specifically, Eq. (9) is accurate close to the source, and Eq. (16) is suitable for analysis in the paraxial and focal region. These equations have to be successfully matched on a plane  $\sigma = \sigma_0 (< 0)$  to utilize their effectiveness. Since Eqs. (9) and (16) are expanded in the same coordinates, it is easy to match them. Subtracting Eq. (12) from Eq. (6), we obtain

$$\tau_s = \tau_p + kb(\sqrt{\sigma_0^2 + \sin^2 \theta} + \sigma_0 \cos \theta). \quad (17)$$

When the observer stays on the axis,  $\tau_s$  becomes equal to  $\tau_p$  because  $\theta = 0$  and  $\sigma_0 < 0$ . It might be expected that the equations are smoothly connected to each other by a simple expression of Eq. (17). Hereafter, we refer to the set of beam



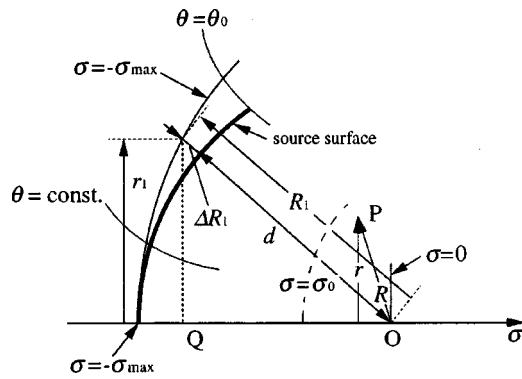


FIG. 3.  $\sigma$ - $\theta$  coordinates in the focusing source.

equations given by Eqs. (9) and (16) with their connecting relation of Eq. (17) as the spheroidal beam equation (SBE).

### C. Boundary conditions at the source

Figure 3 illustrates the source geometry again in the  $\sigma$ - $\theta$  coordinates. The center of the source is located at  $\sigma = -\sigma_{\max} (< 0)$ , where  $\sigma_{\max}$  is defined as the ratio of  $d$  to  $b$ , i.e.,  $\sigma_{\max} = d/b$ , and is larger than unity. The value of  $\sigma_{\max}$  may be somewhat arbitrarily determined independent of the aperture angle and source frequency. However, when the aperture is opened widely and the beam is focused strongly,  $\sigma_{\max}$  has to take on a large value.

Even if  $\sigma_{\max}$  is large, the surface  $\sigma = -\sigma_{\max}$  is not coincident with the actual source surface of  $R = d$ . The distance between the two surfaces is

$$\Delta R_1 = R_1 - d. \quad (18)$$

In order to realize a focusing system with focal length  $d$ , monochromatic sounds have to be radiated from the surface  $\sigma = -\sigma_{\max}$  by leading the phase of  $\omega t$  by  $k\Delta R_1$ . With the aid of Eq. (6), the resultant phase at the surface  $\sigma = -\sigma_{\max}$  becomes

$$\psi = \omega t + k\Delta R_1 = \omega t_s - kR_1 + k\Delta R_1 = \tau_s - kd, \quad (19)$$

in which  $kd$  is constant and may be neglected in actual calculations.

We note that  $r_1$  and  $R_1$  are

$$r_1 = b\sqrt{1 + \sigma_{\max}^2} \sin \theta, \quad R_1 = b\sqrt{\sigma_{\max}^2 + \sin^2 \theta}. \quad (20)$$

The distance  $\overline{OQ}$  is then given by

$$\overline{OQ} = \sqrt{R_1^2 - r_1^2} = b\sigma_{\max} \cos \theta = d \cos \theta, \quad (21)$$

and the following expression relates the half-aperture angle  $\alpha_0$  to the spheroidal angle  $\theta$ ,

$$\tan \alpha_0 = \frac{r_1}{\overline{OQ}} = \sqrt{1 + \frac{1}{\sigma_{\max}^2}} \tan \theta. \quad (22)$$

From Eq. (19), the boundary conditions in the SBE system yield

$$\bar{p}|_{\sigma = -\sigma_{\max}} = \begin{cases} F(\theta) \sin \tau_s, & \theta \leq \theta_0 \\ 0 & \theta > \theta_0, \end{cases} \quad (23)$$

TABLE I. Numerical parameters ( $d = 10$  cm).

$\alpha_0$ [degs]	$a$ [mm]	$\sigma_{\max}$	$\theta_0$ [degs]
10	17.4	2	8.96
20	34.2	2	18
30	50	10/3	28.9
40	64.2	5	39.4

where  $F(\theta)$  is a pressure distribution function on the source. When the pressure is uniformly distributed,  $F(\theta) = 1$ .  $\theta_0$  in Eq. (23) is the angle  $\theta$  that satisfies Eq. (22) and is written as

$$\theta_0 = \tan^{-1} \left( \frac{\tan \alpha_0}{\sqrt{1 + 1/\sigma_{\max}^2}} \right). \quad (24)$$

For example, when the half-aperture angle  $\alpha_0$  is  $30^\circ$  and  $\sigma_{\max}$  is chosen to be  $10/3$ ,  $\theta_0$  becomes  $28.9^\circ$ , a little smaller than  $\alpha_0$ .

## II. NUMERICAL EXAMPLES

We assign the source parameters as follows for comparison with Ystad and Berntsen's numerical examples of using the CM,<sup>11</sup>

frequency = 1 MHz, focal length = 10 cm,  $F(\theta) = 1$ ,

sound speed = 1492 m/s (in water, temperature =  $23^\circ\text{C}$ ).

Our model equation or the SBE is subject to the condition of  $kd \gg 1$  just like the CM. In fact, the present source parameters provide  $kd = 421$ , which is sufficiently larger than unity. Table I presents various aperture angles and their corresponding source radii. The value of  $\sigma_{\max}$  for each aperture angle is not decisive, but representative. For wider aperture angles, however,  $\sigma_{\max}$  usually takes a larger value, as mentioned above.

Although the Rayleigh integral is valid only for a sound field from a planar source, the expression is widely applicable to the beams from a curved radiating source under O'Neil's hypothesis<sup>15</sup> (e.g., the curvature of the source is relatively weak and the frequency is sufficiently high). According to Coulouvrat's theoretical comparison with his exact solution for focusing beams from a uniformly vibrating source shaped as a spherical cap and mounted on an infinite rigid baffle, O'Neil's solution induced from the Rayleigh integral has a good applicability for much wider angles of aperture than generally thought.<sup>16</sup> Actually, favorable agreement between the solutions is found for  $\alpha_0 \leq 60^\circ$  and within a range of  $ka = 10$  to  $50$ . To test the validity of the SBE, therefore, we can safely resort to the following O'Neil solutions for comparison:

$$\frac{p}{\rho_0 c_0 v_0} = j \frac{2 \sin\{k(B - z_1)/2\}}{1 - z_1/d} e^{-jkM} \quad (\text{on the axis}), \quad (25)$$

$$\frac{p}{\rho_0 c_0 v_0} = j \frac{ke^{-jks}}{2\pi R} HF(ka \sin \phi) \quad (\text{near the focal plane}), \quad (26)$$

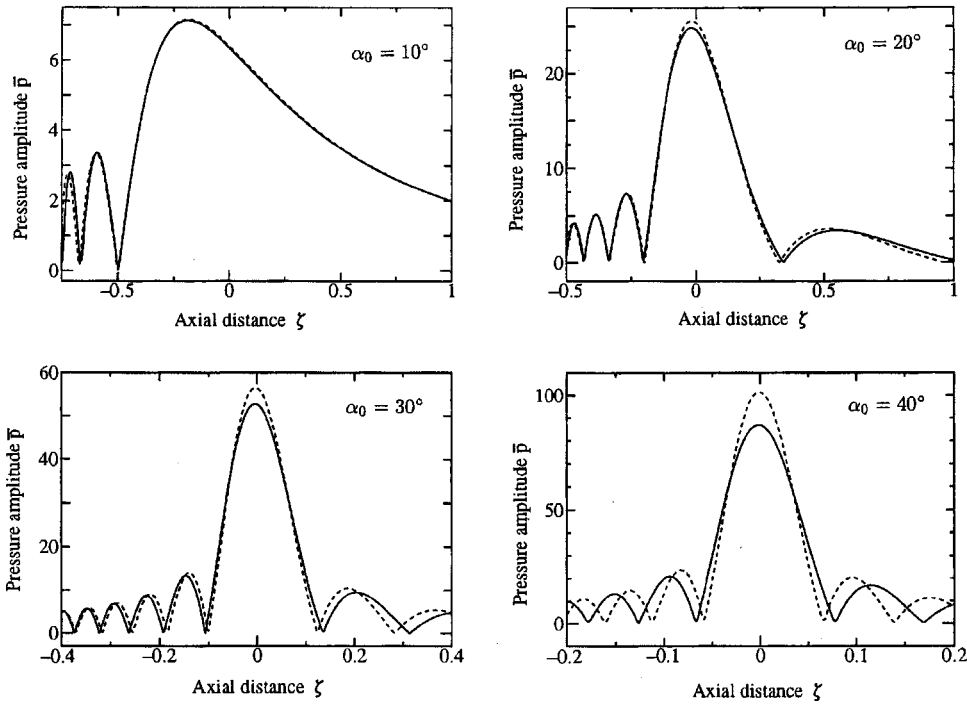


FIG. 4. Comparison of on-axis sound pressure amplitudes by the PE solution (solid curves) and by O'Neil's solution (dotted curves) for various aperture angles.

where  $M = (B + z_1)/2$ ,  $B = \sqrt{(z_1 - h)^2 + a^2}$ ,  $H = 2\pi dh$ ,  $z_1 = z + d$ ,  $\rho_0$  is the medium density,  $v_0$  is the normal velocity amplitude on the source, and

$$F(ka \sin \phi) = \frac{2}{ka \sin \phi} \sum_{n=0}^{\infty} (-1)^n \left(\frac{h}{a}\right)^{2n} \times J_{2n+1}(ka \sin \phi). \quad (27)$$

Other notations ( $R$ ,  $s$ ,  $\phi$ , and  $h$ ) are indicated in Fig. 2. Moreover,  $F(\cdot)$  is a dimensionless function which signifies the so-called source directivity.  $J_{2n+1}(\cdot)$  in the expression is a Bessel function.

In particular, when focusing is still weak, Eqs. (25) and (26) are simplified further. Since  $a \ll d$ , the depth of the aperture surface  $h (= d - \sqrt{d^2 - a^2})$  is approximated by  $h \approx a^2/2d \ll d$ . Taking account of  $a^2 + h^2 = 2dh$ , we derive  $B$  approximately as

$$B = \sqrt{(z_1 - h)^2 + a^2} = \sqrt{z_1^2 - 2h(z_1 - d)} \approx z_1 \left\{ 1 - \frac{h(z_1 - d)}{z_1^2} \right\} \quad (\text{away from the source}). \quad (28)$$

Equations (25) and (26) are then written as

$$\frac{p}{\rho_0 c_0 v_0} = j \frac{G}{z_1/d} \frac{\sin x}{x} e^{-jx - jkz_1} \quad (\text{on the axis}), \quad (29)$$

$$\frac{p}{\rho_0 c_0 v_0} = jG \frac{2}{ka \sin \phi} J_1(ka \sin \phi) e^{-jkd} \quad (\text{in the focal plane}), \quad (30)$$

where  $x = (G/2)(d/z_1 - 1)$ , and  $G = ka^2/2d$ , the focusing gain, is defined as the ratio of the pressure at the focus to the pressure on the source.

For such focusing sources, another straightforward theory has been derived by applying the paraxial approximation to the wave equation. The resultant beam model is called the parabolic equation (PE) and is equivalent to the linearized form of the KZK equation without the dissipation term. Equations (29) and (30) are the same forms as those derived from the PE associated with an appropriate boundary condition.<sup>17</sup>

Sound pressure amplitudes on the beam axis are calculated using O'Neil's solution of Eq. (25) and the PE solution of Eq. (29); their results are shown in Fig. 4 for aperture angles from 10 to 40°. The abscissa is a dimensionless distance in terms of the axial coordinate  $z$  and the focal length  $d = 10$  cm;  $\zeta = z/d$ .  $\zeta = -1$  and 0 correspond to the center of the aperture and the focus, respectively. The ordinate is the amplitude of the dimensionless pressure  $\bar{p} = p/p_0$ .

Incidentally, Tjøtta *et al.*<sup>3</sup> give roughly 16° as an upper limit of the half-aperture angle for the applicability of the PE with equivalent boundary conditions. As can be seen in Fig. 4, their prediction is reasonable since an angle of at most 20° seems to be the upper limit. When  $\alpha_0 = 30^\circ$ , O'Neil's solution attains a peak of 57 near the focus. The value based on the PE solution is, however, 53 and is about 7% less than that by O'Neil's prediction. As the aperture widens, the difference between the peaks increases.

Next, we are concerned with the problem of the validity of the SBE proposed. Figure 5 shows propagation data of the sound pressure amplitude calculated by the SBE and O'Neil's model. The steady-state solution of the SBE is given by a combination of sine and cosine functions,

$$\bar{p} = \begin{cases} g^{(s)} \sin \tau_s + h^{(s)} \cos \tau_s & (\sigma < \sigma_0 < 0) \\ g^{(p)} \sin \tau_p + h^{(p)} \cos \tau_p & (\sigma \geq \sigma_0, \sigma_0 < 0), \end{cases} \quad (31)$$

where  $g^{(s)}$ ,  $h^{(s)}$ ,  $g^{(p)}$ , and  $h^{(p)}$  are amplitude components and are functions of the spatial variables  $\sigma$  and  $\theta$ . Substitut-

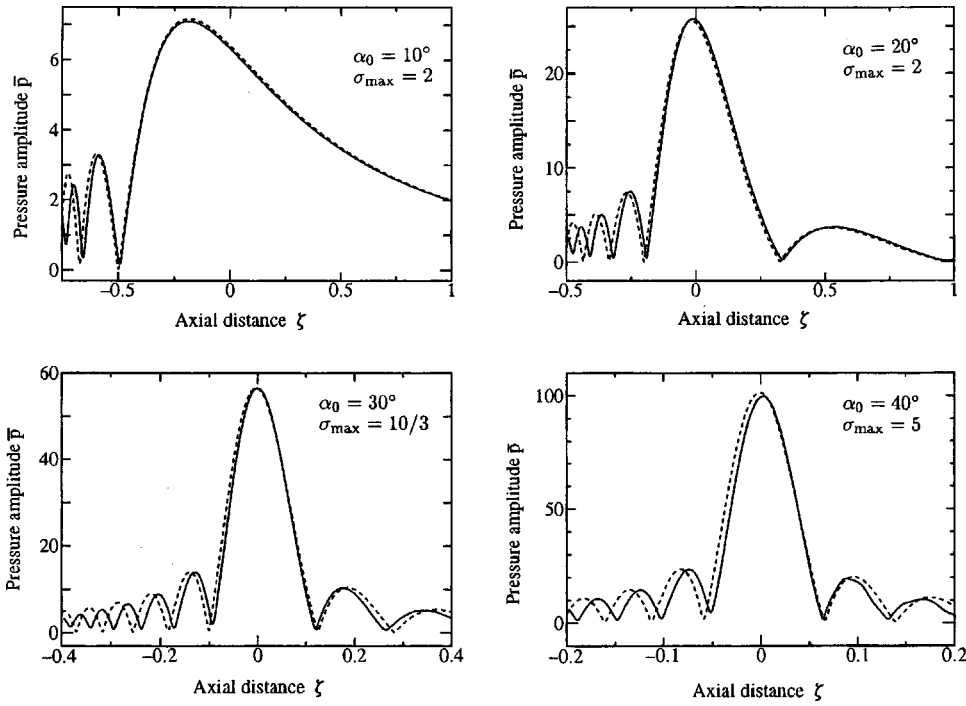


FIG. 5. Comparison of on-axis sound pressure amplitudes by the SBE solution (solid curves) and by O'Neil's solution (dotted curves) for various aperture angles.

ing the two expressions of Eq. (31) into their respective beam equations [Eqs. (9) and (16)], we obtain the partial differential equations in terms of the amplitude components. Initial values of  $g^{(s)}$  and  $h^{(s)}$  at  $\sigma = -\sigma_{\max}$  are imposed through the boundary condition Eq. (23),

$$g^{(s)} = \begin{cases} 1 & (0 \leq \theta \leq \theta_0) \\ 0 & (\theta_0 < \theta \leq \pi/2) \end{cases}, \quad h^{(s)} = 0 \quad (0 \leq \theta \leq \pi/2). \quad (32)$$

$g^{(p)}$  and  $h^{(p)}$  are, on the other hand, connected with  $g^{(s)}$  and  $h^{(s)}$  at  $\sigma = \sigma_0$  via Eq. (17),

$$\begin{aligned} g^{(p)} &= g^{(s)} \cos \gamma - h^{(s)} \sin \gamma \\ h^{(p)} &= g^{(s)} \sin \gamma + h^{(s)} \cos \gamma \quad (\text{at } \sigma = \sigma_0 < 0), \end{aligned} \quad (33)$$

in which  $\gamma = kb(\sqrt{\sigma_0^2 + \sin^2 \theta} + \sigma_0 \cos \theta)$ . Since it is not easy to solve these equations analytically, numerical computation is performed by means of an implicit backward finite difference method. Pressure amplitudes are obtained by  $\sqrt{g^{(s)^2} + h^{(s)^2}}$  or  $\sqrt{g^{(p)^2} + h^{(p)^2}}$ . The numerical integration region ranges from 0 to  $\pi/2 (= 90^\circ)$  on the  $\theta$  axis and from  $-\sigma_{\max}$  ( $\zeta = -1$ ) to the observer on the  $\sigma$  axis. Generally speaking, step sizes in a finite difference method should be as small as possible to attain reasonable accuracy. Three different grid points (450, 600, and 1000) on the  $\theta$  axis were used to determine the accuracy of the numerical solution in advance. There was, however, no great difference in the three solutions. In addition, about ten times more computation run-time was required for the case of 1000 grid points compared with the case of 450 points. Therefore, the step size is equally spaced by  $90/450 = 0.2^\circ$ . The step size in the axial direction is changed with  $\sigma$  in accordance with the convergence condition in the iteration method to solve a large linear matrix. In both the spherical and planar wave regions, the size varies with  $\sigma$  in proportion to  $1 + \sigma^2$ . For  $\alpha_0 = 30^\circ$ , the initial size is set to  $3.5 \times 10^{-3}$ . The transition

location can be numerically determined to be a good choice for  $\sigma_0 = -1$ . Choosing too large or too small a value tends to decrease the model accuracy of the SBE. (See Appendix A.)

It must always be kept in mind that numerical errors really exist, such as the truncation error in a finite difference representation. Even so, the SBE solution represented by solid curves and O'Neil's solution by dotted curves are overall in fairly good agreement independent of the aperture angle. When the aperture angle exceeds  $20^\circ$ , the difference between the solutions is discernible, but not so noticeable as the corresponding difference between O'Neil's solution and the PE solution (Fig. 4). For  $\alpha_0 = 30^\circ$ , the peak pressures for both theories almost equal 57. It is noteworthy that, for wider aperture angles, the SBE solution shifts the propagation curves slightly toward the focus in the prefocal region and slightly toward the source in the post-focal region. As a rule of thumb, the upper limit of effectiveness of the SBE might be an aperture angle of  $40^\circ$ .

An interesting comparison is shown in Fig. 6 for the SBE and CM solutions. The half-aperture angle  $\alpha_0$  is  $26.75^\circ$ , which corresponds to a source radius  $a = 4.5$  cm. Dotted curves in the figures are O'Neil's solution.

We can easily recognize the following features for the SBE solution.

- (i) The peak pressure moves away from the source. However, the shift is not as remarkable as that of the CM.
- (ii) The peak pressure is almost the same as the pressure based on O'Neil's prediction. The CM predicts a little higher amplitude.
- (iii) The minimum pressures on both sides around the focus become almost zero like O'Neil's model; the CM solution does not indicate such minima.

Figure 7 shows beam patterns in the focal plane for  $\alpha_0 = 26.75^\circ$ . The shapes of the main lobes are very close in

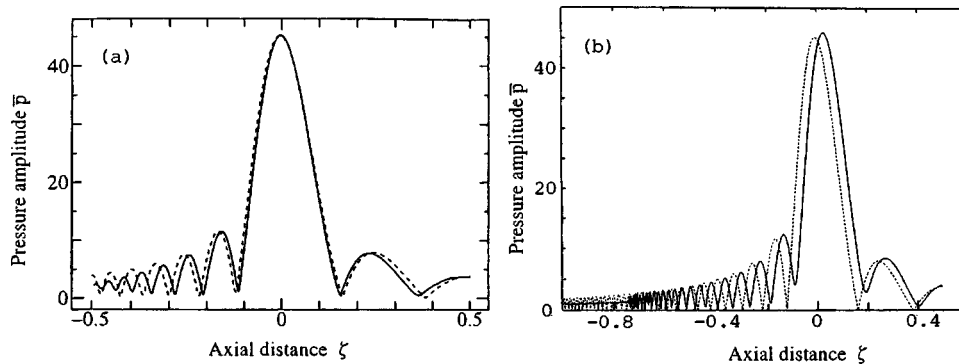


FIG. 6. Comparison of on-axis sound pressure amplitudes by the SBE solution (a) and by the CM solution (b) (Ref. 11). Dotted curves in both figures denote O'Neil's solution. The half-aperture angle is  $26.75^\circ$ .

both solutions. However, the side lobes shift a little to the beam axis compared with O'Neil's pattern. In Fig. 18 of Ystad and Berntsen's article,<sup>11</sup> the main and side lobes are not clearly separated. In contrast with their results, the SBE as well as O'Neil's model exhibit discernible separation among the lobes. Such beam patterns with completely separated lobes are usually observed in the focal plane for a weakly focusing source<sup>18</sup> and in the far-field for an unfocused source.<sup>19</sup>

### III. EXTENSION TO NONLINEAR FIELD

The model equations that we have treated so far exclude the sound absorption and nonlinearity that a real fluid should have. In particular, sound energy concentration in a focusing system is sufficiently high at the focus even if the source pressure level is low. Hence, the nonlinearity in finite-amplitude sound beams has a crucial role to play in generating waveform distortion.

To derive a reliable model equation of a nonlinear beam for high frequency and wide-angle focusing systems, Naze Tjøtta and Tjøtta<sup>1</sup> begin with a nonlinear wave equation that is correct up to the second order in the Mach number and includes thermoviscous dissipation,

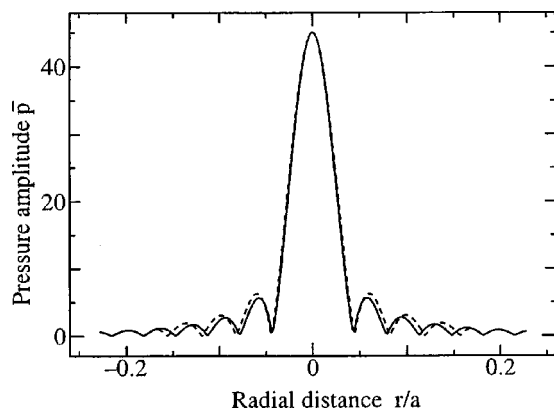


FIG. 7. Beam patterns in the focal plane. Solid curve is obtained by the SBE solution, and dotted curve is O'Neil's solution. The half-aperture angle is  $26.75^\circ$ .

$$\begin{aligned} & \left[ \left( 1 + \frac{\delta}{c_0^2} \frac{\partial}{\partial t} \right) \nabla^2 - \frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} \right] p \\ &= - \frac{\beta}{\rho_0 c_0^4} \frac{\partial^2}{\partial t^2} p^2 - \left( \nabla^2 + \frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} \right) \mathcal{L} \\ &+ \rho_0 \nabla \cdot [\mathbf{v} \times (\nabla \times \mathbf{v})], \end{aligned} \quad (34)$$

where  $\delta$  is the sound diffusivity,  $\beta$  is the coefficient of nonlinearity,  $\mathbf{v}$  is the particle velocity, and  $\mathcal{L}$  is the Lagrangian density defined as  $\mathcal{L} = \rho_0(\mathbf{v} \cdot \mathbf{v})/2 - p^2/2\rho_0 c_0^2$ . In order to derive a simplified model equation for practical purposes, they assume that the fluid particle near the source is vibrating to and fro along the radial line from the origin, and an acoustic impedance similar to that in the far-field of a spherically spreading wave is related to the particle velocity and sound pressure in the MME. Under the above assumption, they conclude that the terms containing the Lagrangian and  $\nabla \cdot [\mathbf{v} \times \nabla \times \mathbf{v}]$  are of order  $1/\sqrt{kd}$  compared to the nonlinear term containing  $\beta$ . Using the following Westervelt equation,<sup>2</sup>

$$\nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} + \frac{\delta}{c_0^4} \frac{\partial^3 p}{\partial t^3} = - \frac{\beta}{\rho_0 c_0^4} \frac{\partial^2 p^2}{\partial t^2}, \quad (35)$$

that is in principle justified when the linear impedance relation of the plane progressive wave holds true, and assuming weak nonlinearity in the beam, Tsuchiya and Kagawa say that a finite element solution of Eq. (34) is very close to an analytical solution of Eq. (35) in water for the second harmonic component [Eqs. (34) and (35) in the absence of losses are used in the article].<sup>4</sup> Moreover, they stress that there is no need to distinguish the two sources of nonlinearities (the thermodynamic and the kinematic nonlinearities) for large values of the nonlinearity parameter  $B/A$ . These statements allow us, in a certain sense, to ignore the second and third nonlinearity terms on the right-hand side of Eq. (34) since they are smaller than the first term. In fact, it can be numerically demonstrated that the second harmonic pressure that originates from the second term is only one percentage of the pressure originating from the full terms for the half-aperture angle of  $30^\circ$  (Appendix B). For higher aperture angles, however, ignoring the nonlinearity terms is questionable because each term is expected to play a comparable role in harmonic generation.<sup>4</sup> For reliability, the authors restrict their numerical simulation to the field emitted from sources with an aperture angle of less than  $30^\circ$ . Thus the relevant nonlinear beam equation is Eq. (35), Westervelt's equation.

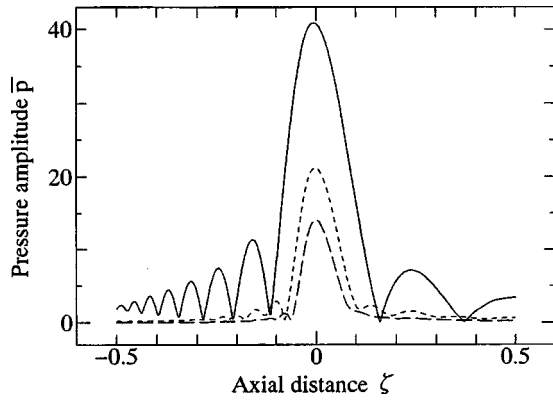


FIG. 8. Harmonic generation in focused beams from a circular concave source with half-aperture angle  $\alpha_0 = 26.75^\circ$ . The source pressure amplitude  $p_0$  is 302 kPa. Solid, dotted, and dashed curves are the axial fundamental, second, and third harmonic pressure amplitudes.

Strictly speaking, the equation obtained by Westervelt is the loss-less form of Eq. (35). The equation should more appropriately be termed the “simple source formation of the nonlinear equation,” as Piquette and Van Buren point out.<sup>20</sup> In the derivation of Eq. (35), it is presumed that the operator  $(\delta/c_0^2)\partial/\partial t$  has a negligibly small effect on the wave propagation.

A similar approximation as in the preceding sections allows us to reduce Eq. (35) to the nonlinear SBE. For example, near the source the first two terms on the left-hand side of Eq. (35) are changed into the expression on the left-hand side of Eq. (9). As a result, we have

$$\begin{aligned} & \frac{\partial^2 \bar{p}}{\partial \tau_s \partial \sigma} + \frac{1}{2} \frac{\sin 2\theta}{\sigma(1+\sigma^2)} \frac{\partial^2 \bar{p}}{\partial \tau_s \partial \theta} + \frac{\varepsilon \sqrt{\sigma^2 + \sin^2 \theta}}{\sigma(1+\sigma^2)} \\ & \times \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right) + \frac{E}{\sigma} \frac{\partial \bar{p}}{\partial \tau_s} \\ & = - \frac{\sqrt{\sigma^2 + \sin^2 \theta}}{\sigma} \left( \alpha b \frac{\partial^3 \bar{p}}{\partial \tau_s^3} + \frac{b}{2l_D} \frac{\partial^2 \bar{p}^2}{\partial \tau_s^2} \right) E \\ & \quad (\sigma < \sigma_0 < 0), \end{aligned} \quad (36)$$

where  $\alpha = \delta\omega^2/2c_0^3$  is the sound attenuation coefficient and  $l_D = \rho_0 c_0^3 / \beta p_0 \omega$  is the shock formation distance for a plane wave.  $E$  in Eq. (36) is a function of  $\sigma$  and  $\theta$ ,

$$E(\sigma, \theta) = \frac{\sigma^2 + \cos^2 \theta}{1 + \sigma^2}. \quad (37)$$

In the same way, another form for a planar wave is obtained from Eq. (16) as follows:

$$\begin{aligned} & \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \sigma} - \frac{\sigma}{1 + \sigma^2} \sin \theta \frac{\partial^2 \bar{p}}{\partial \tau_p \partial \theta} - \frac{\varepsilon(2 - \cos \theta)}{1 + \sigma^2} \left( \frac{\partial^2 \bar{p}}{\partial \theta^2} + \cot \theta \frac{\partial \bar{p}}{\partial \theta} \right) \\ & = \left( \alpha b \frac{\partial^3 \bar{p}}{\partial \tau_p^3} + \frac{b}{2l_D} \frac{\partial^2 \bar{p}^2}{\partial \tau_p^2} \right) E \quad (\sigma \geq \sigma_0, \sigma_0 < 0). \end{aligned} \quad (38)$$

Due to the quadratic nonlinearity of  $\bar{p}^2$ , the wave is distorted during propagation. The distorted waveform is expanded in a Fourier series,

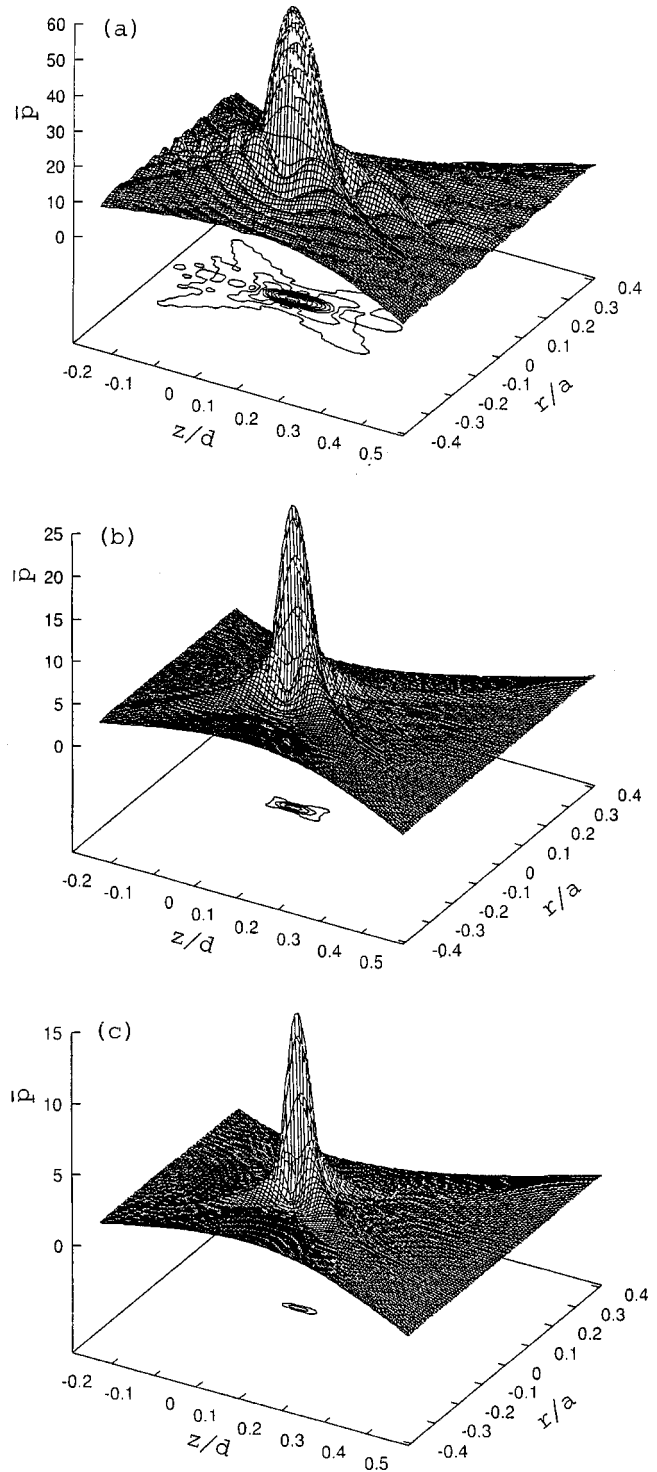


FIG. 9. Three-dimensional diagrams of the fundamental (a), second harmonic (b), and third harmonic (c) sound pressures in focused beams from a circular concave source with half-aperture angle  $\alpha_0 = 30^\circ$ . The source pressure amplitude  $p_0$  is 220 kPa.

$$\bar{p} = \begin{cases} \sum_{n=1}^{\infty} (g_n^{(s)} \sin n\tau_s + h_n^{(s)} \cos n\tau_s) & (\sigma < \sigma_0 < 0) \\ \sum_{n=1}^{\infty} (g_n^{(p)} \sin n\tau_p + h_n^{(p)} \cos n\tau_p) & (\sigma \geq \sigma_0, \sigma_0 < 0), \end{cases} \quad (39)$$

where  $g_n^{(s)}$ ,  $h_n^{(s)}$ ,  $g_n^{(p)}$ , and  $h_n^{(p)}$  are the Fourier coefficients.

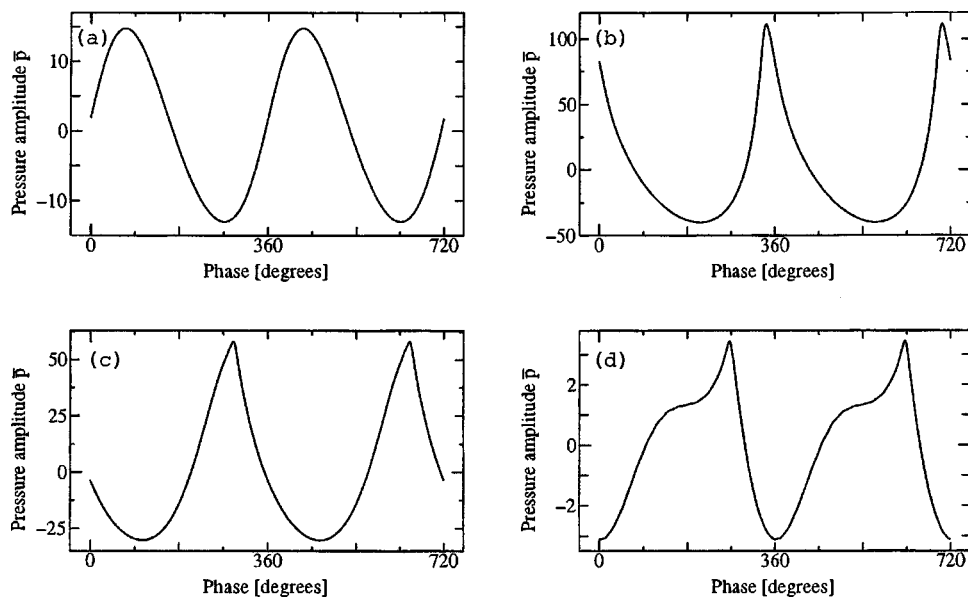


FIG. 10. Time waveform history at  $\zeta = -0.13$  (a), 0 (b), 0.05 (c), and 0.25 (d) on the beam axis. The source conditions are the same as in Fig. 9. The abscissa indicates the phase  $\tau_p$  in degrees.

Substitution of Eq. (39) into Eqs. (36) and (38) results in two sets of coupled and nonlinear partial differential equations for the Fourier coefficients. It seems too intractably involved to solve these equations. However, numerical analysis using the finite difference scheme explained in the previous section makes it possible to obtain the solution of the  $n$ -th harmonic component in a versatile way.

The theoretical source model is the same as in Sec. II: a concave spherical source with a circular aperture of radius  $a$  and with the geometrical focal length  $d = 10$  cm radiates a sinusoidal wave of 1 MHz in water at 23 °C. Pressure on the source face is distributed uniformly. From the literature, the sound absorption coefficient  $\alpha$  at 1 MHz is 0.022 neper/m.<sup>21</sup> The absorption at the  $n$ -th harmonic is thus 0.022  $n^2$  neper/m, in accordance with the frequency-squared law.

Figure 8 shows axial pressures of the first three harmonics for  $\alpha_0 = 26.75^\circ$  and  $p_0 = 302$  kPa. The plane-wave shock formation distance is  $l_D = 50$  cm for  $\rho_0 = 1000$  kg/m<sup>3</sup> and  $\beta = 3.5$ . A sufficient number of harmonics has to be retained in the calculation to ensure negligibly small errors in harmonic components. In particular, the higher the harmonic of interest is, the more harmonics must be retained. Forty harmonics are retained to obtain the propagation data of the first three harmonics. In comparison with the data including 50 harmonics, the maximum difference in the third harmonic amplitude is less than a few percent within our propagation range. These source conditions are the same as those adopted by Ystad and Berntsen.

Figure 6 confirms that the SBE solution provides the propagation curves of linear beams in the shapes and amplitudes differently from the CM simulation. Consequently, a quantitatively different generation of the harmonics is predicted for the two theories. From Fig. 8, the peak pressures for the first to third harmonics are 41, 21, and 14, whereas the corresponding peaks are 43, 20, and 13 from the data in Fig. 21 of Ystad and Berntsen's article,<sup>11</sup> where these peaks are attained at some distance from the focus.

Based on linear theory, the amplitude decay of the fundamental component due to sound absorption is negligibly

small, only  $e^{-0.022 \times 0.1} = 0.998$  (0.02 dB), during propagation from the source to the focus. Nevertheless, we note from Figs. 6 and 8 that the amplitude is decreased from 45 to 41, which is about a 1-dB decay. This extra attenuation is caused by the pumping effect of sound energy from the fundamental to the higher harmonics. Additionally, careful examination of both figures reveals that the peak point of the fundamental moves just a little toward the source. A similar phenomenon is observed theoretically and experimentally for planar sources in a more dramatic way.<sup>22,23</sup>

Three-dimensional diagrams of the fundamental to third harmonic pressures are viewed in Fig. 9 for  $\alpha_0 = 30^\circ$  and  $p_0 = 220$  kPa. Fifty harmonics are retained in the calculation. The bottom map in each figure is an isopressure contour on the  $\sigma$ - $\xi$  plane. The pressure interval of contour lines is  $5 p_0$ . We can ascertain that higher harmonic waves concentrate their energy mostly near the focus.

Figure 10 depicts time waveforms of a finite-amplitude CW along the axis. The source conditions and numerical parameters are the same as those in Fig. 9. The abscissa is scaled in the retarded time  $\tau_p$ . The observation points are located at  $\zeta = -0.13, 0, 0.05,$  and  $0.25$ , which correspond to  $z = -1.3, 0, 0.5,$  and  $2.5$  cm, respectively, in real dimensions. In the prefocal region (a), nonlinearity is weak, and consequent generation of the harmonic distortion is not clearly established in the waveform. At the focus (b), however, the wave distorts its shape asymmetrically. There is sharpening in the positive pressure and the rounding in the negative pressure, and the former peak is about three times larger than the latter peak. Thus nonlinearly and diffractively induced asymmetry of the waveform has been investigated thoroughly by many researchers from theoretical and experimental points of view.<sup>24,25</sup>

Passing through the focus, individual harmonic components shift phases abruptly. It seems that the phase changes result in the reverse waveform in the time scale just behind the focus (c). The wave spreads out in the post-focal region, and the amplitudes of the harmonic components decrease

greatly. Figure 10(d) shows the waveform near the second dip of the fundamental [see Fig. 5]. Parasite shocklike pulses are found in the positive portion of the pressure.

#### IV. CONCLUSION

The authors have proposed a new model equation that is beneficial for the field analysis of ultrasound beams emitted from a concave focusing source with a circular aperture and with a wide-aperture angle. The theory is developed in the oblate spheroidal coordinate system. The reduced beam equation, which is referred to as the spheroidal beam equation (SBE), can be reliably applied to predict the sound pressure profiles along and across the beam axis for half-aperture angles of up to  $40^\circ$  for a linear field. The theory is easily extended to the case where harmonic generation becomes active and significant due to the inherent nonlinearity of the medium. Unfortunately, no analytical solution of the SBE has been found even for a linear field. In nonlinear beams, the solution provides considerable difficulties. However, a straightforward and powerful approach to obtaining a solution of the SBE is feasible through a finite difference scheme. In the present numerical simulation using a VT Alpha 533 computer, it takes no more than 15 min of runtime to obtain all harmonic components even for the heaviest load.

Some numerical examples using the same source parameters as Ystad and Berntsen's article have demonstrated that at least the peak amplitudes of the first three harmonics predicted by the SBE solution are almost the same as those predicted by CM solutions. However, the locations of the peak and the minimum pressures differ between the two solutions. To validate the effectiveness of the present model equation, appropriate experiments and/or numerical comparisons with other theories are needed.

#### ACKNOWLEDGMENTS

This work is in part supported by Grants-in-Aid for Scientific Research, the Ministry of Education, Science, and Sports in Japan (Grant Nos. 08650063 and 10650404).

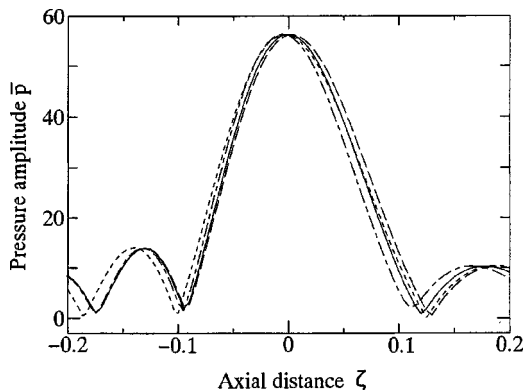


FIG. A1. On-axis sound pressure amplitudes for focused beams from a circular aperture of  $\alpha_0 = 30^\circ$ . The transition locations of spherical to planar wave regions are  $\sigma_0 = -0.8$  (long-dashed line),  $-1.0$  (solid line), and  $-1.2$  (dot-dashed line). The dotted line denotes O'Neil's solution.

#### APPENDIX A

Like the MME solutions,<sup>11</sup> sound pressure profiles along and across the beam axis are dependent on the transition location of  $\sigma_0$ . Figure A1 shows the axial pressure amplitudes when  $\sigma_0 = -0.8, -1,$  and  $-1.2$ . The half-aperture angle is kept at  $30^\circ$ . As can be seen, the pressure peaks are almost independent of the variation of  $\sigma_0$  and are equal to O'Neil's prediction,<sup>57</sup> However, the peak position moves toward the sound source with decreasing values of  $\sigma_0$ . It is worthwhile to note that  $\sigma_0 = -1$  is a reasonable choice in the present numerical computation.

#### APPENDIX B

It is not easy to further simplify each term of the nonlinear sources in Eq. (34) because the planar wave impedance relation between pressure and particle velocity no longer holds in focusing systems with a large aperture angle.<sup>4</sup> Even so, some reasonable assumptions allow us to derive a practical expression for the nonlinear source terms.

Except for a high-viscosity fluid, the sound field is irrotational; i.e.,  $\nabla \times \mathbf{v} = 0$ . Hence the third term on the right-hand side of Eq. (34) is removed from the nonlinear sources. Let the remaining source terms be represented by  $-q$ ,

$$q = \frac{\beta}{\rho_0 c_0^4} \frac{\partial^2}{\partial t^2} p^2 + \left( \nabla^2 + \frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} \right) \mathcal{L}, \quad (\text{B1})$$

or

$$q = q_1 + q_2, \quad q_1 = \frac{\beta}{\rho_0 c_0^4} \frac{\partial^2}{\partial t^2} p^2, \quad q_2 = \left( \nabla^2 + \frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} \right) \mathcal{L}, \quad (\text{B2})$$

where  $\mathcal{L}$  is the Lagrangian density and is given by

$$\mathcal{L} = \frac{\rho_0}{2} \mathbf{v} \cdot \mathbf{v} - \frac{p^2}{2\rho_0 c_0^2}. \quad (\text{B3})$$

For simplicity, cylindrical coordinates  $(z, r)$  are introduced to describe the fields. The radial component of the velocity,  $v$ , is generally smaller than the axial component  $u$  (i.e.,  $u > v$ ). One of the terms obtained by substituting Eq. (B3) into Eq. (B2) becomes approximately

$$\nabla^2(\mathbf{v} \cdot \mathbf{v}) \approx \nabla^2 u^2 = 2 \nabla \cdot (u \nabla u) = 2 \{ u \nabla^2 u + (\nabla u \cdot \nabla u) \}. \quad (\text{B4})$$

Any quadratic nonlinearity term may be replaced by its first-order quantity without destroying the second-order accuracy. Using the linear relations, we then have

$$\nabla^2 u = \frac{1}{c_0^2} \frac{\partial^2 u}{\partial t^2}, \quad \frac{\partial u}{\partial z} = -\frac{1}{\rho_0 c_0^2} \frac{\partial p}{\partial t}, \quad (\text{B5})$$

and we obtain

$$\nabla^2 u^2 = 2 \left\{ \frac{u}{c_0^2} \frac{\partial^2 u}{\partial t^2} + \frac{1}{\rho_0^2 c_0^4} \left( \frac{\partial p}{\partial t} \right)^2 \right\}, \quad (\text{B6})$$

where it is assumed that  $|\partial u / \partial z| > |\partial u / \partial r|$ . In the same way,

$$\nabla^2 p^2 = 2 \{ p \nabla^2 p + (\nabla p \cdot \nabla p) \} = 2 \left\{ \frac{p}{c_0^2} \frac{\partial^2 p}{\partial t^2} + \rho_0^2 \left( \frac{\partial u}{\partial t} \right)^2 \right\}, \quad (\text{B7})$$

where the relation  $\partial p/\partial z = -\rho_0 \partial u/\partial t$  is used. As a result,  $\nabla^2 \mathcal{L}$  in  $q_2$  takes the form

$$\nabla^2 \mathcal{L} = \frac{\rho_0}{c_0^2} u \frac{\partial^2 u}{\partial t^2} + \frac{1}{\rho_0 c_0^4} \left( \frac{\partial p}{\partial t} \right)^2 - \frac{1}{\rho_0 c_0^4} p \frac{\partial^2 p}{\partial t^2} - \frac{\rho_0}{c_0^2} \left( \frac{\partial u}{\partial t} \right)^2. \quad (\text{B8})$$

Applying the same procedure to the differential operator  $\partial^2 \mathcal{L}/\partial t^2$  yields

$$\frac{1}{c_0^2} \frac{\partial^2 \mathcal{L}}{\partial t^2} = \frac{\rho_0}{c_0^2} \left\{ \frac{\partial^2 u}{\partial t^2} u + \left( \frac{\partial u}{\partial t} \right)^2 \right\} - \frac{1}{\rho_0 c_0^4} \left\{ \frac{\partial^2 p}{\partial t^2} p + \left( \frac{\partial p}{\partial t} \right)^2 \right\}. \quad (\text{B9})$$

By adding Eq. (B8) to Eq. (B9), the source term  $q_2$  becomes

$$q_2 = \left( \nabla^2 + \frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} \right) \mathcal{L} = \frac{2\rho_0}{c_0^2} \frac{\partial^2 u}{\partial t^2} u - \frac{2}{\rho_0 c_0^4} p \frac{\partial^2 p}{\partial t^2}. \quad (\text{B10})$$

We now consider the second harmonic pressures in beams with weak nonlinearity to roughly estimate the contribution of the nonlinearities to harmonic generation. First, we put the sound pressure in the form

$$p = p_0 \{ g \sin(\omega t - kz) + h \cos(\omega t - kz) \}, \quad (\text{B11})$$

where  $g$  and  $h$  are both functions containing the spatial variables  $z$  and  $r$ . The corresponding particle velocity becomes

$$u = \frac{p}{\rho_0 c_0} + \frac{p_0}{\rho_0 \omega} \left\{ \frac{\partial g}{\partial z} \cos(\omega t - kz) - \frac{\partial h}{\partial z} \sin(\omega t - kz) \right\}. \quad (\text{B12})$$

The first term of this equation indicates the planar wave impedance relation. The second term, which generally diminishes in the far-field of planar sources, seems to play important roles in harmonic generation in the focusing system. Substituting Eqs. (B11) and (B12) into Eq. (B10) and performing some tedious arrangement, we can obtain

$$q_2 = -\frac{2\omega p_0^2}{\rho_0 c_0^3} \left\{ \left( g \frac{\partial g}{\partial z} - h \frac{\partial h}{\partial z} \right) \sin 2(\omega t - kz) + \frac{\partial(hg)}{\partial z} \cos 2(\omega t - kz) \right\}. \quad (\text{B13})$$

Term  $q_1$  in Eq. (B2) is expressed as

$$q_1 = -\frac{4\beta\omega^2 p_0^2}{\rho_0 c_0^4} \left\{ gh \sin 2(\omega t - kz) + \frac{1}{2} (h^2 - g^2) \cos 2(\omega t - kz) \right\}. \quad (\text{B14})$$

The source magnitudes of  $q_1$  and  $q_2$  are shown in Fig. B1. Both the magnitudes are divided by  $2\omega^2 p_0^2/\rho_0 c_0^4$  to make them nondimensional quantities. No change is made for the source parameters in water ( $d=10$  cm,  $f=1$  MHz, and  $\alpha_0=30^\circ$ ). Sound absorption is not considered here for simplicity. As a whole,  $q_2$  is about two orders of magnitude smaller than  $q_1$  around the focus. However, at the dips, the two sources are comparable in magnitude.

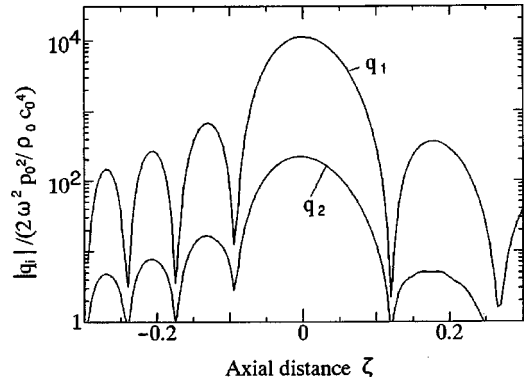


FIG. B1. Amplitude profiles of the nonlinearity source terms  $|q_1|$  and  $|q_2|$  along the beam axis.  $f=1$  MHz,  $d=10$  cm, and  $\alpha_0=30^\circ$ .

On-axis sound pressure levels of the fundamental and second harmonic components are given in Fig. B2, where the source pressure amplitude  $p_0$  is 20 kPa. Other source parameters are the same as in Fig. B1. For the second harmonic, there are two curves. One comes from the total source  $q = q_1 + q_2$ , and the other comes from  $q_2$ . These curves are greatly different in magnitude: curve  $q_2$  is about 40 dB below than curve  $q$ . Therefore, we can reasonably ignore the second term of the nonlinearity sources in Eq. (B1) and can safely use Westervelt's beam equation under our source conditions. It should be noted that the pressure profile of  $q_2$  spreads more around the focus than the profile of  $q$ , and has a rather similar pattern as the primary.

We have derived Eqs. (B4) and (B6) for the conditions that the axial component  $u$  in particle velocity is larger than the radial component  $v$  and the derivative  $\partial u/\partial z$  dominates over  $\partial u/\partial r$  in the whole space, in particular, around the focus. As for this inequality, Eq. (20) in the previous article<sup>26</sup> gives some information about the relationship between  $u$  and  $v$ , although the field is treated as a Gaussian beam and is subject to the paraxial approximation. At the focus, the relations  $u = -(1/\rho_0) \int (\partial p/\partial z) dt$  and  $v = -(1/\rho_0) \int (\partial p/\partial r) dt$  result in  $|u| \approx k|p|/\rho_0 \omega$  and  $|v| \approx (2|p|/\rho_0)(r/a^2)G^2$ , re-

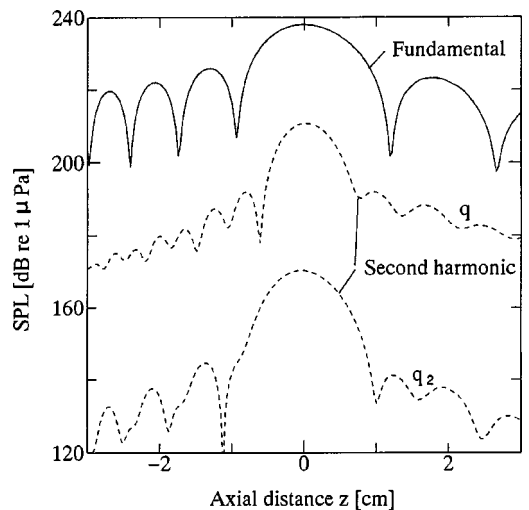


FIG. B2. On-axis pressure amplitudes of the fundamental and second harmonic components. For the second harmonic, two curves, one from the nonlinearity source  $q = q_1 + q_2$ , and the other from  $q_2$ , are given.



spectively, where  $G$  is the focusing gain defined as the ratio of the Rayleigh distance  $ka^2/2$  to the focal length  $d$ . The ratio  $v/u$  then becomes  $(2/k)rG^2/a^2=(r/d)G$ . In Fig. 7 in the present article, for example,  $d=10$  cm,  $a=4.5$  cm, and  $G=45$ ,  $v/u$  becomes approximately 0.5 at  $r=1$  mm. The farther the observer moves away from the axis, the more the ratio increases. However, the velocities  $u$  and  $v$  themselves decay abruptly off the axis. The second requirement of  $\partial u/\partial z$  being larger than  $\partial u/\partial r$  is verified in a similar way. If the source  $q_2$  has hypothetically the same order of magnitude as  $q_1$ , the fields must be different from those theoretically predicted by Westervelt's equation, especially in the post-focal region, because of the cumulative effect of nonlinearity. For a more rigorous discussion about whether Westervelt's equation still describes focused beams from a widely opening aperture well, further investigation is essential to evaluate nonlinearity source terms adequately.

<sup>1</sup>J. Naze Tjøtta and S. Tjøtta, "Model equation and boundary conditions for the sound field from a high-frequency, strongly curved and highly intense transducer," *Acta Acustica* **1**, 69–87 (1993).  
<sup>2</sup>M. F. Hamilton and C. L. Morfey, "Model equations," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, CA, 1998), Chap. 3.  
<sup>3</sup>J. Naze Tjøtta, S. Tjøtta, and E. H. Vefring, "Effects of focusing on the nonlinear interaction between two collinear finite amplitude sound beams," *J. Acoust. Soc. Am.* **89**, 1017–1027 (1991).  
<sup>4</sup>T. Tsuchiya and Y. Kagawa, "Finite element analysis of focusing field in nonlinear acoustic waves," *Jpn. J. Appl. Phys., Suppl.* **30**, 51–53 (1991).  
<sup>5</sup>V. P. Kuznetsov, "Equations of nonlinear acoustics," *Sov. Phys. Acoust.* **16**, 467–470 (1971).  
<sup>6</sup>T. Yano, private communication, 1996.  
<sup>7</sup>P. T. Christopher and K. J. Parker, "New approaches to nonlinear diffractive field propagation," *J. Acoust. Soc. Am.* **90**, 488–499 (1991).  
<sup>8</sup>P. T. Christopher, "Modeling the Dornier HM3 Lithotripter," *J. Acoust. Soc. Am.* **96**, 3088–3095 (1994).  
<sup>9</sup>J. Tavakkoli, D. Cathignol, R. Souchon, and O. A. Sapozhnikov, "Modeling of pulsed finite-amplitude focused sound beams," *J. Acoust. Soc. Am.* **104**, 2061–2072 (1998).  
<sup>10</sup>K. Naugolnykh and L. A. Ostrovsky, *Nonlinear Wave Processes in Acous-*

*tics* (Cambridge University Press, Cambridge, 1998), Sec. 4.5.  
<sup>11</sup>B. Ystad and J. Berntsen, "Numerical solution of parabolic equations for strongly curved focusing sources," *Acust. Acta Acust.* **82**, 698–706 (1996).  
<sup>12</sup>C. Flammer, *Spheroidal Wave Functions* (Stanford University Press, Stanford, 1957).  
<sup>13</sup>T. Kamakura, T. Ishiwata, and K. Matsuda, "A new theoretical approach to the analysis of nonlinear sound beams using the oblate spheroidal coordinate system," *J. Acoust. Soc. Am.* **105**, 3083–3086 (1999).  
<sup>14</sup>G-P. J. Too and J. H. Ginsberg, "Cylindrical and spherical coordinate versions of NPE for transient and steady-state sound beams," *Trans. ASME* **114**, 420–424 (1992).  
<sup>15</sup>H. T. O'Neil, "Theory of focusing radiators," *J. Acoust. Soc. Am.* **21**, 516–526 (1949).  
<sup>16</sup>F. Coulouvrat, "Continuous field radiated by a geometrically focused transducer: Numerical investigation and comparison with an approximate model," *J. Acoust. Soc. Am.* **94**, 1663–1675 (1993).  
<sup>17</sup>B. G. Lucas and T. G. Muir, "The field of a focusing source," *J. Acoust. Soc. Am.* **72**, 1289–1296 (1982).  
<sup>18</sup>M. A. Averkiou and M. F. Hamilton, "Measurements of harmonic generation in a focused finite-amplitude sound beam," *J. Acoust. Soc. Am.* **98**, 3439–3442 (1995).  
<sup>19</sup>J. A. TenCate, "An experimental investigation of the nonlinear pressure field produced by a plane circular piston," *J. Acoust. Soc. Am.* **94**, 1084–1089 (1993).  
<sup>20</sup>J. C. Piquette and A. L. Van Buren, "Nonlinear scattering of acoustic waves by vibrating surfaces," *J. Acoust. Soc. Am.* **76**, 880–889 (1984).  
<sup>21</sup>J. Saneyoshi, K. Kikuchi, and O. Nomoto, *Chouonpa Gijutsu Binran* (Nikkan Kogyo, Tokyo, 1980), Chap. 4.  
<sup>22</sup>S. Nachev, D. Cathignol, J. Naze Tjøtta, A. M. Berg, and S. Tjøtta, "Investigation of a high intensity sound beam from a plane transducer. Experimental and theoretical results," *J. Acoust. Soc. Am.* **98**, 2303–2323 (1995).  
<sup>23</sup>T. Kamakura, N. Hamada, K. Aoki, and Y. Kumamoto, "Nonlinearly generated spectral components in the nearfield of a directive sound source," *J. Acoust. Soc. Am.* **85**, 2331–2337 (1989).  
<sup>24</sup>H. Hobæk and B. Ystad, "Experimental and numerical investigation of shock wave propagation in the post focal region of a focused sound field," *Acust. Acta Acust.* **83**, 978–986 (1997).  
<sup>25</sup>M. F. Hamilton, "Sound beams," in *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, CA, 1998), Chap. 8.  
<sup>26</sup>T. Kamakura, K. Matsuda, Y. Kumamoto, and M. A. Breazeale, "Acoustic streaming induced in focused Gaussian beams," *J. Acoust. Soc. Am.* **97**, 2740–2746 (1995).

# Propagation of acoustic pulses in material with hysteretic nonlinearity

Vitalyi Gusev

Laboratoire de Physique de l'Etat Condensé, UPRESA-CNRS 6087, Faculté des Sciences, Ecole Nationale Supérieure d'Ingénieurs du Mans, Université du Maine, 72085 Le Mans, France

(Received 16 April 1999; accepted for publication 12 March 2000)

Evolution equations for propagation of both unipolar and bipolar acoustic pulses are derived by using hysteretic stress-strain relationships. Hysteretic stress-strain loops that incorporate quadratic nonlinearity are derived by applying the model of Preisach–Mayergoyz space for the characterization of structural elements in a micro-inhomogeneous material. Exact solutions of the nonlinear evolution equations predict broadening in time and reduction in amplitude of a unipolar finite-amplitude acoustic pulse. In contrast with some earlier theoretical predictions, the transformation of the pulse shape predicted here satisfies the law of “momentum” conservation (the “equality of areas” law in nonlinear acoustics of elastic materials). A bipolar pulse of nonzero momentum first transforms during its propagation into a unipolar pulse of the same duration. This process occurs in accordance with the “momentum” conservation law and without formation of shock fronts in the particle velocity profile. © 2000 Acoustical Society of America.

[S0001-4966(00)04406-4]

PACS numbers: 43.25.Dc [MFH]

## INTRODUCTION

There exists nowadays a strong consensus that the general theory of nonlinear acoustic wave propagation in micro-inhomogeneous materials should incorporate a stress-strain relationship with hysteresis.<sup>1–4</sup> The presence of an additional quadratic nonlinearity (so-called hysteretic nonlinearity) in micro-inhomogeneous materials provides an explanation for observations such as the proportionality of the shift in the resonance frequency of a vibrating bar to the amplitude of oscillations,<sup>1,5</sup> and the observation of a quadratic dependence of the third harmonic amplitude on the excitation amplitude in some experiments.<sup>6</sup>

For the description of resonance bar experiments, the theoretical analysis of the wave equation with hysteretic nonlinearity is usually based on one or another method of successive approximations (perturbation methods).<sup>1,5,6</sup> These perturbation methods were also applied for analytical<sup>4,7,8</sup> and numerical<sup>9</sup> analysis of initially harmonic sound wave propagation in a hysteretic half-space. In particular, stimulated backscattering of acoustic waves in materials with hysteretic nonlinearity was predicted.<sup>8</sup> In the case of a medium with hysteresis and relaxation, the developed theory<sup>4</sup> provides a possible explanation of the experimentally observed dependence of the nonlinear properties of piezoelectric transducer (PZT) ceramics on frequency.<sup>10</sup>

Recently, exact analytical solutions of the evolution equation for the propagation of initially harmonic plane waves in a hysteretic medium were established.<sup>4,11</sup> Exact solutions for quasiperiodic waves in hysteretic materials predict a possibility of self-focusing of acoustic beams and existence of shear-horizontal surface acoustic wave (i.e., localization due to dynamic softening of the hysteretic medium in the regions of high sound level).<sup>11</sup>

The present paper is devoted to exact analytical description of acoustic pulse evolution in a medium with hysteretic

nonlinearity. This research was motivated by two main reasons.

First, in general, acoustic waves can be excited in a pulsed mode with higher amplitudes than in continuous mode, and this favors an observation of nonlinear effects in their propagation. In addition to explosion applications in geophysical research, finite-amplitude acoustic pulses and even shock waves can be generated in laboratory experiments using laser action on materials (see, for example, Ref. 12 for references). Both unipolar strain pulses and bipolar strain pulses can be generated by laser sources.<sup>12,13</sup> Plain acoustical pulses of complicated profiles can be, in principle, excited when different physical mechanisms<sup>12,13</sup> are involved in the optoacoustic transformation (for example, the thermoelastic mechanism and ablation). Though laser-generated acoustic pulses have proved to be useful for evaluation of nonlinear properties of materials in diverse applications<sup>12</sup> (see also Ref. 25 for a recent experiment in a granular medium), they are clearly not the only method of generating high-amplitude acoustic pulses. For example, experiments on the nonlinear transformation of acoustic pulse excited in aluminum by an electromagnetic-induction type transducer were reported as long as 25 years ago.<sup>14</sup>

Second, both previous attempts to theoretically treat acoustic pulse propagation in a medium with hysteretic nonlinearity<sup>15,16</sup> predict the diminishing of motion momentum of an acoustic disturbance as it propagates [see Fig. 2 of Ref. 15 and Eqs. (18), (20), and (24) of Ref. 16]. In our opinion, this prediction contradicts the basic equation for plane motion of a solid,

$$\rho_0 \frac{\partial v}{\partial t} = \frac{\partial \sigma}{\partial x}. \quad (1)$$

Here  $\rho_0$  is the equilibrium density of a solid,  $v$  is the displacement (or particle) velocity,  $\sigma$  is the stress associated

with an acoustic wave. In fact, Eq. (1) for pulsed processes is equivalent to the equation for momentum. By integrating Eq. (1) over the space coordinate  $x$ , and by taking into account that for a pulsed process  $\sigma(x = \pm\infty) = 0$ , it can be derived that

$$\frac{\partial}{\partial t} \left( \int_{-\infty}^{\infty} v dx \right) = 0. \quad (2)$$

Thus the area of the particle velocity profile, which is known to be proportional to the momentum of the acoustic pulse,<sup>17</sup> is conserved during its propagation. For the weak nonlinear disturbances analyzed in Ref. 16, the linear strain  $s = \partial u / \partial x$  associated with a unidirectional wave is proportional to the particle velocity  $v = \partial u / \partial t$  (where  $u$  is the mechanical displacement). Consequently, in our opinion, the predictions of Ref. 16 formulated for unipolar strain pulses are in contradiction with Eq. (2). In principle, Eq. (2) proves as well the conservation of the area of stress profile when variations of mechanical modulus caused by hysteresis are significantly smaller than the linear elastic value. However, by integrating Eq. (1) over time  $t$ , and by taking into account that  $v(t = \pm\infty) = 0$  (since we can choose  $t = -\infty$  before the initial moment of the sound excitation and we can assume that at  $t = \infty$  the acoustic energy is completely absorbed), it can be derived that

$$\frac{\partial}{\partial x} \left( \int_{-\infty}^{\infty} \sigma dt \right) = 0. \quad (3)$$

The conservation law of Eq. (3) is valid for an arbitrary stress-strain relationship. In our opinion, the prediction in Fig. 2 of Ref. 15 (where the area of the unipolar stress pulse diminishes with propagation distance) contradicts the conservation law of Eq. (3).

In the next section the mathematical model for description of hysteretic nonlinearity is briefly introduced. The fundamental assumption in acoustic pulse propagation prediction (as related to the existence of memory effects in microinhomogeneous materials) is discussed. The particular hysteretic stress-strain loops necessary to describe acoustic pulse evolution are derived theoretically. In Sec. II the description of the evolution of unipolar sound pulses is given on the basis of the exact solution of the nonlinear evolution equation. In Sec. III the exact description of the propagation of bipolar acoustic pulses is obtained and analyzed. It should be noted that, though the momentum conservation law of Eqs. (2) and (3) has not been used as a condition for the identification of the “physical” wave profile in cases of mathematically multivalued solutions of nonlinear equations, nevertheless the predicted evolution of pulses does take place in accordance with this conservation law. The proposed theory predicts broadening of a unipolar pulse. A bipolar acoustic pulse (of nonzero initial momentum) in its propagation through a hysteretic medium, first, transforms into a unipolar pulse of the same duration, and only after that broadening of the pulse starts.

## I. MATHEMATICAL MODEL

In the analysis presented here it is assumed that in a material with hysteresis the stress associated with an acoustic wave,  $\sigma$ , depends not only on the strain associated with the acoustic wave,  $s$ , but also on the variation  $p'_0$  of a parameter  $p_0$  responsible for the hysteretic behavior of a medium<sup>2</sup>

$$\sigma \approx \sigma_{ss} + \sigma_{p_0} p'_0. \quad (4)$$

Here the derivatives  $\sigma_s \equiv (\partial \sigma / \partial s)_{p_0}$  and  $\sigma_{p_0} \equiv (\partial \sigma / \partial p_0)_s$  describe the weighted sensitivity of the stress to the strain  $s$  and the parameter  $p_0$ , respectively. Importantly, no assumptions have to be made about the physical identity of the introduced parameter  $p_0$ .<sup>2</sup>

In the mathematical description of the parameter  $p_0$  assumed in this paper (the model of Preisach–Mayergoyz space<sup>2,18,19</sup>) it appears that the second term in Eq. (4) is nonlinear.<sup>2,4</sup> It can be presented in the form

$$\sigma_{p_0} p'_0 = \sigma_{p_0} p_s [s, s_c] s, \quad (5)$$

where the sensitivity  $p_s \equiv \partial p_0 / \partial s$  of the hysteretic parameter  $p_0$  to strain is a linear function of the strain  $s$  and of a characteristic value  $s_c$  of the strain. Consequently, the second term in Eq. (4) is quadratic in strain and describes quadratic hysteretic nonlinearity.<sup>2,4,11</sup> In the present investigation the hysteretic nonlinearity is considered to be dominant in comparison with quadratic elastic nonlinearity [this is the reason why the term  $(1/2)\sigma_{ss}s^2$  is neglected in the expansion of Eq. (4)]. Kinematic nonlinearity of the relationship between the strain  $s$  and spatial derivatives of the mechanical displacement  $u$  is also considered to be negligible. Consequently, for plane longitudinal acoustic waves propagating along the  $x$ -axis we are using the representation  $s = \partial u / \partial x$  of linear strain.

Under the assumptions of Eqs. (4) and (5) the equation of motion Eq. (1) can be presented in the form

$$\rho_0 \frac{\partial^2 u}{\partial t^2} = \sigma_s \frac{\partial^2 u}{\partial x^2} + \sigma_{p_0} p_s \left[ \frac{\partial u}{\partial x}, \left( \frac{\partial u}{\partial x} \right)_c \right] \frac{\partial^2 u}{\partial x^2}. \quad (6)$$

Using the method of a slowly varying wave profile,<sup>19</sup> Eq. (6) is transformed into an evolution equation for the strain wave propagating in the positive direction of the  $x$ -axis,

$$\frac{\partial s}{\partial x} - \frac{1}{2c_0} \left( \frac{\sigma_{p_0}}{\sigma_s} \right) p_s [s, s_c] \frac{\partial s}{\partial \tau} = 0. \quad (7)$$

Here  $\tau \equiv t - (x/c_0)$  is the retarded time and  $c_0 \equiv \sqrt{\sigma_s / \rho_0}$  is the adiabatic velocity of sound. In accordance with Eq. (7), to describe sound propagation it is necessary to specify the dependence of the hysteretic parameter  $p_0$  on strain (i.e., the function  $p_s [s, s_c]$  should be modeled).

In the present investigation we adopt an approach developed in Refs. 2 and 18–20 for description of hysteresis and memory effects. In accordance with this paradigm,<sup>2,20</sup> the equilibrium macroscopic internal parameter  $p_0(s)$  of a medium with hysteresis is described as a linear superposition of corresponding parameters  $p_0(s, s_1, s_2)$  of unspecified individual microscopic mechanical elements contributing to the inelasticity of a material,

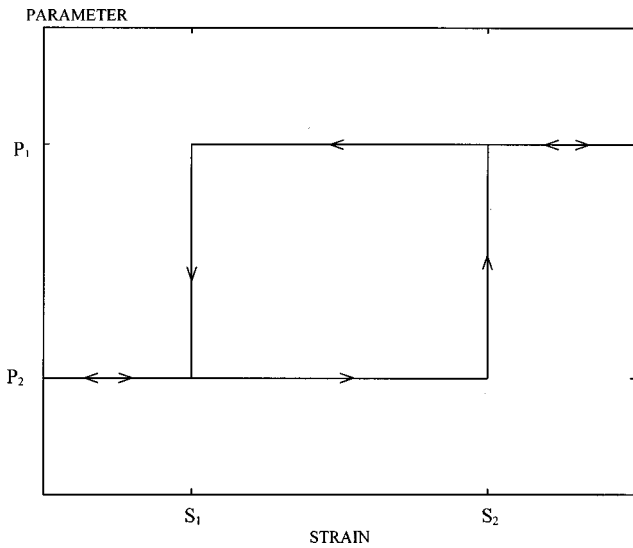


FIG. 1. Dependence of the internal parameter of individual hysteretic mechanical elements on strain. The arrowheads indicate the direction of strain variation.

$$p_0(s) = \int_{-\infty}^{s_2} ds_1 \int_{s_1}^{\infty} ds_2 p_0(s, s_1, s_2) f_{12}(s_1, s_2). \quad (8)$$

In the superposition of Eq. (8) it is assumed that the dependence of the parameter  $p_0(s, s_1, s_2)$  of an individual mechanical element on the strain  $s$  contains hysteresis (Fig. 1). The individual mechanical element can only be in two possible states characterized by the values  $p_2$  and  $p_1$  of the internal parameter. Transition from the first to the second state with increasing strain occurs when  $s = s_2$ ; the inverse transition occurs with decreasing strain when  $s = s_1 < s_2$  (Fig. 1). The fact that  $s_1 < s_2$  (i.e., the fact of hysteretic behavior) is taken into account mathematically by imposing limits on the integration in Eq. (8). Although the parameters  $p_2$  and  $p_1$  can vary from one individual element to another, it is assumed that  $p_2 - p_1 \equiv \Delta p = \text{const}$ , while a possible variation of this difference among the elements is incorporated in the distribution function  $f_{12}(s_1, s_2)$  of mechanical elements in the  $(s_1, s_2)$ -coordinates (Preisach–Mayergoyz coordinates<sup>2,18,19</sup>).

Though specification of the mechanism of hysteresis is not necessary it is sometimes suitable to associate the transition of an element from state  $p_1$  to state  $p_2$  with the opening of this element (a crack, for example<sup>20</sup>). Then, the inverse transition  $p_2 \rightarrow p_1$  is associated with the closing of this element (Fig. 1).

It should be noted that assuming infinitely fast transitions between the state  $p_1$  and the state  $p_2$  following the strain path (Fig. 1), we neglect in this paper all relaxation processes which can accompany these transitions.<sup>4,21–23</sup>

It was demonstrated that the model of Eq. (8) and of Fig. 1 presented above describes hysteresis and memory effects.<sup>2</sup> Memory effects are formally incorporated in the model through the requirement that each individual element be in one of two possible states  $p_1$  or  $p_2$ . The current state of the element depends on the history of material loading [i.e., on strain (or stress) history]. Thus the parameter  $p_0(s, s_1, s_2)$  in Eq. (8) depends, in fact, on the history of the acoustic strain,

not just on the current value of the acoustic strain. In other words, to describe acoustic wave propagation it is not sufficient to know the distribution function  $f_{12}(s_1, s_2)$  of hysteretic elements. It is also necessary for each element which can be affected by the wave (i.e., for each element that can change its state under the action of the wave) to know its initial state before the arrival of the wave (i.e., to know if it is open or closed).

In general, to identify the initial state of all the elements contributing to Eq. (8), knowledge of the whole strain history of the material under consideration is required. The absence of this information can be overcome in the case of periodic wave processes,<sup>2,4,11</sup> because the history of loading of the elements affected by the wave is known for periodic phenomena (that is a periodic wave “prepares” the material for its transition). It is assumed that the periodic process started an infinitely long time ago at some boundary and, consequently, at the observation instant the process is periodic not only at the boundary but along the whole path of acoustic wave propagation. In this case it is possible to predict the hysteretic stress-strain relationship along the whole path of the wave propagation.

For example, in the case of a periodic process with a single maximum and a single minimum during a period, the stress-strain relationship of Eqs. (4), (5), and (8) predicts a classical closed hysteretic loop.<sup>2,19,20</sup> In the particular case of a periodic process symmetrical relative to the average strain level (i.e., with symmetrical positive and negative phases relative to average level), the stress-strain relation predicted by Eqs. (4), (5), and (8) has the form<sup>4</sup>

$$\frac{\sigma'}{\rho_0 c_0^2} = s' - h \left[ s'^A s' + \frac{1}{2} (s'^2 - (s'^A)^2) \text{sgn} \left( \frac{\partial s'}{\partial t} \right) \right]. \quad (9)$$

Here the functions  $\sigma' \equiv \sigma - \langle \sigma \rangle$  and  $s' \equiv s - \langle s \rangle$  denote the changes of stress and strain relative to their average levels defined by  $\langle \sigma \rangle \equiv (\sigma_{\max} + \sigma_{\min})/2$  and  $\langle s \rangle \equiv (s_{\max} + s_{\min})/2$ , where the indexes “max” and “min” are used to define the maximum and minimum values of the functions. The amplitude  $s'^A$  of strain variation relative to average level is defined in Eq. (9) by  $s'^A \equiv (s_{\max} - s_{\min})/2$ . It should be noted that (for the assumed symmetrical periodic process) the amplitudes of both positive and negative phases of the strain wave  $s'$  are equal to  $s'^A$ . In Eq. (9),  $h$  denotes the dimensionless parameter of hysteretic nonlinearity,<sup>4</sup>  $h \equiv -(\sigma_{p_0}/\sigma_s)\Delta p \langle f_{12} \rangle$ , where  $\langle f_{12} \rangle$  is the average value of the distribution function in the region of the  $(s_1, s_2)$ -plane defined by  $\langle s \rangle - s'^A \leq s_1 \leq s_2 \leq \langle s \rangle + s'^A$ .<sup>4</sup> A positive value of the parameter  $h$  assumed later in the paper corresponds to hysteresis being dissipative.

The hysteretic loop described by Eq. (9) is qualitatively presented in Fig. 2 in the normalized coordinates  $\sigma'_h \equiv h\sigma'/\rho_0 c_0^2$ ,  $s'_h \equiv h s'$  for the particular case  $h s'^A = 1/2$  ( $\sigma'_h = \{s'_h - [s'^2_h - (1/2)^2] \text{sgn}(\partial s'_h/\partial t)\}/2$ ). It is worth noting, however, that the general presentation of Fig. 2 for the hysteretic loop [in particular, the orientation of the loop on the  $(\sigma'_h, s'_h)$  plane and its width] depends on the parameter  $h s'^A$ . For physical analysis, it is more instructive to separate linear

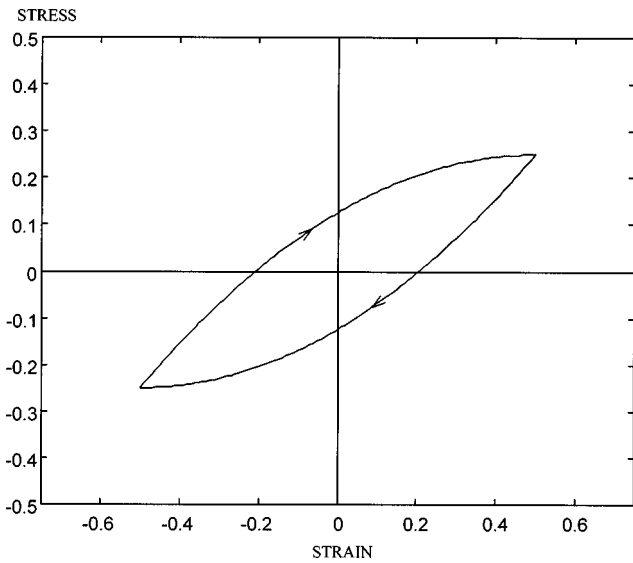


FIG. 2. Hysteresis stress-strain ( $\sigma'_h - s'_h$ ) loop for a periodic process with a single maximum and a single minimum during a period. The hysteretic loop is presented in the normalized coordinates  $\sigma'_h \equiv h\sigma'/\rho_0 c_0^2$ ,  $s'_h \equiv hs'$  for the particular case  $hs'^A = 1/2$  ( $\sigma'_h = \{s'_h - [s_h^2 - (1/2)^2] \text{sgn}(\partial s'_h / \partial t)\} / 2$ ). The arrowheads indicate the path direction with increasing time.

$\sigma'_L$  and nonlinear (hysteretic)  $\sigma'_{NL}$  parts in the stress-strain relationship of Eq. (9),

$$\frac{\sigma'_{NL}}{\rho_0 c_0^2} = -h \left[ s'^A s' + \frac{1}{2} (s'^2 - (s'^A)^2) \text{sgn} \left( \frac{\partial s'}{\partial t} \right) \right].$$

The presentation of the latter relationship in the normalized coordinates  $\sigma_N \equiv \sigma'_{NL} / h (s'^A)^2 \rho_0 c_0^2$ ,  $s_N \equiv s'_h / s'^A$  (Fig. 3) eliminates all the parameters

$$\sigma_N = - \left[ s_N + \frac{1}{2} (s_N^2 - 1) \text{sgn} \left( \frac{\partial s_N}{\partial t} \right) \right].$$

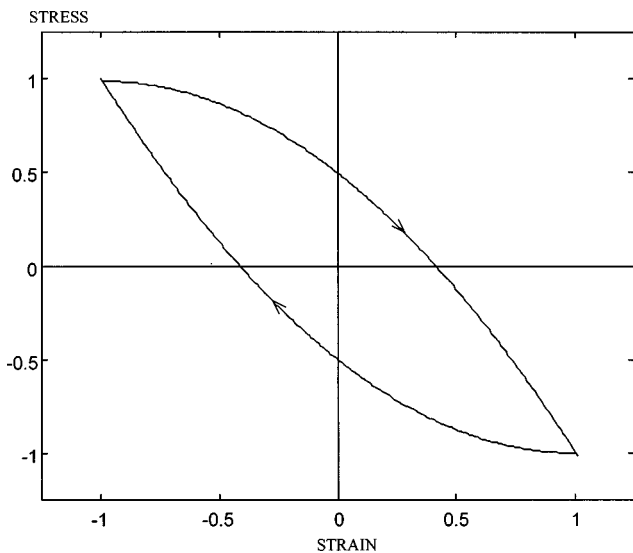


FIG. 3. Universal presentation of the nonlinear (hysteretic) part of the stress-strain relationship (in normalized coordinates  $\sigma_N - s_N$ ) for a periodic process with a single maximum and a single minimum during a period. The normalized stress variation and the normalized strain variation are defined by  $\sigma_N \equiv \sigma'_{NL} / h (s'^A)^2 \rho_0 c_0^2$ ,  $s_N \equiv s'_h / s'^A$ , where  $s'^A \equiv (s_{\max} - s_{\min}) / 2$  is the amplitude of the strain variation. The arrowheads indicate the path direction with increasing time.

The true contribution of the universal loop shown in Fig. 3 to the actual hysteretic loop (Fig. 2) depends, of course, on the hysteretic parameter  $h$  and on the amplitude  $s'^A$  of strain variation.

When using Eq. (9) to describe periodic acoustic wave propagation, the dependence of the wave amplitude on the distance  $x$  from the acoustic source,  $s'^A = s'^A(x)$ , should be taken into account.<sup>4,9,11</sup> In order to apply the stress-strain relationship of Eq. (9) in the analysis of initially harmonic processes<sup>4,11</sup> it is of crucial importance that the average level of the strain  $\langle s \rangle$  for these processes does not depend on the spatial coordinate  $x$ . The independence of the average strain on the propagation distance follows from conservation of symmetry of positive and negative phases even in the process of nonlinear attenuation.<sup>4,11</sup>

Clearly, the stress-strain relationship of Eq. (9) (Fig. 2) is similar to the one described by the hysteretic loop in Fig. 1 of Ref. 15. It can be used for the analysis of the nonlinear evolution of acoustic pulses; however, not via direct identification of  $\sigma'$  and  $s'$  in Eq. (9) with the variations of stress and strain relative to their undisturbed values in a material (as was done in Ref. 15). In fact, following the procedure of Ref. 15 and assuming that the stress-strain loop on the coordinate plane  $(\sigma, s)$  is identical to the theoretical hysteretic loop of Eq. (9) on the coordinate plane  $(\sigma', s')$  (Fig. 2), it is impossible to use this loop for the analysis of acoustic pulse propagation. It is reasonable to accept that the acoustic pulse starts from the point  $(\sigma = 0, s = 0)$ . But the stress-strain loop of Eq. (9) after identification of the variables  $\sigma' = \sigma$ ,  $s' = s$  does not pass through this point (see Fig. 1 of Ref. 15). So, it does not contain information which is necessary for the prediction of the acoustic pulse evolution.

From a physics point of view the problem is, of course, deeper [that is, in general, it can not be solved just by a correct choice of the variable transformation in Eq. (9)]. For the description of a single acoustic pulse propagation (single-shot experiments) the stress-strain relationship of Eq. (9) (which corresponds to the loading history of a periodic process with a single minimum and a single maximum over a period<sup>4,9,11</sup>) is useless. This is because the stress-strain relationship that is required to describe acoustic pulse propagation will depend on the particular method of the transition from periodic loading to single-shot loading (i.e., on how we switch off the periodic signal before exciting the single pulse). However, fortunately, single-shot experiments are very rare. Most experiments involving acoustic pulses are, in fact, performed in a pulsed-periodic regime. By repeating the pulsed action and by applying the methods of signal stacking and averaging, the signal-to-noise ratio can be significantly improved. Consequently, the analysis of pulsed-periodic acoustic signal propagation is the most important problem. For this kind of process it can be assumed that the history of loading corresponds to the history of loading by identical pulses of finite length arriving earlier in time. Thus this is just a particular case of the general periodic process. As a result, it is possible to describe the pulsed-periodic process using the stress-strain loop of Eq. (9) provided there is only a single maximum and minimum in a single period.

It should be pointed out that, in fact, for the validity of

the models derived below it is not required for a pulsed process to be strictly periodic. It is sufficient that the pulsed process consists of repetitive pulsed action by similar pulses, where the separation in time of subsequent pulses significantly exceeds the duration of the pulses (in order to avoid overlapping upon the nonlinear transformation).

For analysis of such a repetitive pulsed process, Eq. (9) should be first rewritten in the original coordinates  $\sigma$  and  $s$ ,

$$\frac{\sigma - \langle \sigma \rangle}{\rho_0 c_0^2} = s - \frac{s_{\max} + s_{\min}}{2} - h \left[ \frac{s_{\max} - s_{\min}}{2} \left( s - \frac{s_{\max} + s_{\min}}{2} \right) + \frac{(s - s_{\max})(s - s_{\min})}{2} \operatorname{sgn} \left( \frac{\partial s}{\partial t} \right) \right]. \quad (10)$$

The choice of the parameter  $\langle \sigma \rangle$  in Eq. (10) provides the opportunity to identify the stress-strain hysteretic loop for a particular repetitive process. In fact, the parameter  $\langle \sigma \rangle$  should be chosen in accordance with the detected signal.

For example, we view a periodic process with symmetrical positive and negative phases ( $s_{\max} = -s_{\min} \equiv s^A$ ) as the process in which the stress varies relative to zero level  $\langle \sigma \rangle = 0$ . By substituting  $s_{\max} = -s_{\min} \equiv s^A$  and  $\langle \sigma \rangle = 0$  in Eq. (10), we derive the stress-strain relationship

$$\frac{\sigma}{\rho_0 c_0^2} = s - h \left[ s^A s + \frac{s^2 - (s^A)^2}{2} \operatorname{sgn} \left( \frac{\partial s}{\partial t} \right) \right],$$

which is identical to Eq. (10) if we identify  $\sigma$  with  $\sigma'$ , and  $s$ ,  $s^A$  with  $s'$ ,  $s'^A$ . This confirms that the hysteretic loop in Fig. 1 of Ref. 15 is valid for the description of a symmetrical periodic process with average levels of stress and strain equal to zero.

If we consider an acoustic pulse as the disturbance from some constant level (which is fixed before the arrival of the acoustic pulse), then, in fact,  $\sigma$  and  $s$  in Eq. (10) are the deviations of stress and strain from this level. Following this consideration the average stress level  $\langle \sigma \rangle$  should be chosen such that the acoustic pulse starts from the point  $\sigma = 0$ ,  $s = 0$  in the  $(\sigma, s)$ -plane. This yields

$$\frac{\langle \sigma \rangle}{\rho_0 c_0^2} = \frac{s_{\max} + s_{\min}}{2} + h \left[ -\frac{s_{\max}^2 - s_{\min}^2}{2} + \frac{s_{\max} s_{\min}}{2} \operatorname{sgn} \left( \frac{\partial s}{\partial t} (\sigma = 0, s = 0) \right) \right]. \quad (11)$$

In accordance with Eq. (11), the average stress induced by a repetitive pulsed process depends on the maximum and minimum values of the strain as well as on the sign of the first phase of the acoustic pulse (i.e.,  $\langle \sigma \rangle$  depends on the sign of the derivative  $\partial s / \partial t$  at the leading front of the acoustic pulse). Substitution of Eq. (11) into Eq. (10) leads to the general description of the stress-strain relationship for acoustic pulses with a single maximum and a single minimum

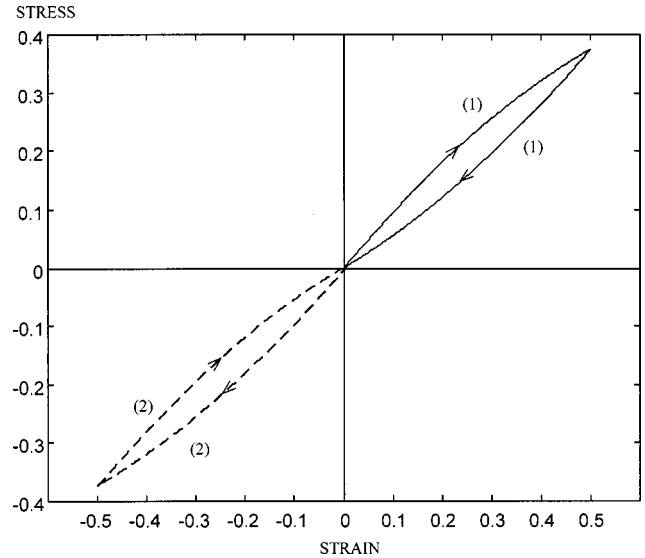


FIG. 4. Hysteretic stress-strain ( $\sigma_h - s_h$ ) relationship for a sequence of a unipolar dilatation [loop (1)] and a compression [dashed loop (2)] acoustic pulse with a single extremum of equal amplitude. The arrows indicate the direction of strain variation in time.

$$\frac{\sigma}{\rho_0 c_0^2} = s - \frac{h}{2} \left[ (s_{\max} - s_{\min}) s + (s - s_{\max})(s - s_{\min}) \times \operatorname{sgn} \left( \frac{\partial s}{\partial t} \right) - s_{\max} s_{\min} \operatorname{sgn} \left( \frac{\partial s}{\partial t} (0) \right) \right]. \quad (12)$$

In Secs. II and III, the stress-strain relationship as derived in Eq. (12) is applied to describe unipolar and bipolar acoustic pulses, respectively.

## II. PROPAGATION OF UNIPOLAR ACOUSTIC PULSES

For a unipolar dilatation pulse ( $s \geq 0$ ) the stress-strain relationship of Eq. (12) (because of the restricting condition  $s_{\min} = 0$ ) additionally simplifies

$$\frac{\sigma}{\rho_0 c_0^2} = s \left\{ 1 - \frac{h}{2} [s_{\max} + (s - s_{\max}) \operatorname{sgn}(\partial s / \partial t)] \right\}. \quad (13)$$

The hysteretic loop of Eq. (13) is presented in Fig. 4 [loop (1)] in its normalized variables  $\sigma'_h \equiv h\sigma' / \rho_0 c_0^2$ ,  $s_h \equiv hs$  for the particular case when  $hs_{\max} = 1/2$  ( $\sigma'_h = s_h \{ (3/2) - [s_h - (1/2)] \operatorname{sgn}(\partial s_h / \partial t) \} / 2$ ). Comparison of the theoretically derived hysteretic loop (1) in Fig. 4 with the hysteretic model of the medium presented in Fig. 1 of Ref. 16 demonstrates their qualitative similarity. Thus the theoretical conclusions of Ref. 16 [those contradicting the conservation law of Eq. (2)] are not due to a qualitatively different model applied for the description of the properties of the hysteretic medium but, perhaps, are due to a mistake in the particular solution of the pulse propagation problem derived in Ref. 16.

For a unipolar compression pulse ( $s \leq 0$ ) the stress-strain relationship of Eq. (12) (because of the restricting condition  $s_{\max} = 0$ ) can also be additionally simplified to

$$\frac{\sigma}{\rho_0 c_0^2} = s \left\{ 1 - \frac{h}{2} [-s_{\min} + (s - s_{\min}) \operatorname{sgn}(\partial s / \partial t)] \right\}. \quad (14)$$

The hysteretic loop of Eq. (14) is presented in Fig. 4 [loop (2)] in its normalized variables  $\sigma_h \equiv h\sigma/\rho_0 c_0^2$ ,  $s_h \equiv hs$  for the particular case when  $hs_{\min} = -1/2$  ( $\sigma_h = s_h\{(3/2) - [s_h + (1/2)]\text{sgn}(\partial s_h/\partial t)\}/2$ ).

From the evident symmetry of the derived expressions [Eqs. (13) and (14)] for both hysteretic loops (see also Fig. 4), it follows that the description of the propagation of dilatation and compression acoustic pulses is similar. Here, we choose to analyze the propagation of compression pulses in more detail because they are more easily generated by laser action on materials.<sup>12,13</sup> In the case of a unipolar compression pulse the description of Eq. (14) for a hysteretic stress-strain relationship leads to a breakup of the evolution equation of Eq. (7) into two different equations,

$$\begin{aligned} \partial s/\partial x - (h/2c_0)s\partial s/\partial \tau &= 0 \quad \text{if } \partial s/\partial \tau < 0, \\ \partial s/\partial x - (h/2c_0)(s_{\min} - s)\partial s/\partial \tau &= 0 \quad \text{if } \partial s/\partial \tau > 0. \end{aligned} \quad (15)$$

The first equation of Eqs. (15) describes the nonlinear transformation of the leading front ( $\partial s/\partial \tau < 0$ ) of the acoustic pulse. The second equation of Eqs. (15) describes the nonlinear transformation of the trailing front ( $\partial s/\partial \tau > 0$ ) of the acoustic pulse. For convenience of comparison with the classical results for gases and for fluids,<sup>17</sup> Eqs. (15) are rewritten in terms of the particle velocity  $v$  of the acoustic wave ( $v = -c_0 s$  for a pulse propagating in the positive direction),

$$\begin{aligned} \partial v/\partial x + (h/2c_0^2)v\partial v/\partial \tau &= 0 \quad \text{if } \partial v/\partial \tau > 0, \\ \partial v/\partial x + (h/2c_0^2)(v_{\max} - v)\partial v/\partial \tau &= 0 \quad \text{if } \partial v/\partial \tau < 0. \end{aligned} \quad (16)$$

By normalizing the particle velocity with respect to the initial pulse amplitude  $v_0 = v_{\max}(x=0)$  ( $V \equiv v/v_0$ ), the retarded time to the initial characteristic duration  $\tau_0$  of the acoustic pulse ( $\theta \equiv \tau/\tau_0$ ) and the spatial coordinate to the characteristic length  $x_{NL} = 2c_0^2\tau_0/hv_0$  of the nonlinear transformation ( $\zeta \equiv x/x_{NL}$ ), the system of Eqs. (16) is presented in the form

$$\begin{aligned} \partial V/\partial \zeta + V\partial V/\partial \theta &= 0 \quad \text{if } \partial V/\partial \theta > 0, \\ \partial V/\partial \zeta + (V_{\max} - V)\partial V/\partial \theta &= 0 \quad \text{if } \partial V/\partial \theta < 0. \end{aligned} \quad (17)$$

The exact analytical solution of this system [Eqs. (17)] of quasilinear partial differential equations in implicit form is

$$\begin{aligned} V &= V_L(\theta - V\zeta) \quad \text{if } \partial V/\partial \theta > 0, \\ V &= V_T\left(\theta - \int_0^\zeta V_{\max}(\zeta')d\zeta' + V\zeta\right) \quad \text{if } \partial V/\partial \theta < 0, \end{aligned} \quad (18)$$

where the functions  $V_L(\theta)$  and  $V_T(\theta)$  describe the leading and the trailing fronts of the acoustic pulse at the boundary ( $\zeta=0$ ) of the medium. In comparison with solutions of qualitatively similar equations proposed in Ref. 16 [see Eqs. (6)–(8) in Ref. 16] the solution of Eqs. (18) precisely accounts for the fact that the pulse amplitude  $V_{\max}$  changes during the pulse propagation [ $V_{\max} = V_{\max}(\zeta)$ ]. This important point has been neglected in the theory of Ref. 16, and therefore, the predicted results showed a diminishing acoustic pulse momentum during the pulse propagation through a hysteretic medium.

In accordance with Eqs. (18), to describe the pulse shape evolution, it is necessary to know the dependence  $V_{\max}$

$= V_{\max}(\zeta)$  of the acoustic pulse amplitude on propagation distance. To this extent, we use the condition  $V(\theta \rightarrow \theta_{\max}, \partial V/\partial \theta > 0) = V(\theta \rightarrow \theta_{\max}, \partial V/\partial \theta < 0) = V_{\max}$  of the particle velocity continuity at the peak of the acoustic pulse.<sup>4,16</sup> Here,  $\theta_{\max}$  denotes the position of the pulse peak in time. In fact, at the point  $V = V_{\max}$ ,  $\theta = \theta_{\max}$ , the solution of Eqs. (18) provides a set of two equations

$$\begin{aligned} V_{\max} &= V_L(\theta_{\max} - V_{\max}\zeta), \\ V_{\max} &= V_T\left(\theta_{\max} - \int_0^\zeta V_{\max}(\zeta')d\zeta' + V_{\max}\zeta\right), \end{aligned} \quad (19)$$

for two unknown functions [i.e., for  $V_{\max} = V_{\max}(\zeta)$  and  $\theta_{\max} = \theta_{\max}(\zeta)$ ]. The unknown function  $\theta_{\max} = \theta_{\max}(\zeta)$  can be easily solved for. Equations (19) lead to

$$\int_0^\zeta V_{\max}(\zeta')d\zeta' - 2V_{\max}\zeta + V_T^{-1}(V_{\max}) + V_L^{-1}(V_{\max}) = 0, \quad (20)$$

where the notations  $V_T^{-1}$  and  $V_L^{-1}$  were introduced for inverse functions. It is convenient to transfer the integral equation of Eq. (20) into a differential equation

$$\left[ \frac{1}{V_T'(V_{\max})} - \frac{1}{V_L'(V_{\max})} - 2\zeta \right] \frac{dV_{\max}}{d\zeta} - V_{\max} = 0, \quad (21)$$

where  $V_T'$  and  $V_L'$  denote the derivatives of the functions  $V_T$  and  $V_L$  over their argument (i.e., over  $V_{\max}$ ). Equation (21) is a closed first-order ordinary differential equation for the determination of  $V_{\max} = V_{\max}(\zeta)$ . The boundary condition is  $V_{\max}(\zeta=0) = 1$ .

It is important for a qualitative understanding of the nonlinear process of acoustic pulse transformation in a medium with hysteresis, that precise analytical integration of Eq. (21) can be done for pulses of triangular profile. In this case  $V(\zeta=0) = V_L(\theta) = \theta/\theta_{\max}(0)$  for  $\partial V/\partial \theta > 0$  and  $V(\zeta=0) = V_T(\theta) = (1-\theta)/(1-\theta_{\max}(0))$  for  $\partial V/\partial \theta < 0$ , where  $\theta_{\max}(0) \equiv \theta_{\max}(\zeta=0)$  is the initial time coordinate of the acoustic pulse peak. Then Eq. (21) takes the form  $(1+2\zeta)dV_{\max}/d\zeta + V_{\max} = 0$  and its solution predicts a reduction of the pulse amplitude upon propagation in the following manner:

$$V_{\max}(\zeta) = 1/\sqrt{1+2\zeta}. \quad (22)$$

At large distances  $\zeta \gg 1$  the asymptotic behavior is  $V_{\max}(\zeta) \propto 1/\sqrt{\zeta}$ . It is worth noting that the law of amplitude reduction is the same for all triangular pulses [i.e., it does not depend on the initial position  $\theta_{\max}(0) \equiv \theta_{\max}(\zeta=0)$  of the pulse peak]. With the use of Eq. (22), the position  $\theta_{\max}$  of the pulse can be found from any one of Eqs. (19),

$$\theta_{\max}(\zeta) = [\theta_{\max}(0) + \zeta]/\sqrt{1+2\zeta}. \quad (23)$$

In accordance with the solution of Eq. (18) the leading (starting) point ( $V=0, \partial V/\partial \theta > 0$ ) of the acoustic pulse is invariant (immobile). This statement also follows from the representation of the equation for the evolution of the leading front in Eqs. (17) in an equivalent system of ordinary differential equations,

$$d\theta/d\zeta = V, \quad dV/d\zeta = 0. \quad (24)$$

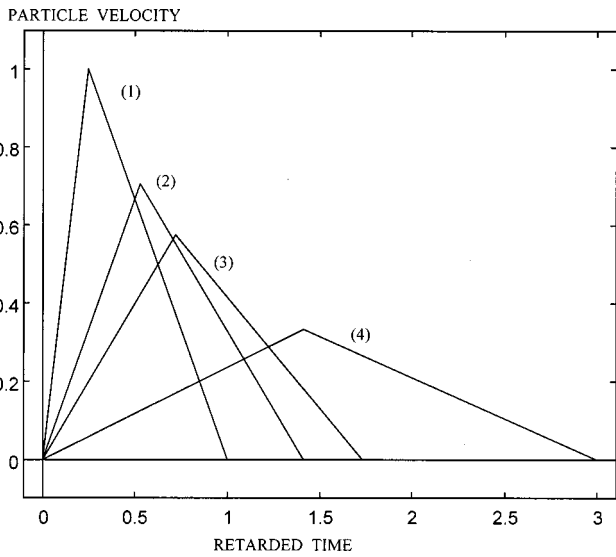


FIG. 5. Evolution of a triangular compression pulse in a medium with hysteresis. The profiles of the particle velocity are presented at distances  $\zeta = 0$  [profile (1)],  $\zeta = 0.5$  [profile (2)],  $\zeta = 1$  [profile (3)], and  $\zeta = 4$  [profile (4)].

The first of these equations predicts the invariance (immobility) of the leading point (of course, in the coordinate system moving at sonic velocity  $c_0$ ). As a consequence, Eq. (23) is the solution for the duration of the leading front, and it describes broadening of the leading front upon propagation. The equivalent system of ordinary differential equations for the evolution of the trailing front is

$$d\theta/d\zeta = V_{\max} - V, \quad dV/d\zeta = 0. \quad (25)$$

The first of these equations predicts that the trailing point ( $V=0, \partial V/\partial\theta < 0$ ) of the acoustic pulse moves along the  $\theta$ -axis. By using Eq. (22) and  $V=0$  to integrate this equation, the dependence of the acoustic pulse duration on the propagation distance is obtained,

$$\theta_0(\zeta) = \sqrt{1 + 2\zeta}. \quad (26)$$

Consequently, it follows from Eqs. (26) and (23) that the trailing front of the pulse also broadens during pulse propagation.

In accordance with the solutions in Eqs. (26) and (22), the pulse area  $V_{\max}(\zeta)\theta_0(\zeta)/2 = 1/2$  (related to its momentum) is conserved in pulse propagation. In line with the discussion presented in the Introduction, this proves the validity of the derived solutions. An example of the transformation of a triangular compression pulse in a medium with hysteresis [described by Eqs. (18) and (22)] is presented in Fig. 5. It can be checked straightforwardly that, to describe the transformation of a dilatation pulse of the same shape, it is sufficient to mirror the profiles in Fig. 5 relative to the horizontal axis. In fact, the peaks of a compression pulse as well as of a dilatation pulse both propagate at subsonic velocity in hysteretic media.

In conclusion of this Sec. II it is worth noting that even though analytical integration of Eq. (21) is only possible for some particular initial profiles of acoustic pulses, qualitative analysis of the acoustic pulse transformation can be obtained for an arbitrary initial pulse shape. At a distance  $d\zeta \ll 1$  from

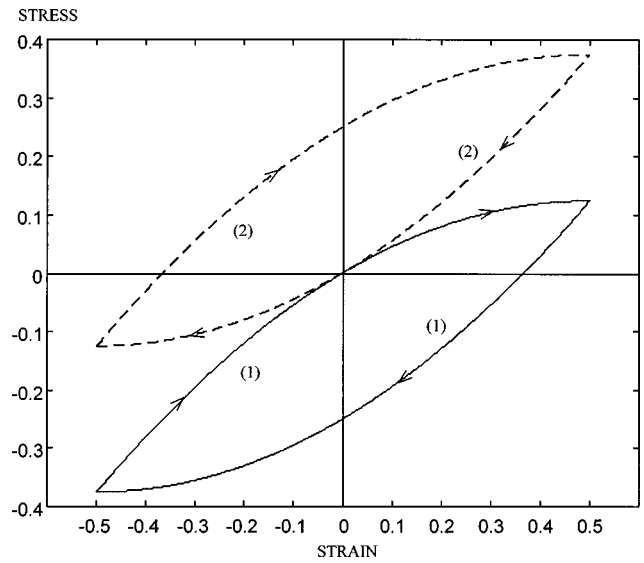


FIG. 6. Hysteresis stress-strain ( $\sigma_h - s_h$ ) relationship for a sequence of equal amplitude bipolar acoustic pulses. Loop (1) corresponds to pulses starting with a dilatation phase. The dashed loop (2) corresponds to pulses starting with the compression phase. The arrows indicate the direction of strain variation in time.

the boundary, the points of the leading front are shifted relative to the initial profile  $V_L(\theta)$  at  $\zeta=0$  in accordance with the rules of Eq. (24), that is  $V \cong \text{const} \cong V_L(\theta, \zeta=0)$  and  $d\theta \cong V_L(\theta, \zeta=0)d\zeta$ . By doing so, we can find the position of the leading front at  $d\zeta \ll 1$  by moving each point of the initial front horizontally [ $V \cong \text{const} \cong V_L(\theta, \zeta=0)$ ] by distance in time  $d\theta \cong V_L(\theta, \zeta=0)d\zeta$  along the  $\theta$ -axis. The peak of the pulse (which is defined as the intersection of the leading and trailing fronts) also changes its position. With the updated coordinates [ $V_{\max}(d\zeta), \theta_{\max}(d\zeta)$ ] of the pulse peak, the trailing edge can be shifted in accordance with the rules  $V \cong \text{const} \cong V_T(\theta, \zeta=0)$  and  $d\theta \cong [V_{\max}(d\zeta) - V_T(\theta, \zeta=0)]d\zeta$  that follow from Eq. (25). As a result, we obtain the approximate profile of the acoustic pulse at a distance  $d\zeta$  from the boundary. The same procedure can be applied in an iterative manner to this newly derived profile, and so on.

### III. PROPAGATION OF BIPOLAR ACOUSTIC PULSES

The stress-strain relationship expressed in Eq. (12) for the description of propagation of bipolar pulses is different for pulses starting with a dilatation phase ( $\partial s/\partial t(s=0) > 0$ ) than for pulses starting with a compression phase [ $\partial s/\partial t(s=0) < 0$ ]. The hysteretic loop of Eq. (12) is presented in Fig. 6 in normalized variables  $\sigma_h \equiv h\sigma/\rho_0 c_0^2$ ,  $s_h \equiv hs$  for the case  $hs_{\max} = -hs_{\min} = 1/2$  ( $\sigma_h = \{s_h - [s_h^2 - (1/2)^2] \text{sgn}(\partial s_h/\partial t) - (1/2)^2 \text{sgn}(\partial s_h(0)/\partial t)\}/2$ ). Loop (1) in Fig. 6 corresponds to a bipolar pulse starting with a dilatation phase [ $\partial s/\partial t(s=0) > 0$ ]. Loop (2) in Fig. 6 corresponds to a bipolar pulse starting with a compression phase [ $\partial s/\partial t(s=0) < 0$ ]. Of course, in a general case, the orientation and the width of the hysteretic loops described by Eq. (12) depend on the values of the parameters  $h$ ,  $s_{\max}$ , and  $s_{\min}$ .

Application of the stress-strain relationship, given by Eq. (12), in the analysis of wave propagation leads to a



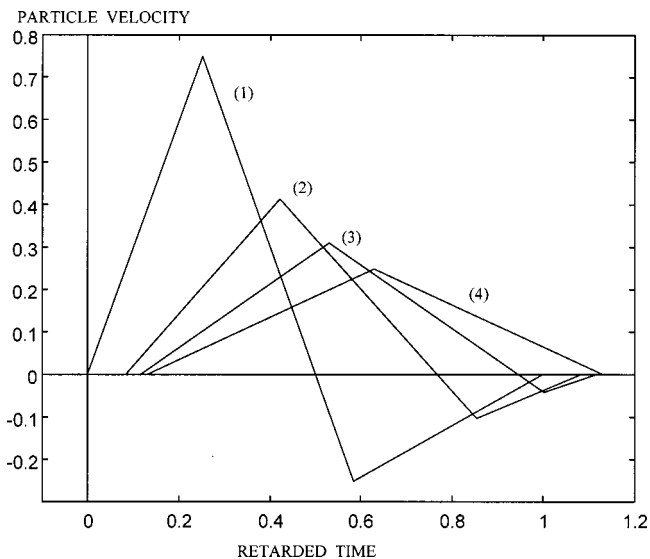


FIG. 7. Transformation of a bipolar pulse into a unipolar pulse upon propagation through a medium with hysteretic nonlinearity in a case where the initial amplitude of the compression phase exceeds the initial amplitude of the dilatation phase. The net area of the pulse in the (particle velocity, retarded time) plane is positive. Wave profiles are presented at distances  $\zeta = 0$  [profile (1)],  $\zeta = 0.5$  [profile (2)],  $\zeta = 1$  [profile (3)], and  $\zeta = \zeta_- = 5/3$  [profile (4)].

breakup of the evolution equation [Eq. (7)] into two equations,

$$\partial s / \partial x + (h/2c_0)(s_{\max} - s) \partial s / \partial \tau = 0 \quad \text{if } \partial s / \partial \tau < 0,$$

$$\partial s / \partial x + (h/2c_0)(s - s_{\min}) \partial s / \partial \tau = 0 \quad \text{if } \partial s / \partial \tau > 0.$$

In terms of the particle velocity  $v$  of the acoustic pulse, these equations can be rewritten as

$$\partial v / \partial x + (h/2c_0^2)(v - v_{\min}) \partial v / \partial \tau = 0 \quad \text{if } \partial v / \partial \tau > 0,$$

$$\partial v / \partial x + (h/2c_0^2)(v_{\max} - v) \partial v / \partial \tau = 0 \quad \text{if } \partial v / \partial \tau < 0.$$

In normalized variables  $V \equiv v / [v_{\max}(0) - v_{\min}(0)]$ ,  $\theta \equiv \tau / \tau_0$ ,  $\zeta \equiv x / x_{NL}$  (where  $v_{\max}(0) - v_{\min}(0) \equiv v_{\max}(x=0) - v_{\min}(x=0)$  is the characteristic peak-to-peak amplitude of the pulse at the boundary,  $\tau_0$  is the characteristic duration of the pulse at the boundary, and  $x_{NL} = 2c_0^2 \tau_0 / h[v_{\max}(0) - v_{\min}(0)]$  is the characteristic nonlinear length) the system of evolution equations takes the form

$$\begin{aligned} \partial V / \partial \zeta + (V - V_{Z\min}) \partial V / \partial \theta &= 0 \quad \text{if } \partial V / \partial \theta > 0, \\ \partial V / \partial \zeta + (V_{\max} - V) \partial V / \partial \theta &= 0 \quad \text{if } \partial V / \partial \theta < 0. \end{aligned} \quad (27)$$

The equivalent systems of ordinary differential equations are

$$d\theta / d\zeta = V - V_{\min}, \quad dV / d\zeta = 0 \quad \text{if } \partial V / \partial \theta > 0, \quad (28)$$

and

$$d\theta / d\zeta = V_{\max} - V, \quad dV / d\zeta = 0 \quad \text{if } \partial V / \partial \theta < 0. \quad (29)$$

In the following the propagation of bipolar pulses starting with a compression phase is chosen for detailed analysis only because they are more easily generated by laser action on the materials.<sup>12,13</sup> In this case both the leading front and the trailing front correspond to an increasing particle velocity, i.e., to  $\partial v / \partial \tau > 0$  (Fig. 7). In accordance with Eq. (28) both the motion of the leading point of the pulse [i.e., the

point with initial coordinates  $V=0$ ,  $\theta(\zeta=0) = \theta_L(\zeta=0) = 0$ ] and the motion of the trailing point [i.e., the point with initial coordinates  $V=0$ ,  $\theta(\zeta=0) = \theta_T(\zeta=0) = 1$ ] along the  $\theta$ -axis are described by the same equation,

$$d\theta_L = d\theta_T = -V_{\min}(\zeta) d\zeta. \quad (30)$$

Consequently, as long as the pulse is bipolar, the duration of the pulse does not change as it propagates [ $d\theta_0(\zeta) = d(\theta_L - \theta_T) \equiv 0$ , where  $\theta_0(\zeta)$  denotes the normalized duration of the acoustic pulse,  $\theta_0(0) = 1$ ]. Since by definition  $V_{\min}(\zeta) \leq 0$ , it can be concluded that a bipolar pulse propagates at subsonic velocity (assuming softening of a hysteretic material by an acoustic field, i.e.,  $h > 0$ ). For  $V = \text{const} = 0$ , Eq. (29) describes the motion along the  $\theta$ -axis of the transition point [ $V=0, \theta_s = \theta_s(\zeta)$ ] separating positive and negative phases,

$$d\theta_s = V_{\max}(\zeta) d\zeta. \quad (31)$$

By combining Eqs. (30) and (31), the change in duration  $\theta_+ = \theta_+(\zeta) \equiv \theta_s(\zeta) - \theta_L(\zeta)$  of the compression phase of the pulse is obtained,

$$d\theta_+ = d[\theta_s(\zeta) - \theta_L(\zeta)] = [V_{\max}(\zeta) + V_{\min}(\zeta)] d\zeta. \quad (32)$$

In accordance with Eq. (32), the duration of the compression phase broadens if the amplitude  $V_+(\zeta) \equiv V_{\max}(\zeta)$  of this phase is larger than the amplitude  $V_-(\zeta) \equiv -V_{\min}(\zeta)$  of the dilatation phase. The compression phase shortens in duration if  $V_+(\zeta) < V_-(\zeta)$ . So, in the general case  $V_+(\zeta) \neq V_-(\zeta)$ , both the duration  $\theta_+(\zeta)$  of the compression phase and the duration  $\theta_-(\zeta) \equiv 1 - \theta_+(\zeta)$  of the dilatation phase are changing during bipolar pulse propagation, while the total duration of the pulse is conserved,

$$\theta_+(\zeta) + \theta_-(\zeta) \equiv 1. \quad (33)$$

The exact analytical solutions of Eq. (27) in implicit form are

$$\begin{aligned} V &= V_L \left( \theta + \int_0^\zeta V_{\min}(\zeta') d\zeta' - V\zeta \right) \\ &\quad \text{if } \theta_L(\zeta) \leq \theta \leq \theta_{\max}(\zeta), \\ V &= V_S \left( \theta - \int_0^\zeta V_{\max}(\zeta') d\zeta' + V\zeta \right) \\ &\quad \text{if } \theta_{\max}(\zeta) \leq \theta \leq \theta_{\min}(\zeta), \\ V &= V_T \left( \theta + \int_0^\zeta V_{\min}(\zeta') d\zeta' - V\zeta \right) \\ &\quad \text{if } \theta_{\min}(\zeta) \leq \theta \leq \theta_T(\zeta). \end{aligned} \quad (34)$$

Here the functions  $V_L(\theta)$  and  $V_T(\theta)$  describe the leading and trailing fronts of the acoustic pulse at the boundary ( $\zeta = 0$ ) of the medium, while the function  $V_S(\theta)$  describes the front that separates the two phases [i.e., the latter function describes the transition from compression to dilatation in the time interval  $\theta_{\max}(\zeta) \leq \theta \leq \theta_{\min}(\zeta)$ , where  $\theta_{\max}(\zeta)$  and  $\theta_{\min}(\zeta)$  denote the positions of the peaks of the pulse].

The conditions  $V(\theta \rightarrow \theta_{\max}, \partial V / \partial \theta > 0) = V(\theta \rightarrow \theta_{\max}, \partial V / \partial \theta < 0) = V_{\max}$  and  $V(\theta \rightarrow \theta_{\min}, \partial V / \partial \theta < 0) = V(\theta \rightarrow \theta_{\min}, \partial V /$

$\partial\theta>0$ )= $V_{\min}$  expressing the continuity of the particle velocity in the pulse amplitude and the derived solution Eqs. (34), provide four equations,

$$\begin{aligned} V_{\max} &= V_L \left( \theta_{\max} + \int_0^\zeta V_{\min}(\zeta') d\zeta' - V_{\max}\zeta \right), \\ V_{\max} &= V_S \left( \theta_{\max} - \int_0^\zeta V_{\max}(\zeta') d\zeta' + V_{\max}\zeta \right), \\ V_{\min} &= V_S \left( \theta_{\min} - \int_0^\zeta V_{\max}(\zeta') d\zeta' + V_{\min}\zeta \right), \\ V_{\min} &= V_T \left( \theta_{\min} + \int_0^\zeta V_{\min}(\zeta') d\zeta' - V_{\min}\zeta \right), \end{aligned} \quad (35)$$

for the unknown coordinates  $[\theta_{\max}(\zeta), V_{\max}(\zeta)]$  and  $[\theta_{\min}(\zeta), V_{\min}(\zeta)]$  of pulse peaks in the  $(\theta, V)$  plane. Elimination of  $\theta_{\max}(\zeta)$  and  $\theta_{\min}(\zeta)$  in Eqs. (35) leads to two integral equations,

$$\begin{aligned} \int_0^\zeta V_{\min}(\zeta') d\zeta' + \int_0^\zeta V_{\max}(\zeta') d\zeta' - 2V_{\max}\zeta \\ + V_S^{-1}(V_{\max}) - V_L^{-1}(V_{\max}) &= 0, \\ \int_0^\zeta V_{\min}(\zeta') d\zeta' + \int_0^\zeta V_{\max}(\zeta') d\zeta' - 2V_{\min}\zeta \\ + V_S^{-1}(V_{\min}) - V_T^{-1}(V_{\min}) &= 0, \end{aligned} \quad (36)$$

or equivalently to two differential equations,

$$\begin{aligned} \left[ \frac{1}{V_S'(V_{\max})} - \frac{1}{V_L'(V_{\max})} - 2\zeta \right] \frac{dV_{\max}}{d\zeta} + V_{\min} - V_{\max} &= 0, \\ \left[ \frac{1}{V_S'(V_{\min})} - \frac{1}{V_T'(V_{\min})} - 2\zeta \right] \frac{dV_{\min}}{d\zeta} + V_{\max} - V_{\min} &= 0, \end{aligned} \quad (37)$$

for the maximum and minimum velocities in the acoustic pulse. We can deduce an algebraic relationship between  $V_{\max}(\zeta)$  and  $V_{\min}(\zeta)$  by taking the difference of the two equations in Eqs. (36). Subsequently, any of Eqs. (37) can be transformed into a closed first-order differential equation for a single unknown.

For a qualitative understanding of the nonlinear transformation of bipolar pulses in hysteretic media a rather simple analytical description can be obtained in the case where the profile of the pulse at the boundary can be presented by a combination of linear profiles (sawtooth profile). For example, in the simplest case in Fig. 7 the bipolar pulse can be modeled by using only three parameters  $V_{\max}(0)$ ,  $\theta_{\max}(0)$  and  $\theta_{\min}(0)$ ,

$$\begin{aligned} V_L(\theta) &= V_{\max}(0) \frac{\theta}{\theta_{\max}(0)}, \\ V_T(\theta) &= V_{\min}(0) \frac{1-\theta}{1-\theta_{\min}(0)}, \\ V_S(\theta) &= \frac{V_{\max}(0)[\theta_{\min}(0)-\theta] + V_{\min}(0)[\theta-\theta_{\max}(0)]}{\theta_{\min}(0)-\theta_{\max}(0)}. \end{aligned} \quad (38)$$

Note that, because of the adopted normalization, there is a relation  $V_{\min}(0)=V_{\max}(0)-1$  between  $V_{\min}(0)$  and  $V_{\max}(0)$ . For an initial pulse as described by Eqs. (38) the algebraic equation obtained from Eqs. (36) can be suitably written as

$$(2\zeta + \alpha_+^{-1})V_+ + (2\zeta + \alpha_-^{-1})V_- = 1. \quad (39)$$

Here the compact notations  $\alpha_+ \equiv V_+(0)/\theta_+(0)$  and  $\alpha_- \equiv V_-(0)/\theta_-(0)$  are introduced. Using the same notations, the system of Eqs. (37) takes the form

$$\begin{aligned} [2\zeta + \alpha_+^{-1}] \frac{dV_+}{d\zeta} + V_+ + V_- &= 0, \\ [2\zeta + \alpha_-^{-1}] \frac{dV_-}{d\zeta} + V_+ + V_- &= 0. \end{aligned} \quad (40)$$

By substitution of Eq. (39) in the first of the equations in Eqs. (40) and by integration, the expression for the compression amplitude as a function of the propagation distance is derived,

$$\begin{aligned} V_{\max}(\zeta) &= V_{\max}(0) [(\sqrt{(1+2\alpha_-\zeta)/(1+2\alpha_+\zeta)} - \alpha_-)/ \\ & (1-\alpha_-)]. \end{aligned} \quad (41)$$

The expression for the amplitude of the dilatation peak follows from Eq. (41) and from the symmetry of Eqs. (39) and (40) relative to an interchange of the indexes “-” and “+,”

$$\begin{aligned} V_{\min}(\zeta) &= V_{\min}(0) [(\sqrt{(1+2\alpha_+\zeta)/(1+2\alpha_-\zeta)} - \alpha_+)/ \\ & (1-\alpha_+)]. \end{aligned} \quad (42)$$

The solution in Eq. (41) predicts that the compression phase can completely disappear at some finite distance  $\zeta = \zeta_+$  from the boundary  $[V_{\max}(\zeta=\zeta_+)=0]$ ,

$$\zeta_+ = \frac{1}{2\alpha_-} \frac{(\alpha_-^2 - 1)}{(1 - \alpha_+ \alpha_-)} \equiv -\frac{\theta_+}{4\alpha_-} \frac{(\alpha_- + 1)}{A}. \quad (43)$$

Here  $A \equiv (V_+ \theta_+ - V_- \theta_-)/2 = \int_{-\infty}^{\infty} V d\theta$  is the net area of the pulse profile in the  $(\theta, V)$  plane at the boundary. It should be mentioned that Eq. (33) and the condition of normalization  $V_+(0) + V_-(0) = 1$  were extensively used to achieve the compact presentation of  $\zeta_+$  in Eq. (43). In accordance with Eq. (43), the coordinate  $\zeta_+$  is positive (and, consequently, the compression phase disappears at a finite distance  $\zeta_+$  from the boundary) only if  $A$  is negative. This prediction supports the law of net area conservation described in the Introduction. Evidently, if the initial net area is positive ( $A > 0$ ), the compression phase can not disappear, because the disappearance of the compression phase would mean that only the dilatation phase would be left and, thus,  $A(\zeta_+) < 0$ , which is in contradiction with the area conservation law. In contrast, when  $A > 0$  the dilatation phase of a bipolar pulse

disappears at a finite distance  $\zeta_- = \theta_-(\alpha_+ + 1)/(4\alpha_+A)$ . If  $A=0$  (net momentum equal to zero)  $\zeta_{\pm} = \infty$ , i.e., the pulse length remains the same at all times and distances.

It is worth noting that the variation of the compression pulse duration in Eq. (32) is not entirely controlled by the law of net area conservation. Surely, in the case when  $A < 0$ , for example, at the final stage ( $\zeta \rightarrow \zeta_+$ ) of the transformation of a bipolar pulse into a unipolar pulse, the compression phase is shortening [ $\theta_+(\zeta) \rightarrow 0$ ]. However, at the initial stage of a bipolar pulse transformation ( $\zeta \rightarrow 0$ ), the compression phase broadens in accordance with Eq. (32) if  $V_{\max}(0) > -V_{\min}(0)$ , i.e., if  $V_+ > V_-$ . In this case we define the coordinate  $\zeta_S$  as the transition distance between broadening of the compression pulse and shortening. This coordinate can be obtained from the condition  $V_{\max}(\zeta_S) = -V_{\min}(\zeta_S)$  and the derived expression in Eqs. (41) and (42),

$$\zeta_S = \frac{1}{2} \left( \frac{y^2 - 1}{\alpha_+ - \alpha_- y^2} \right),$$

$$y = \alpha_+ [1 + \text{sgn}(1 - \alpha_+) \sqrt{1 - \alpha_+^{-1} \alpha_-^{-1}}].$$

The expressions (41) and (42), when substituted in Eqs. (34), provide a complete description of the transformation of a bipolar pulse composed of two triangular phases [Eq. (38)].

We now evaluate some characteristic features of these solutions. In particular, in accordance with Eqs. (30) and (42), the position of the leading (starting) point of the acoustic pulse can be expressed as

$$\begin{aligned} \theta_L(\zeta) &= - \int_0^{\zeta} V_{\min}(\zeta') d\zeta' \\ &= - \frac{V_{\min}(0)}{2(1 - \alpha_+)} \left\{ \frac{1}{\alpha_-} \left[ \sqrt{(1 + 2\alpha_- \zeta)(1 + 2\alpha_+ \zeta)} - 1 \right. \right. \\ &\quad \left. \left. + \left( \sqrt{\frac{\alpha_-}{\alpha_+}} - \sqrt{\frac{\alpha_+}{\alpha_-}} \right) \right. \right. \\ &\quad \left. \left. \times \ln \left( \frac{\sqrt{\alpha_+} \sqrt{(1 + 2\alpha_- \zeta)} + \sqrt{\alpha_-} \sqrt{(1 + 2\alpha_+ \zeta)}}{\sqrt{\alpha_+} + \sqrt{\alpha_-}} \right) \right] \right. \\ &\quad \left. - 2\alpha_+ \zeta \right\}. \end{aligned} \quad (44)$$

The expression for the position of the phase separation point can be described with the help of Eq. (44) as

$$\begin{aligned} \theta_S(\zeta) - \theta_S(0) &= \int_0^{\zeta} V_{\max}(\zeta') d\zeta' \\ &= \theta_L(\zeta, V_{\min}(0) \rightarrow -V_{\max}(0), \text{“+”} \leftrightarrow \text{“-”}). \end{aligned} \quad (45)$$

Finally, with the help of Eqs. (34) and (44), the positions of the compression and the dilatation peaks are given by

$$\begin{aligned} \theta_{\max}(\zeta) &= \theta_L(\zeta) + \left[ \zeta + \frac{\theta_{\max}(0)}{V_{\max}(0)} \right] V_{\max}(\zeta), \\ \theta_{\min}(\zeta) &= \theta_L(\zeta) + 1 + \left[ \zeta - \frac{1 - \theta_{\min}(0)}{V_{\min}(0)} \right] V_{\min}(\zeta). \end{aligned}$$

PARTICLE VELOCITY

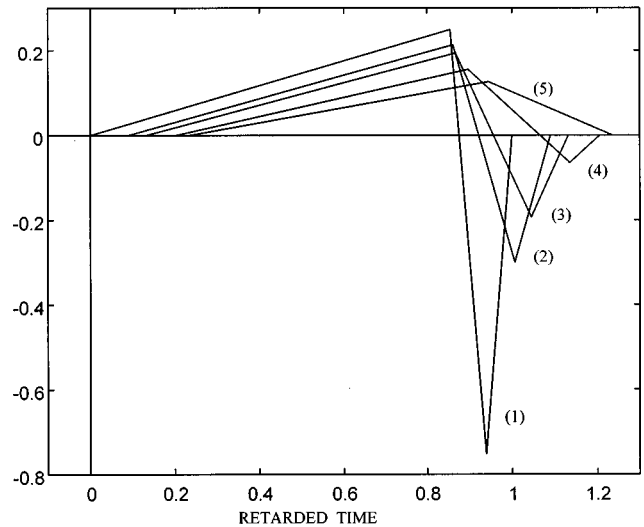


FIG. 8. Transformation of a bipolar pulse into a unipolar pulse upon propagation through a medium with hysteretic nonlinearity in a case where the initial amplitude of the compression phase is smaller than the initial amplitude of the dilatation phase. The net area of the pulse in the (particle velocity, retarded time) plane is positive. Wave profiles are presented at distances  $\zeta=0$  [profile (1)],  $\zeta=0.2$  [profile (2)],  $\zeta=\zeta_S=0.3743$  [profile (3)],  $\zeta=1$  [profile (4)], and  $\zeta=\zeta_-=2.25$  [profile (5)].

The analytical solutions are illustrated in Figs. 7 and 8 for two particular cases. In the first example (Fig. 7), the initial conditions [ $V_{\max}(0) > -V_{\min}(0)$ ,  $A > 0$ ] cause gradual broadening of the compression phase until the bipolar pulse transforms into a unipolar pulse. In the second example (Fig. 8), due to the initial conditions [ $V_{\max}(0) < -V_{\min}(0)$ ,  $A > 0$ ] the pulse transformation starts with a shortening of the compression phase. However, the amplitude-dependent attenuation causes a faster reduction of the dilatation phase amplitude. As a result, at a distance  $\zeta = \zeta_S$  [curve (3) in Fig. 8] the amplitudes of both phases become equal and broadening of the compression phase starts.

For the sawtooth pulses illustrated in Figs. 7 and 8 [with the use of the formulas Eqs. (41), (42) for  $V_{\max}(\zeta)$ ,  $V_{\min}(\zeta)$ , and of the formulas Eqs. (44), (45) for finding  $\theta_+(\zeta) = \theta_S(\zeta) - \theta_L(\zeta)$ ,  $\theta_-(\zeta) = 1 - \theta_+(\zeta)$ ], the conservation of net area  $A(\zeta) = A(\zeta=0)$  was carefully checked. This serves as an additional verification of the validity of the obtained analytical description. In principle, the law of net area conservation can be used as a starting point (i.e., as an additional condition instead of one of the conditions used in this paper) for deriving the solution of pulse propagation in a hysteretic medium.

It is worth noting that in accordance with the description in Eqs. (27)–(29) for nonlinear wave propagation in hysteretic media, the transformation of a bipolar pulse starting with a dilatation phase differs from the results presented in Figs. 7 and 8 just by an inversion relative to the  $V=0$  axis. In particular, these pulses are also subsonic (for  $h > 0$ ). Importantly, in the case of bipolar pulses of arbitrary initial profile (i.e., of non-sawtooth shape but still with single maximum and single minimum), a qualitative analysis of the pulse transformation based on the ordinary differential equations

(28), (29) can be performed (using a similar approach as described at the end of Sec. II).

Finally, it should be noted that after the transformation of a bipolar pulse into a unipolar pulse [i.e., at the distances  $\zeta \geq \zeta_+$  if  $A < 0$  or at the distances  $\zeta \geq \zeta_-$  if  $A > 0$ ] for the description of its propagation beyond the critical distance  $\zeta \geq \zeta_+$  or  $\zeta \geq \zeta_-$  one should apply the theory presented in Sec. II.

#### IV. CONCLUSION

This paper is devoted to the presentation of the exact analytical description of the transformation of acoustic pulses in a medium with hysteretic quadratic nonlinearity. The evolution equations for the propagation of both unipolar and bipolar acoustic pulses are derived by using hysteretic stress-strain relationships. Hysteretic stress-strain loops which incorporate quadratic nonlinearity are derived starting from the model of Preisach–Mayergoyz space for the characterization of structural elements in micro-inhomogeneous material.<sup>2,18,19</sup> The exact solution of the nonlinear evolution equations predict broadening in duration and reduction of the amplitude of a unipolar finite-amplitude acoustic pulse. In contrast with earlier theoretical predictions<sup>15,16</sup> the transformation of the pulse shape predicted in this paper satisfies the law of momentum conservation (the “equality of areas” law in nonlinear acoustics of elastic materials<sup>17</sup>).

Unfortunately we have not found any experimental results that allow direct comparison with the theoretical predictions derived here. In the experiments<sup>14</sup> the nonlinear transformation of the profile of a unipolar acoustic pulse in aluminum has been observed but not as a function of the pulse propagation distance but as a function of the annealing temperature for a fixed distance of the propagation. It has been found that increase of the annealing temperature during the preparation of a specimen causes reduction of the amplitude of the pulse and broadening of its duration without formation of a shock front (Fig. 14 in Ref. 14). These observations are in qualitative agreement with the theoretical predictions in Eq. (22) [ $V_{\max} = (1 + xv_0 h / c_0^2 \tau_0)^{-1/2}$ ] and Eq. (26) [ $\theta_0(\zeta) = (1 + xv_0 h / c_0^2 \tau_0)^{1/2}$ ] if we make an assumption that the parameter of the hysteretic quadratic nonlinearity,  $h$ , increases with the rise of the annealing temperature. Note that this assumption is in accordance with the established dependence of the parameter  $h$  in copper on the characteristics of the annealing process.<sup>6,24</sup> However, any quantitative comparison with the results of the experiments<sup>14</sup> is currently impossible.

In the case of a bipolar acoustic pulse the theory predicts that a bipolar pulse of nonzero momentum first transforms in its propagation into a unipolar pulse of the same duration. This process occurs in accordance with the momentum conservation law and without shock formation in the particle velocity profile. The theory predicts that the compression or dilatation phase of an acoustic pulse with the highest amplitude broadens upon propagation. However due to nonlinear amplitude reduction the ratio of the amplitudes of the compression and dilatation phases changes during propagation. Therefore initial broadening of a particular phase at short

propagation distances can be followed by shortening at large propagation distances.

In our opinion the theory presented in this paper is the necessary step toward the development of a general theory of finite-amplitude acoustic pulse propagation in micro-inhomogeneous materials. If necessary, ordinary elastic quadratic nonlinearity can be incorporated in this theory in a straightforward manner.<sup>4</sup> The approach described in this paper should find applications in geophysics (seismology) and nondestructive characterization of materials (from rocks to advanced nanocrystalline materials).

- <sup>1</sup> V. E. Nazarov, L. A. Ostrovsky, I. A. Soustova, and A. M. Sutin, “Nonlinear acoustics of micro-inhomogeneous media,” *Phys. Earth Planet. Inter.* **50**, 65–73 (1988).
- <sup>2</sup> R. A. Guyer, K. P. McCall, and G. N. Boitnott, “Hysteresis, discrete memory, and nonlinear wave propagation in rock: A new paradigm,” *Phys. Rev. Lett.* **74**, 3491–3494 (1995).
- <sup>3</sup> A. Kadish, P. A. Johnson, and B. Zinszner, “Evaluating hysteresis in earth materials under dynamic resonance,” *J. Geophys. Res. B* **101**, 25139–25147 (1996).
- <sup>4</sup> V. Gusev, W. Lauriks, and J. Thoen, “Dispersion of nonlinearity, nonlinear dispersion, and absorption of sound in micro-inhomogeneous materials,” *J. Acoust. Soc. Am.* **103**, 3216–3226 (1998).
- <sup>5</sup> V. E. Nazarov, L. A. Ostrovskii, I. A. Soustova, and A. M. Sutin, “Anomalous acoustic nonlinearity in metals,” *Akust. Zh.* **34**, 491–499 (1988) [English transl.: *Sov. Phys. Acoust.* **34**, 285–289 (1988)].
- <sup>6</sup> V. E. Nazarov, “Nonlinear acoustic effects in annealed copper,” *Akust. Zh.* **37**, 150–156 (1991) [English transl.: *Sov. Phys. Acoust.* **37**, 75–78 (1991)].
- <sup>7</sup> V. E. Nazarov and A. M. Sutin, “Harmonic generation in the propagation of elastic waves in nonlinear solid media,” *Akust. Zh.* **35**, 711–716 (1989) [English transl.: *Sov. Phys. Acoust.* **35**, 410–413 (1989)].
- <sup>8</sup> V. Gusev, H. Bailliet, P. Lotton, and M. Bruneau, “Interaction of the counter-propagating acoustic waves in the media with nonlinear dissipation and in the hysteretic media,” *Wave Motion* **29**, 211–221 (1999).
- <sup>9</sup> K. E.-A. Van den Abeele, P. A. Johnson, R. A. Guyer, and K. R. McCall, “On the quasianalytic treatment of hysteretic nonlinear response in elastic wave propagation,” *J. Acoust. Soc. Am.* **101**, 1885–1898 (1997).
- <sup>10</sup> J. K. Na and M. A. Breazeale, “Ultrasonic nonlinear properties of lead zirconate-titanate ceramics,” *J. Acoust. Soc. Am.* **95**, 3213–3221 (1994).
- <sup>11</sup> V. Gusev, C. Glorieux, W. Lauriks, and J. Thoen, “Nonlinear bulk and surface shear acoustic waves in materials with hysteresis and end-point memory,” *Phys. Lett. A* **232**, 77–86 (1997).
- <sup>12</sup> S. A. Akhmanov and V. Gusev, “Laser excitation of ultrashort acoustic pulses: New possibilities in solid-state spectroscopy, diagnostics of fast processes, and nonlinear acoustics,” *Usp. Fiz. Nauk* **162**, 3–87 (1992) [English transl.: *Sov. Phys. Usp.* **35**, 153–191 (1992)].
- <sup>13</sup> V. Gusev and A. Karabutov, *Laser Optoacoustics* (American Institute of Physics, New York, 1993).
- <sup>14</sup> Y. Yasumoto, A. Nakamura, and R. Takeuchi, “Developments in the use of acoustic shock pulses in the study of elastic properties of solids,” *Acustica* **30**, 260–267 (1974).
- <sup>15</sup> A. Kadish, “On the problem of source identification from the far-field observations in hysteretic materials,” *J. Acoust. Soc. Am.* **103**, 3196–3202 (1998).
- <sup>16</sup> V. E. Nazarov, “Propagation of a unipolar pulse in a medium with hysteretic nonlinearity,” *Akust. Zh.* **43**, 225–229 (1997) [English transl.: *Acoust. Phys.* **43**, 192–195 (1997)].
- <sup>17</sup> O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Consultants Bureau, New York, 1977).
- <sup>18</sup> I. D. Mayergoyz, “Hysteretic models from the mathematical and control theory point of view,” *J. Appl. Phys.* **57**, 3803–3805 (1985).
- <sup>19</sup> K. R. McCall and R. A. Guyer, “Equation of state and wave propagation in hysteretic nonlinear elastic materials,” *J. Geophys. Res.* **99**, 23887–23897 (1994).
- <sup>20</sup> D. J. Holcomb, “Memory, relaxation, and microfracturing in dilatant rock,” *J. Geophys. Res.* **86**, 6235–6248 (1981).
- <sup>21</sup> R. A. Guyer and P. A. Johnson, “Nonlinear mesoscopic elasticity: Evidence for a new class of materials,” *Phys. Today* **52**, 30–35 (1999).
- <sup>22</sup> J. A. TenCate and T. J. Shankland, “Slow dynamics in the nonlinear

- elastic response of Berea sandstone,” *Geophys. Res. Lett.* **23**, 3019–3022 (1996).
- <sup>23</sup>R. A. Guyer, K. R. McCall, and K. Van Den Abeele, “Slow elastic dynamics in a resonant bar of rock,” *Geophys. Res. Lett.* **25**, 1585–1588 (1998).
- <sup>24</sup>V. E. Nazarov, “Sound damping by sound in metals,” *Acoust. Lett.* **15**, 22–25 (1991).
- <sup>25</sup>I. B. Esipov, V. V. Zosimov, and K. I. Matveev, “Propagation of an acoustic pulse of finite amplitude in a granular medium,” *Akust. Zh.* **43**, 648–653 (1997) [English transl.: *Acoust. Phys.* **43**, 558–563 (1997)].

# Scattering of sonic booms by anisotropic turbulence in the atmosphere

Mark Kelly, Richard Raspet, and Henry E. Bass

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

(Received 1 June 1999; revised 1 January 2000; accepted 5 February 2000)

An earlier paper [J. Acoust. Soc. Am. **98**, 3412–3417 (1995)] reported on the comparison of rise times and overpressures of sonic booms calculated with a scattering center model of turbulence to measurements of sonic boom propagation through a well-characterized turbulent layer under moderately turbulent conditions. This detailed simulation used spherically symmetric scatterers to calculate the percentage of occurrence histograms of received overpressures and rise times. In this paper the calculation is extended to include distorted ellipsoidal turbules as scatterers and more accurately incorporates the meteorological data into a determination of the number of scatterers per unit volume. The scattering center calculation overpredicts the shifts in rise times for weak turbulence, and still underpredicts the shift under more turbulent conditions. This indicates that a single-scatter center-based model cannot completely describe sonic boom propagation through atmospheric turbulence. © 2000 Acoustical Society of America. [S0001-4966(00)02005-1]

PACS numbers: 43.25.Jh, 43.28.Mw [MAB]

## INTRODUCTION

Boulinger *et al.*<sup>1</sup> reported on the use of a scattering center-based calculation to develop distributions of sonic boom rise time and overpressure for different atmospheric conditions. The results of this calculation were compared to measured data from the JAPE-2 tests.<sup>2,3</sup> These tests provided simultaneous sonic boom measurements from T-38 aircraft along with detailed meteorological data. A scattering center-based calculation was employed, since this method had proven accurate in earlier studies of continuous wave propagation near the Earth's surface.<sup>4</sup> In addition, this method was flexible enough to allow the prediction of sonic boom propagation through a realistic representation of the measured turbulence field, including its spectrum and altitude dependence.

Turbulence was represented in Ref. 1 by spherical turbules; the resulting predictions of rise times and overpressures exhibited an increase in the peak of the rise time distribution and a decrease in the peak of the overpressure distribution, as did the measured data (compared to rise time and overpressure distributions from measurements made under calm conditions). The predicted shifts of the distribution peaks, however, were not as great as shifts shown by the histograms of the measured data.

Recent advances in meteorology have led to the modeling of the stretching of moderate scale turbulence in the atmosphere.<sup>5–7</sup> In this paper we report on the development of a single scatter formulation for ellipsoidal turbules and its consequent application to calculate rise time distributions for sonic boom propagation through the planetary boundary layer. The spherical turbule calculations displayed an excess of unperturbed waves relative to the measured results. A spherical turbule produces significant perturbations within a small elliptical region on the ground around the direction of incidence. An elliptical turbule produces significant perturbations over a much larger and irregular footprint on the

ground. It was expected that the more realistic turbules with a larger footprint per turbule would lead to improved predictions.

The calculated histograms of rise time occurrences are compared to measurements from the second Joint Acoustic Propagation Experiments (JAPE-2) conducted in 1991 at White Sands Missile Range. In particular, acoustic data<sup>3</sup> gathered below the flight path of T-38 aircraft operated at Mach 1.2 along with extensive simultaneous meteorological data<sup>2</sup> from JAPE-2 are used in this paper. The measured rise time histograms display rise times that are calculated using the definition

$$t_{\text{rise}} = \frac{(0.9p_{\text{max}} - 0.1p_{\text{max}})}{(dp/dt)_{\text{max}}} \quad (1)$$

The calculated rise times from the discrete numerical data approximate the maximum rise times using the points  $t(0.4p_{\text{max}})$ ,  $0.4p_{\text{max}}$  and  $t(0.6p_{\text{max}})$ ,  $0.6p_{\text{max}}$  on the leading edge of the calculated waveforms. This approximation was tested and found to satisfactorily approximate  $(dp/dt)_{\text{max}}$  in Eq. (1) while avoiding numerical problems due to spikes in waveforms.

## I. PROPAGATION MODEL

### A. Scattering formulation

To accommodate anisotropic turbulence, we describe local deviations in index of refraction are described in Cartesian coordinates by the relation

$$\mu(x, y, z, s_x, s_y, s_z) = q_i e^{-(x/s_x)^2 - (y/s_y)^2 - (z/s_z)^2}, \quad (2)$$

where  $q_i$  is the turbule strength, the  $s$  are the characteristic dimensions of the turbule, and the coordinates  $(x, y, z)$  are measured from the center of the  $i$ th turbule. The strength may be either positive or negative, and the  $x$  axis is taken to be the direction of mean wind. As with the spherical scatter-

ers previously used, the anisotropic Cartesian refractive fluctuation profile may be incorporated into the equation of propagation, which is consequently solved for the scattered wave pressure via a first Born approximation. This offers a closed-form expression for the pressure which is convenient for use in computer simulation. The scattering contribution from one turbule (subscript  $i$ ) having characteristic scales  $(s_x, s_y, s_z)$  in Cartesian coordinates may be expressed as

$$\hat{p}_i^B = \frac{q_i \sqrt{\pi}}{2} k^2 \frac{e^{ik(r_{st} + r_{tr})}}{r_{st} r_{tr}} s_{\text{eff}}^3 \exp(-C_{\text{el}} k^2 s_{\text{eff}}^6 / 4), \quad (3)$$

so that the total pressure at the receiver due to propagation through  $N$  turbules is

$$\hat{p}(r) = \hat{p}_0(r) + \sum_{i=1}^N \hat{p}_i^B, \quad (4)$$

where  $r$  is the source–receiver distance and  $\hat{p}_0(r)$  is the complex unperturbed (“direct”) spherical wave pressure amplitude  $e^{ikr}/r$ . The superscript  $B$  denotes “Born,”  $r_{st}$  and  $r_{tr}$  are the source–turbule and turbule–receiver distances (from the center of the turbule), and  $k$  is the wave number. Each scattered wave is modeled as emanating from the center of a turbule (center-based scattering). The complex constant  $s_{\text{eff}}$  behaves somewhat like a composite scale, and depends upon the shape and size of the turbule as well as its position relative to the source and receiver. This term has units of distance, and depends on the wave number  $k$ . The constant  $C_{\text{el}}$  is another complex geometric term that depends on  $k$ , and implicitly contains the scattering angle. Hence the pressure amplitude as well as the phase of the scattered wave depend on the scattering angle, with the largest scattering amplitudes occurring at the smallest scattering angles.

The scattering amplitude tends to increase with turbule size for both spherical and ellipsoidal turbules, while the factors  $C_{\text{el}}$  and  $s_{\text{eff}}$  in (3) can cause significant phase changes for turbules with scales on the order of 20–40 m. It should be noted that use of Eq. (3) also avoids an approximation of the off-center polar angle by the center-based scattering angle used in Refs. 1 and 4. A definition of the constants  $C_{\text{el}}$  and  $s_{\text{eff}}$  may be found in Appendix A, while a full derivation from the wave equation is provided in Ref. 8.

## B. Simulation of the planetary boundary layer using a range of scales

In order to perform a realistic propagation computation, it is necessary to calculate the scattering effects of many randomly placed turbules in a chosen simulation volume. To arrive at an appropriate turbule size distribution, we obtained a number density of scatterers by fitting a von Karman spectrum to measured meteorological (wind) spectra, following the procedure used by de Wolf<sup>9</sup> in his study of electromagnetic wave propagation through permittivity fluctuations. It was shown by de Wolf that if the spectrum  $\Phi$  of fluctuations of the dielectric permittivity is of a von Karman type,

$$\Phi(\mathbf{k}) = \frac{\gamma C_n^2}{(\mathbf{k}^2 + L_0^{-2})^{11/6}} e^{-\mathbf{k}^2 l_m^2}, \quad (5)$$

then a density of scatterers  $n(s)$  can be found to describe the turbulent atmosphere. The microscale  $l_m$  is negligible (1 mm or less),  $\gamma$  is a constant, and the macroscale  $L_0$  and refractive index structure constant  $C_n^2$  ultimately give information about the scales and strengths of turbules needed for atmospheric simulation.

The use of meteorological data from JAPE-2 to obtain  $n(s)$  is facilitated by transforming (5) to a one-dimensional form, instead of dealing with three-dimensional wave vectors  $\mathbf{k}$ . Reducing Eq. (5) to one dimension (corresponding to the direction of mean wind) is suitable as long as the lengths  $(s_x, s_y, s_z)$  are all of roughly the same order of magnitude for any given turbule. Fitting the one-dimensional von Karman spectral form to wind spectra from JAPE-2 gives values of  $C_n^2$  and  $L_0$ , which allow a calculation of the density of scatterers:

$$n(s) = \frac{2^{-5/6} \gamma}{\pi^3 \Gamma(\frac{11}{6})} C_n^2 s^{-10/3} e^{-s^2/2L_0^2}. \quad (6)$$

Only wind fluctuations were included in the spectral fitting because the variance of temperature fluctuations was consistently at least one order of magnitude smaller than the variance of wind speed fluctuations.<sup>2</sup> The function  $n(s)$ , whose behavior is almost entirely dominated by the  $s^{-10/3}$  term, is a measure of the relative number of eddies per size per volume, and thus has units of  $\text{m}^{-4}$ . From Eq. (6) a quasi-random one-dimensional (spherical) turbule size distribution function may be obtained through the use of a Monte Carlo method.

After assigning a quasirandom size to a given turbule based on (6), the turbule is given a random position within an atmospheric simulation volume (explained in the next section). The shape of a turbule may be obtained from its altitude, size, and the character of meteorological conditions as described by the Monin–Obukhov length  $L_{\text{MO}}$ . The reshaping of initially spherical turbules follows from a turbule stretching formulation developed by Wilson *et al.*<sup>7</sup> Atmospheric sciences literature suggests that eddies may be horizontally elongated in the direction of mean wind when shearing mechanisms dominate turbulence production (near the ground), but are vertically elongated when convective conditions dominate (farther from the ground).<sup>5,10,11</sup> The parameter  $L_{\text{MO}}$  is a surface-layer length scale that describes the relative importance of these mechanisms, and is negative under the unstable turbulent conditions represented by the JAPE data used for this work. The dimensionless height  $\zeta = (z/L_{\text{MO}})$  is a ratio of convective (buoyant) to shear (inertial) forces, and facilitates an expression of the importance of these forces with respect to altitude. Wilson *et al.* use a diagonal stretching matrix to model the elongation of turbules in terms of  $\zeta$ , where each element of the matrix represents the stretch in a given direction with respect to the vertical element. Since the wind fluctuations from JAPE were measured in the direction of mean wind, it is sensible to reexpress the turbule shaping as the vertical ( $z$ ) or transverse horizontal ( $y$ ) stretching with respect to the turbulence scale in the  $x$  direction:

$$f_y(\zeta) = \begin{cases} \kappa^2, & (-\zeta) \leq 1, \\ 1, & 1 < (-\zeta) < \kappa_{fc}^3, \\ 1, & \kappa_{fc}^3 \leq (-\zeta) \end{cases} \quad (7)$$

and

$$f_z(\zeta) = \begin{cases} \kappa^2, & (-\zeta) \leq \kappa^3, \\ (-\zeta)^{2/3}, & \kappa^3 < (-\zeta) < \kappa_{fc}^3, \\ \kappa_{fc}^2, & \kappa_{fc}^3 \leq (-\zeta), \end{cases}$$

where the von Karman constant  $\kappa=0.4$  and the free-convection von Karman constant  $\kappa_{fc}$  is roughly equal to 1.2. In order to conveniently account for turbulence isotropy in the inertial subrange, we applied this stretching only to turbules having  $x$ -axis lengths larger than a spectrum-dependent cutoff,  $s_{cut}$ . Looking at (7) and (6) it is evident that a continuous range of scale lengths is now possible, with large turbules near the surface having cigar-like shapes oriented in the direction of mean wind, larger turbules possessing vertically stretched egg-like shapes in the upper majority of the PBL ( $\zeta > 1$ ), and small spherical turbules occurring (more frequently) at any height. The range of sizes is chosen to be limited at the high end by  $L_0$ , while the smallest  $x$  scale allowed is one meter. The minimum size is so chosen because turbules having sizes less than one meter scatter incident waves negligibly, while the maximum turbule scale depends on the fit of measured wind spectra and thus changes with meteorological conditions.

### C. Physical representation of the atmosphere

As mentioned in the previous section, it is necessary to fill a realistic volume with quasirandomly sized, randomly positioned turbules in order to represent the turbulent atmosphere in a physically meaningful manner. In order to realistically model sonic boom propagation from significant altitudes, it is also necessary to address the nonturbulent upper atmosphere that exists above the turbulent PBL.

#### 1. Meteorological parameters and upper atmosphere

The JAPE-2 data used for this paper (as well as Ref. 8) was taken during flybys of T-38 aircraft traveling with speed Mach 1.2 and altitude 10 km, under two different sets (types) of weather conditions occurring on the 19th and 28th of August, 1991.<sup>2,3</sup> In addition to the wind spectra obtained from meteorological data taken on these two days, the height of the PBL was evaluated using the altitude  $z_i$  of the capping temperature inversion, as measured by tethersonde. The two sets of conditions (both corresponding to sunny, windy afternoons in the desert) can be described as moderate turbulence dominated by convective production, and mild turbulence with a significantly thick shear-dominated surface layer, respectively. Since the wind fluctuations dominated those of temperature on these days, the index of refraction fluctuation variance can be described by using only the variance of wind speed in the direction of mean wind:

$$\langle \mu^2 \rangle = \sigma_u^2 / c_0^2. \quad (8)$$

Generally, conditions marked by  $\langle \mu^2 \rangle > 3 \times 10^{-6}$  are regarded as moderately turbulent, while those having smaller

variances are considered mildly turbulent. The moderately turbulent conditions possessed a value of  $\langle \mu^2 \rangle = 9.6 \times 10^{-6}$ , inversion height  $z_i = 1$  km, negligible surface layer ( $L_{MO} = -6$  m), and outer scale length  $L_0 = 300$  m. The mildly turbulent conditions were described by  $\langle \mu^2 \rangle = 1.5 \times 10^{-6}$ , a significant shear/surface layer ( $L_{MO} = -64.8$  m), inversion height  $z_i = 1400$  m, and outer scale  $L_0 = 37$  m.

In the upper atmosphere, nonturbulent effects such as nonlinear steepening, dispersion, and absorption must be addressed. This is accomplished by the use of a version of the program SHOCKN,<sup>12</sup> which incorporates the improved Anderson algorithm<sup>13</sup> to propagate an  $N$  wave from the flight altitude of 10 km to the top of the PBL in each case. The program also takes into account the nonvertical propagation of a supersonic boom by using an appropriately lengthened path according to the Mach angle  $\theta_m = \sin^{-1}(M^{-1}) = 56^\circ$  for  $M = 1.2$ .

#### 2. Geometry and simulation volume

A set of coaxial cylinders were chosen to represent the turbulent boundary layer for each set of conditions, and served as the scattering volume for propagation of the sonic boom from the inversion height to the ground. Because larger turbules have a larger scattering effect than small eddies, they scatter appreciably over a larger angle; consequently, turbules with  $s_x > s_{cut}$  were randomly positioned in the larger cylinder, while turbules with  $s_x < s_{cut}$  were confined to a smaller cylinder of half the radius. This was done to reduce computing time, in light of the larger number density of smaller turbules. The cylinder radius was chosen such that any turbules outside it would not have a detectable effect on the total scattered pressure. After randomly placing  $N$  turbules into the coaxial volumes, the cylinders were skewed  $56^\circ$  from the vertical to account for the direction of incidence of the sonic boom.

Before doing a single-scatter propagation calculation, the strength and number of turbules must be chosen in accord with meteorological data reflecting the weather conditions current to each sonic boom case. Based on the theoretical altitude dependence of the velocity structure parameter  $C_u^2$ , a height dependence for  $q^2$  may be obtained. In the surface layer ( $z < -L_{MO}$ ) and mixing layer (above  $\sim 0.1z_i$ ) of the PBL,  $C_u^2$  is essentially constant, while Chintawongvanich *et al.*<sup>14</sup> show that  $C_u^2 \propto z^{-4/3}$  in the intermediate matching layer, which roughly corresponds to the middle region of Eq. (7). The altitude-dependent behavior of turbule strength can thus be summarized as

$$q_i^2 = \begin{cases} q^2(z_{q0}), & z \leq z_{q0}, \\ q^2(z_{q0})(z/z_{q0})^{-4/3}, & z_{q0} < z < 0.1z_i, \\ q^2(z_{q0})(0.1z_i/z_{q0})^{-4/3}, & 0.1z_i \leq z. \end{cases} \quad (9)$$

The height  $z_{q0}$  was obtained by choosing the smallest altitude at which the data first started displaying height-dependent ( $z^{-4/3}$ ) behavior;  $z_{q0}$  was equal to 16 m for the mildly turbulent data, and 30 m for the moderately turbulent data. It should also be noted that a different  $s_{cut}$  was picked for each day, with a value of 3 m for the mild situation, and 5 m for the moderate case.



The value of both  $N$  and the “ground” strength  $q^2(z_{q0})$  were consequently obtained through an iterative procedure similar to one described in Ref. 1, using values of  $\langle \mu^2 \rangle$  and  $\langle \mu^4 \rangle$  calculated from wind fluctuations measured on each day. A modified relation based on the fraction of a given volume filled with ellipsoidal turbules according to  $n(s)$  was used as a guideline to choose pairs of  $N$  and  $q^2(z_{q0})$  for the previously described simulation volume. Thus  $N$  turbules were randomly positioned in the cylinder of height  $z_i$ , shaped according to altitude, given strengths according to (9), whereupon their contributions to  $\langle \mu^2 \rangle$  and  $\langle \mu^4 \rangle$  were summed and averaged along five horizontal lines in the direction of mean wind placed randomly below  $z_{q0}$ . This was repeated 2000 times for each pair and set of conditions, with the final values chosen for  $N$  and  $q^2(z_{q0})$  being those that produced the best agreement between the numerically modeled and measured refractive moments.<sup>8</sup> Each repetition, or realization, involved differently (randomly) placed turbules, as well as averaging lines. It should be noted that Boulanger *et al.*<sup>1</sup> did not account for the effect of the height dependence of  $q_i$  upon the measured  $\langle \mu^2 \rangle$  and  $\langle \mu^4 \rangle$ .

The subsequent propagation of the output boom from SHOCKN through the anisotropic PBL scattering model was performed on 500 realizations, each for both moderate and mild turbulence, with the horizontal component of the propagation path at angles of 0°, 45°, and 90° to the mean wind, and also for spherical turbules within the improved atmospheric formulation given above. The propagation consisted of summing the complex pressures scattered from each turbule and the pressure due to the direct wave, all in the frequency domain. The resultant frequency component sums are Fourier transformed back to the time domain, giving sonic boom waveforms as modified by turbulence. The rise time of each waveform (500 for each case) was calculated using Eq. (1), and approximating  $(dp/dt)_{\max}$  by the slope from 40% to 60% of the peak overpressure. These rise times were then collected to create rise time distributions that are compared in the next section to distributions from JAPE-2.

## II. RESULTS AND CONCLUSION

Figure 1 compares the histogram of measured rise times (a) to the results of the scattering center calculations for both spherical (b) and ellipsoidal (c) turbules, under conditions of mild turbulence. The conditions are such that the turbulence is mechanically driven, so that moderately sized turbules are stretched in the direction of the wind. The distorted turbules used to produce Fig. 1(c) are aligned along a wind direction that is skewed 45° from the line of flight. Neither spherical (b) nor ellipsoidal (c) turbule results accurately match the measured data; both overpredict the upward shift of rise times. It should be mentioned, however, that the shape of the distribution shown for a 45° azimuthal angle in Fig. 1(c) matches the shape of the measured distribution better than the distributions predicted for the cases of mean wind parallel or perpendicular to the line of flight.

Figure 2 compares the distribution of rise times measured on a moderately turbulent day (a) to histograms showing the results of spherical (b) and ellipsoidal (c) scattering center calculations. The turbulence is convectively driven,

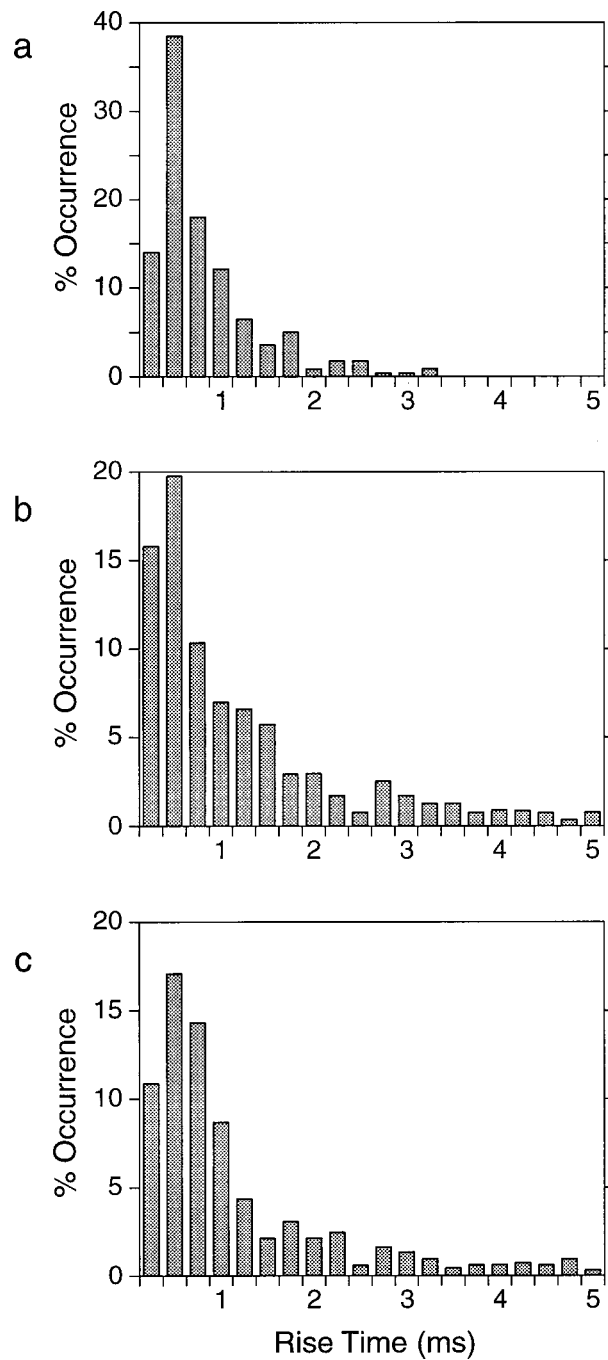


FIG. 1. Rise time distributions for the T-38 aircraft under mild, mechanically driven turbulence: (a) histogram of measured rise times, and predictions using (b) spherical turbules and (c) ellipsoidal turbules.

causing moderately sized turbules to become elongated vertically. Again, the predictions do not agree with the data. In this case, the shift to longer rise times shown in Figs. 2(b) and (c) is underpredicted. Comparing parts (b) and (c) of the first two figures, it can be seen that there is not much difference in the shape of the predicted histograms using spherical and ellipsoidal turbules.

The promising results of Boulanger *et al.*<sup>1</sup> have not been significantly improved by the introduction of nonspherical turbules. Evidently, the single-scatter, scattering center calculation cannot accurately predict the statistics of sonic boom propagation through the turbulent planetary boundary

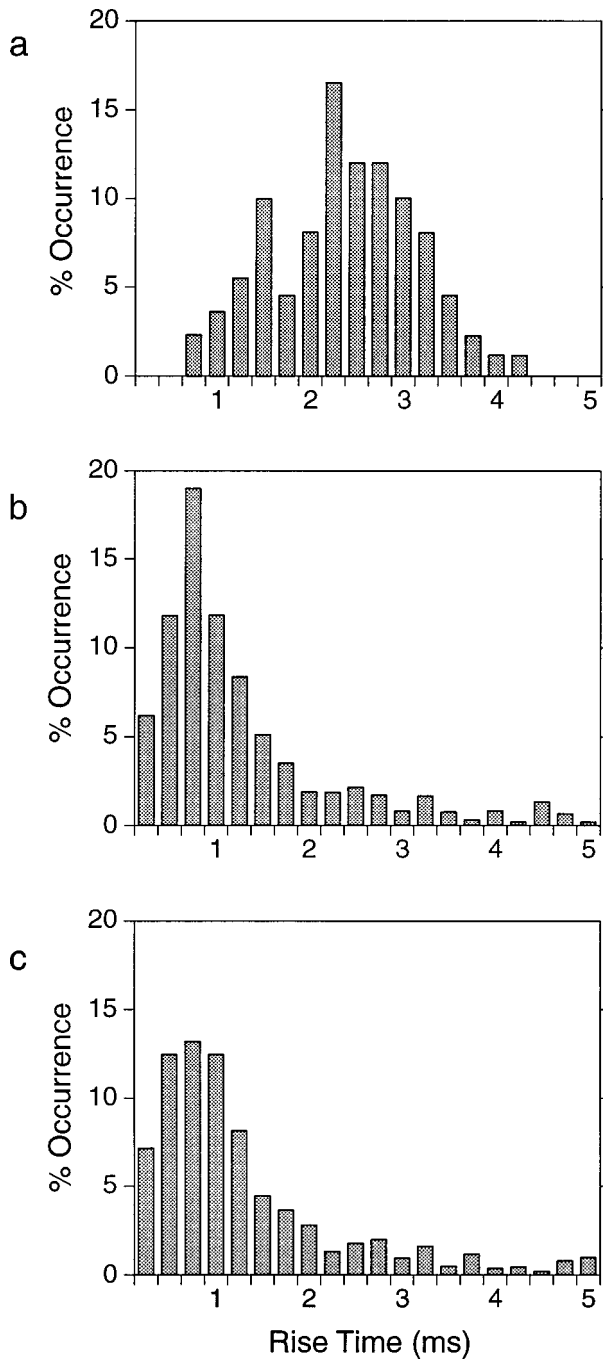


FIG. 2. Rise time distributions of T-38 aircraft under moderate convective turbulence conditions: (a) histogram of measured rise times, and predictions using (b) spherical turbules and (c) ellipsoidal turbules.

layer. The calculation produces histograms which display some features of the measured data, but these predictions are not quantitatively accurate. Higher-order effects such as multiple scatter and changes in the time of flight within turbules may be necessary to model air-to-ground boom propagation.

The single-scatter method did produce accurate results when applied by McBride *et al.*<sup>4</sup> to study the statistics of ground-to-ground propagation of continuous waves. The formulation developed for scattering by ellipsoidal turbules (presented in Appendix A) may thus be useful for future studies of acoustic scattering in the atmosphere, such as the effect of distorted turbules upon scattering for near-ground propagation.

## APPENDIX: DEFINITION OF SCATTERING CONSTANTS $s_{\text{eff}}$ AND $C_{\text{el}}$

In order to obtain a closed-form solution to the inhomogeneous wave equation in Cartesian coordinates, a Taylor expansion is performed on two “geometrical” variables, with retention of terms up to second order in the coordinates  $(x, y, z)$ . There are nine such terms, the coefficients of which are eventually reduced to two useful parameters  $s_{\text{eff}}$  and  $C_{\text{el}}$ , as they appear in Eq. (3). The constant  $C_{\text{el}}$  and effective scale  $s_{\text{eff}}$  are thus defined as

$$-C_{\text{el}} = k^2(2DEHI + 2DFGI + 2EFGH - D^2I^2 - E^2H^2 - F^2G^2) + 2ik(DGH\gamma^2 + EGI\beta^2 + FHI\alpha^2) - (G\beta\gamma)^2 - (H\alpha\gamma)^2 - (I\alpha\beta)^2 \quad (\text{A1})$$

and

$$s_{\text{eff}}^{-3} = \sqrt{(\alpha^2\beta^2\gamma^2 + k^2(D^2\gamma^2 + E^2\beta^2 + F^2\alpha^2 - 2ikDEF))}, \quad (\text{A2})$$

with

$$\alpha = \sqrt{s_x^{-2} - ikA^2}, \quad \beta = \sqrt{s_y^{-2} - ikB^2}, \quad \gamma = \sqrt{s_z^{-2} - ikC^2}, \quad (\text{A3})$$

where the geometric coefficients are given by

$$A^2 = \frac{Y_s^2 + Z_s^2}{2r_{\text{st}}^3} + \frac{Y_r^2 + Z_r^2}{2r_{\text{tr}}^3}, \quad B^2 = \frac{X_s^2 + Z_s^2}{2r_{\text{st}}^3} + \frac{X_r^2 + Z_r^2}{2r_{\text{tr}}^3},$$

$$C^2 = \frac{X_s^2 + Y_s^2}{2r_{\text{st}}^3} + \frac{X_r^2 + Y_r^2}{2r_{\text{tr}}^3}, \quad D = \frac{X_s + Y_s}{2r_{\text{st}}^3} + \frac{X_r + Y_r}{2r_{\text{tr}}^3},$$

$$E = \frac{X_s + Z_s}{2r_{\text{st}}^3} + \frac{X_r + Z_r}{2r_{\text{tr}}^3}, \quad F = \frac{Y_s + Z_s}{2r_{\text{st}}^3} + \frac{Y_r + Z_r}{2r_{\text{tr}}^3},$$

$$G = \frac{X_s}{r_{\text{st}}} + \frac{X_r}{r_{\text{tr}}}, \quad H = \frac{Y_s}{r_{\text{st}}} + \frac{Y_r}{r_{\text{tr}}}, \quad \text{and} \quad I = \frac{Z_s}{r_{\text{st}}} + \frac{Z_r}{r_{\text{tr}}}.$$

The  $(X_s, Y_s, Z_s)$  and  $(X_r, Y_r, Z_r)$  are the source and receiver positions with respect to the center of a given turbule, while  $r_{\text{st}}$  and  $r_{\text{tr}}$  are the source–turbule and turbule–receiver distances.

<sup>1</sup>P. Boulanger, R. Raspet, and H. E. Bass, “Sonic boom propagation through a realistic turbulent atmosphere,” *J. Acoust. Soc. Am.* **98**, 3412–3417 (1995).

<sup>2</sup>B. W. Kennedy, R. O. Olsen, J. R. Fox, and G. M. Mitchler, *Joint Acoustic Propagation Experiment Project Summary*, The Bionetics Corporation, Las Cruces, NM, September 1991.

<sup>3</sup>W. L. Willshire, Jr., D. P. Garber, and D. W. deVilbiss, “The effect of turbulence on the propagation of sonic booms,” *Proceedings of the 5th International Symposium on Long Range Sound Propagation*, The Open University, Milton Keynes, England, May 1992.

<sup>4</sup>W. E. McBride, H. E. Bass, R. Raspet, and K. E. Gilbert, “Scattering of sound by atmospheric turbulence: A numerical simulation above a com-

- plex impedance boundary,” *J. Acoust. Soc. Am.* **90**, 3314–3325 (1991).
- <sup>5</sup>G. J. Boer and T. G. Shepherd, “Large-scale two-dimensional turbulence in the atmosphere,” *J. Atmos. Sci.* **40**, 164 (1983).
- <sup>6</sup>W. Wilken, “Experimental study of the influence of varying atmospheric conditions on sound propagation close to the ground,” *Acustica* **62**, 55–65 (1986).
- <sup>7</sup>D. K. Wilson and D. W. Thomson, “Acoustic propagation through anisotropic, surface-layer turbulence,” *J. Acoust. Soc. Am.* **96**, 1080–1095 (1994).
- <sup>8</sup>M. Kelly, “Scattering of sonic booms by anisotropic turbulence in the atmosphere,” M. S. thesis, Graduate Program in Physics, The University of Mississippi, 1997.
- <sup>9</sup>D. A. de Wolf, “A random-motion model of fluctuations in a nearly transparent medium,” *Radio Sci.* **18**, 138–142 (1975).
- <sup>10</sup>H. E. Cramer, “Measurements of turbulence structure near the ground within the frequency range from 0.5 to 0.1 cycles/sec,” *Adv. Geophys.* **6**, 75–96 (1959).
- <sup>11</sup>H. A. Panofsky, “Scale analysis of atmospheric turbulence at 2 m,” *Q. J. R. Meteorol. Soc.* **88**, 57–69 (1962).
- <sup>12</sup>R. O. Cleveland, J. P. Chambers, H. E. Bass, R. Raspet, D. T. Blackstock, and M. F. Hamilton, “Comparison of computer codes for the propagation of sonic boom waveforms through isothermal atmospheres,” *J. Acoust. Soc. Am.* **100**, 3017–3027 (1996).
- <sup>13</sup>M. O. Anderson, “The propagation of a spherical  $N$  wave in an absorbing medium and its diffraction by a circular aperture,” *Appl. Res. Lab. Tech. Rep. No. 77-5, ARL-TR-74-25, AD 787878 Applied Research Laboratories, The University of Texas, Austin, TX, August 1974.*
- <sup>14</sup>P. Chintawongvanich and R. Olsen, “Experimental study of temperature structure parameter in the surface layer,” preprint volume of the *7th International Symposium on Meteorological Observations and Instrumentation, Laser Atmospheric Studies*, 1991.

# The interaction of a laser-generated cavity with a solid boundary

S. J. Shaw

*Department of AAETS, Loughborough University, Loughborough LE11 3TU, United Kingdom*

W. P. Schiffers, T. P. Gentry, and D. C. Emmony

*Department of Physics, Loughborough University, Loughborough LE11 3TU, United Kingdom*

(Received 8 August 1998; accepted for publication 7 March 2000)

In this paper new observations of a laser-generated cavitation bubble interacting with an inertial boundary are presented. Employing schlieren photography techniques and a thin film transducer placed on the surface of the boundary, the pressure stresses induced in the solid boundary and the surrounding fluid by collapsing bubbles, created very close to the solid surface, are experimentally measured. Liquid jet development, shock wave emission, and “splash” phenomena are identified. For different creation sites close to the boundary, the relevance of each of these phenomena with respect to potentially damaging pressure stresses in the boundary is speculated on. © 2000 Acoustical Society of America. [S0001-4966(00)03606-7]

PACS numbers: 43.25.Yw, 43.30.Yj, 43.35.Ei [MAB]

## INTRODUCTION AND MOTIVATION

In this paper we are interested in the phenomenon of cavitation and the possible mechanisms which may lead to the observed associated damage. Over the last 100 years a large amount of work has been published on possible mechanisms explaining how cavitation damage occurs. Reviews of this work can be found in Plesset and Prosperetti,<sup>1</sup> Mørch,<sup>2</sup> Blake and Gibson,<sup>3</sup> and Leighton.<sup>4</sup> At the present time, both experimental and theoretical work confirms that when a bubble collapses near a solid boundary, provided it is sufficiently close, the boundary destroys the symmetry of the collapse process. The result is that a liquid jet is formed which threads the bubble and impacts on the boundary.<sup>5-14</sup> Many workers have attributed the erosion of solid boundaries to this liquid jet impact or the “shock” waves that are emitted when the bubble achieves minimum volume. Indeed Shima *et al.*<sup>15</sup> provided ranges where they thought each or both of these effects is important.

In a recent paper<sup>16</sup> using schlieren photography, photoelasticity, and a thin film transducer placed on a solid boundary, we recorded the pressure stresses at the surface of the boundary of a laser-generated bubble in water, particularly during the first collapse. For a range of values of stand-off parameter  $\gamma$  (the ratio of the distance the bubble is created from the boundary to the maximum bubble radius), we studied the importance of the liquid “jet” with respect to possible damage processes. It is found that for  $\gamma = \frac{5}{6}$  and  $\gamma = 1$ , the pressure transducer records a double peak structure. For  $\gamma = \frac{5}{6}$ , the pressure is observed to rise to a maximum over approximately 20  $\mu\text{s}$  after the initial jet impact, whose magnitude as measured by the transducer is surprisingly small. Superimposed upon this pressure is a sharp spike with an upward gradient, similar to the laser initial breakdown. The smallness of the initial jet diameter (0.2 mm) relative to the circular transducer diameter (3.0 mm) could partially explain the relative smallness of the transducer signal. We interpreted the sharp peak due to the passage of shock waves,

emitted when the bubble achieves minimum volume, while the longer duration pressure was attributed to fluid flow effects through the now toroidal shaped cavity. For larger values of  $\gamma$  ( $\frac{7}{6}, \frac{4}{3}$ ), the transducer output appeared to be dominated by acoustic transient signals.

Tong *et al.*<sup>17</sup> conducted numerical simulations to provide an interpretation of experimental results presented for three values of the stand-off parameter  $\gamma = 0.85, 0.92,$  and  $1.2$ . The transducer results for  $\gamma = 0.85$  and  $0.92$  indicated a double peak structure of a sharp spike superimposed on a slower, but longer lasting pressure profile. Assuming an inviscid, incompressible model for the fluid and a simple adiabatic model for the variation of the cavity pressure, the boundary integral method was employed to calculate bubble shapes and induced pressure stresses. A new phenomenon was identified, which was termed a “splash” effect and it is found to precede the liquid jet impact on the boundary. The “splash” effect is only valid for a small stand-off parameter range and is dependent on whether there is a liquid layer between the bubble surface closest to the boundary, and the boundary itself, during the collapse phase and the thinness of this layer. For  $\gamma = 0.85$  and  $0.92$  it is predicted that when the liquid jet threads the bubble and impacts into this layer, the closeness of the boundary results in a radial flow away from the jet axis. This flow interacts with the flow setup by the contracting bubble and a “splash” is projected away from the thin liquid layer, in a direction opposite to the liquid jet motion. Comparing the numerical simulations with the experimentally deduced results, it was concluded that in the case of  $\gamma = 0.85$  and  $0.92$ , the first pressure peak is indeed due to shock wave emission, resulting from this “splash” effect subsequently impacting violently on the bubble, and is not due to the bubble reaching minimum volume. The second peak is attributed to the resulting flow through the toroidal bubble. For stand-off parameters of  $\gamma = 1.2$  and larger, the liquid layer is shown to be too thick and the full devel-

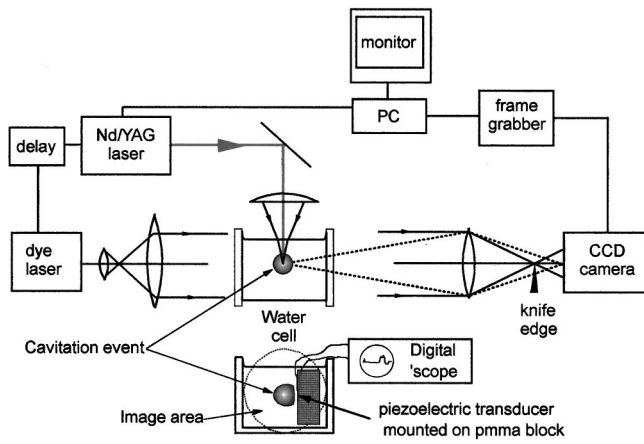


FIG. 1. Experimental setup.

opment of “splash” effects is prevented for the case of the first collapse cycle.

Another recently published paper with relevance to this work is the experimental study of Philipp and Lauterborn.<sup>18</sup> They recorded the interaction of a laser-generated bubble with a solid boundary for a range of values of the stand-off parameter using high-speed photography with framing rates of up to one million frames/s, and investigated the resultant surface damage on a flat metal specimen (the boundary). They concluded through their study that the liquid jet impact on the solid boundary is only important with respect to potential damage for the stand-off parameter range  $\gamma > 0.7$ , based on predicting water hammer pressures.

In the present work we investigate experimentally the collapse of laser-generated cavities near boundaries for two values of a stand-off parameter less than 0.7. The aim of the study is to determine the pressure stresses induced on the solid boundary (and the surrounding fluid) and how these can be interpreted with respect to the newly postulated “splash” phenomena. We want to deduce the lower stand-off parameter limit for the “splash” effect and determine if there is any other information we can glean about erosion processes. We also note that the numerous uses of the laser in medical applications is making it increasingly important to understand laser-generated cavities, although whether bubbles are created mechanically, acoustically, or optically it is felt that the general mechanisms behind the resultant damage will be the same.

## I. EXPERIMENTAL SETUP

The experimental setup is displayed in Fig. 1 and can be divided into main components. The first part comprises a *Q*-switched Nd:Yag laser operating at  $1.06 \mu\text{m}$  which was used to generate the cavitation bubbles. The laser output was about 100 mJ, which is sufficient to easily obtain breakdown in both air and water, and had a pulse width of 20–30 ns. The laser beam was expanded using a 25-mm plano-concave focal length lens and then focused by a 28-mm focal length,  $f/2.0$  camera lens. In the experiments, the cavitation interaction events were studied both optically and through the use of pressure sensitive transducers. The optical observations of the transient events were conducted by the second and third

parts of the experimental setup, which is the high-speed image recording system. The ultrafast light source for the schlieren photography of the bubble evolutions was a nitrogen pumped dye laser with an output energy of approximately  $15 \mu\text{J}$  and pulse width of 500 ps. The dye employed was green Coumarin 500 which can be tuned to emit visible radiation at wavelengths centered around 514 nm and having a coherence length of about  $20 \mu\text{m}$ . The light emitted by the dye laser was spatially filtered by a  $\times 40$  microscope objective and a  $25\text{-}\mu\text{m}$  pinhole. It was then passed through a 50-mm focal length camera lens producing a parallel beam of light to illuminate the cavity events.

The resultant images were formed on the sensitive area of a charge coupled device (CCD) camera by a 135-mm focal length camera lens and subsequently recorded by a video-frame-grabber unit. To obtain Schlieren images a knife-edge was placed between the camera lens and the CCD camera, positioned half way in the waist of the beam. The analog video signal from the camera was digitized and stored in a  $512 \times 512$  pixel array with 64 gray levels. The frame-grabbed images were in turn passed to the computer and displayed as analog video signals on a monitor. Hard copy output was achieved by transferring the pictures to a personal computer, where they were processed using a software package and subsequently printed. For each shot of the Nd:Yag laser only one image was recorded, the time of which was determined by manually varying the delay between the Nd and dye laser pulses. The good reproducibility of the laser breakdown process in water ensures that a sequence of images of the evolution of a cavitation bubble created close to a solid boundary was obtained over a time span of hundreds of microseconds, but with a temporal resolution in each image of the order of a nanosecond.

The test area was a transparent glass cell with the dimensions  $5 \times 5 \times 5 \text{ cm}$ , filled to the brim with deionized water to provide a flat surface. The solid boundary employed was a polymethylmethacrylate (perspex) block immersed in the water. To measure the pressure development induced in the solid boundary by the oscillations of the laser-generated cavity, a pressure transducer consisting of a polyvinylidene fluoride (pvdf) film ( $20\text{-}\mu\text{m}$  thick) was bonded to the surface. The active area of the transducer had a diameter of 3 mm. The output signal from the transducer was displayed as a voltage, which was proportional to pressure, trace on a digital storage oscilloscope.

The experimental setup allows for the simultaneous recording of the optical interaction between the bubble and boundary, and the resultant induced pressure stresses in the solid.

## II. EXPERIMENTAL RESULTS AND DISCUSSION

In this section we present the experimentally determined results consisting of transducer pressure (voltage) traces and corresponding Schlieren images displaying bubble shapes of the interaction of a laser-generated cavity in water with a solid boundary for small values of the stand-off parameter ( $\gamma$ ). Two sets of results are presented corresponding to  $\gamma$  values of 0.65 and 0.54.

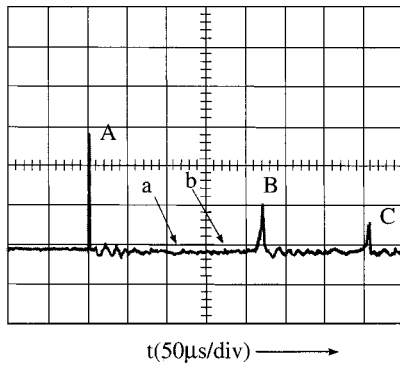


FIG. 2. Transducer plot at low resolution for  $\gamma=0.65$ .

Figure 2 displays a series of pressure spikes at low temporal resolution for a cavitation bubble interacting with a solid boundary created a distance,  $s$ , of 0.71 mm from the solid wall with a maximum bubble radius,  $R_{\max}$ , of 1.09 mm (i.e.,  $\gamma=0.65$ ). In the horizontal direction one division represents  $50 \mu\text{s}$ . Spike A corresponds to the pressure signal due to the initial laser breakdown acoustic transient, while spikes B and C correspond to pressure signals around the first and second minimum volumes of the bubble, respectively. For the purposes of the present study we are interested in spike B only, as it is felt that for the chosen stand-off parameter values, this region is the one which will have most relevance with respect to potential damage processes. We also noted above that this work is a follow-up to the work of Shaw *et al.*<sup>16</sup> and Tong *et al.*,<sup>17</sup> both of whom concentrate on this time interval of the bubble oscillation cycle, but for larger values of stand-off parameter.

To investigate this region in detail, the pressure profile in Fig. 3, which is a magnification of spike B in Fig. 2, will be considered in conjunction with a sequence of seven schlieren images, displayed in Fig. 4. Note that in Fig. 3 in the horizontal direction one division now represents  $5 \mu\text{s}$ , while in the vertical direction the signal is magnified five times. The transducer output illustrates that around the first minimum volume, for this chosen value of stand-off parameter, the transducer is subjected to several distinct pressure stresses which are different in intensity and duration. The main pressure signal rises to a maximum over approximately  $15 \mu\text{s}$ , and then falls away rapidly over a time span of 3–4  $\mu\text{s}$ . Therefore it is evident that the transducer surface feels a

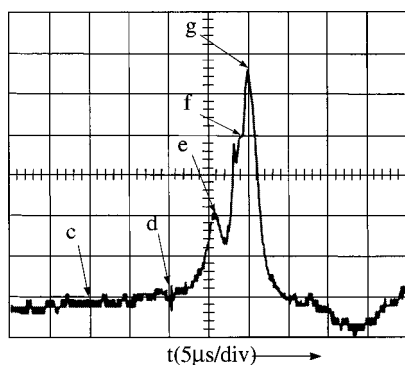


FIG. 3. Transducer plot of the highly resolved pressure peak due to the first collapse of the cavitation bubble for  $\gamma=0.65$ .

“push” pressure over a total time span of about 18–19  $\mu\text{s}$ . Superimposed upon this pressure development are a series of faster spikes. The first spike (label e) is quite distinct (starting at about 215  $\mu\text{s}$  after initial breakdown), having a relatively sharp rise time and a peak which lasts approximately 1  $\mu\text{s}$ . About 2  $\mu\text{s}$  later, a second spike is observed, although it is harder to distinguish, on the sharp rising slope of the longer pressure pulse. A third spike (f) is just visible a further 1  $\mu\text{s}$  later. After the “push” or positive pressure, the transducer records a negative pressure (or “pull”) starting at approximately 220  $\mu\text{s}$  after initial breakdown, and lasting for about 12  $\mu\text{s}$ .

To try and understand how the laser-generated cavitation bubble induces this form of pressure profile on the transducer we now consider the schlieren sequence in Fig. 4. Each of the pictures was photographed individually, i.e., each schlieren image represents a different cavitation event. The delay between firing the Nd:Yag laser and the illuminating dye laser was varied, yielding the set of photographs showing the bubble between maximum radius and a few microseconds after minimum volume. Each time the Nd:Yag laser was fired, the resultant pressure signal, as recorded by the transducer (Fig. 2), was observed. It is found that despite every single laser shot generating a different cavitation event, by ensuring the transducer records are the same each time, then the good experimental reproducibility of the bubble collapse process is confirmed. Hence the information provided by the schlieren images can be regarded as sequential. Remembering this, the labels (a)–(g) in Figs. 2 and 3 correspond to the individual photographs in Fig. 4. The switching of the nitrogen laser was recorded as an electromagnetic interference (EMI) pulse on the transducer trace. However, the actual time each cavitation event is illuminated by the dye laser is 600 ns after the nitrogen laser was triggered due to the internal delay.

At maximum radius [Fig. 4(a)] the bubble surface farthest from the solid boundary is spherical while the opposite surface is flat along a length of 1.2 mm. It is noted, however, that the bubble is still detached from the solid wall as indicated by the gray region between the bubble and the boundary. The white dot in the middle of the bubble is the laser-generated self-illuminating plasma caused by the optical breakdown process and indicates the initial geometrical center of the bubble. The plasma exposes the CCD camera at breakdown and is therefore visible on every image even though each picture may be exposed by the dye laser up to hundreds of microseconds later. As the bubble collapses, restrictions placed on the right-hand side of the bubble due to the proximity of the solid boundary result in the left-hand bubble surface moving much faster toward the bubble center. Consequently the collapse process follows the “classical” asymmetric route to liquid “jet” development [Fig. 4(b)–(d)] as has been observed by many authors<sup>5–14</sup> for a range of  $\gamma$  values. In Fig. 4(c) the surface farthest from the boundary has started to flatten, indicating the formation of the jet, which will then thread the bubble, impacting upon the boundary. Because of the illumination technique employed in the experimental setup (a parallel beam of light rather than diffuse scattered light), it is not possible to observe the jet

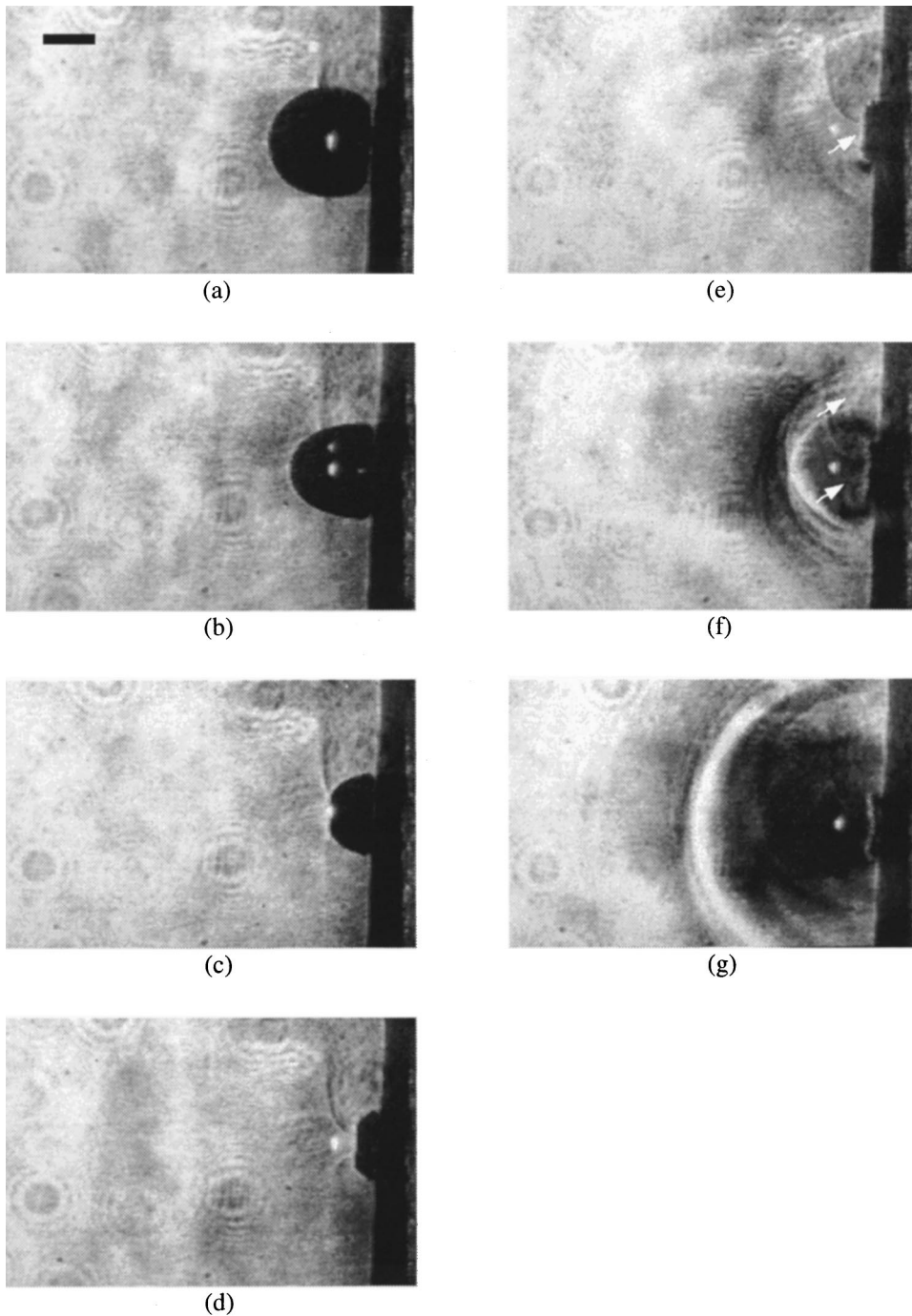


FIG. 4. Schlieren images of an oscillating cavitation bubble in front of a solid boundary for  $\gamma=0.65$ , at times  $t \cdot$  (a) 110, (b) 170, (c) 200, (d) 210, (e) 216, (f) 219, (g) 220  $\mu\text{s}$ . The horizontal bar in image (a) measures 1 mm.

actually passing through the bubble as other research groups<sup>7,9-11,13</sup> have. However, the shapes of the bubble combined with the transducer results give a clear indication of jet development. Examining the transducer plot (Fig. 3), the jet probably impacts the boundary 4–5  $\mu\text{s}$  before Fig. 4(c). In Fig. 4(d), the bubble has the shape of a trapezium sitting on another trapezium. The gray region which was sandwiched between the bubble and boundary is no longer visible. However, a separate line is visible which enters images 4(c) and (d) from above the bubble, and is bent toward the collapsing bubble and in line with the laser-generated plasma. This line is caused by absorption effects along the original laser path used to create the bubble. These absorption effects, which are also just visible in previous schlieren pictures, can be

regarded as an experimental streamline, i.e., it is a method by which the movement of the fluid can be followed nonobtrusively, experimentally. The mechanism behind this effect is described in more detail in Schiffers *et al.*<sup>19</sup> In Fig. 4(d) the marked laser path indicates fluid flow through the toroidal shaped bubble toward the solid wall.

A couple of microseconds after photograph 4(d) was taken, the transducer indicates a noticeable rise in pressure. Superimposed upon this pressure buildup is the first spike region. Photograph 4(e) which was taken about 1  $\mu\text{s}$  after this spike region first peaked, shows the bubble as it continues to collapse. The bubble is now almost rectangular in cross section, elongated along a line parallel to the solid boundary. The aforementioned gray/white line which

“marks” the fluid flow is still visible above the laser-induced plasma and is still bent toward the collapsing bubble. Two faint shock fronts are just visible at the top of the picture (the lower front being marked by the small arrows). Since these acoustic transient fronts can be traced back to emission sites on the bubble volume and it is evident that they are emitted less than  $2 \mu\text{s}$  before photograph 4(e) was taken, then it can be concluded that the first spike region is caused by the passage of these shock waves through the transducer.

As we have noted above in the Introduction, in Shaw *et al.*<sup>16</sup> for the larger stand-off parameters of  $\frac{5}{6}$  and 1, a double peak pressure profile was recorded consisting of a sharp spike superimposed on a longer lasting pressure stress. Because the rise time of the pressure spike was similar to the initial breakdown we concluded that this transducer signal was due to the passage of shock fronts identified in the corresponding schlieren images. We postulated that the shock wave emission was a result of the bubble achieving minimum volume. The subsequent work of Tong *et al.*<sup>17</sup> suggests that these shock wave emissions are not due to minimum volume being reached but are in actual fact due to the “splash” impacting violently with the contracting bubble. Indeed in both the work of Shaw *et al.*<sup>16</sup> and Tong *et al.*,<sup>17</sup> the schlieren images indicate that the bubble has smaller volumes after the first spike pressure emission, i.e., the bubble is continuing to collapse. For the stand-off parameter value of  $\gamma=0.92$ , the numerical model predicts an impact between the “splash” and the bubble surface on a ring around the “jet base.” The numerical model used is limited to incompressible flow and therefore gives no information about shock wave emission. However, by careful comparison between the theoretical and experimentally determined bubble shapes and by tracing the emission sites of the shock fronts in the schlieren pictures to the surface farthest from the boundary of the still collapsing bubbles, it was concluded that this impact must be the cause of the acoustic emissions. For the smaller stand-off parameter value of  $\gamma=0.81$ , it was again postulated that the pressure spike was due to the splash impacting on the bubble surface, although the impact in this case was predicted to be on the bubble sides. For the results presented here for  $\gamma=0.65$ , it is felt that the first spike is still due to splash impact effects. Indeed Fig. 4(a)–(c) indicate a thin film of liquid between the bubble and boundary, although the thickness does decrease, before the jet completely threads the cavity, which is deemed essential by Tong *et al.*<sup>17</sup> for the development of splash effects. The bubble shape along the boundary in Fig. 4(d) is also consistent with the flow patterns predicted in Tong *et al.*<sup>17</sup> for larger stand-off parameter values. Last, the bubble continues to collapse as indicated by Fig. 4(f). Tracing back the shock front marked by the arrows in Fig. 4(e), the emission site appears to be on the upper bubble surface near the boundary, i.e., the splash impact is on the bubble sides for this  $\gamma$  value, consistent with the observations of Tong *et al.*<sup>17</sup>

Below the rectangular shaped cavitation bubble a dark area is visible which is marked by a separate arrow.

Comparing Figs. 3 and 4(f), it is felt that the second and third spikes are associated with effects around minimum vol-

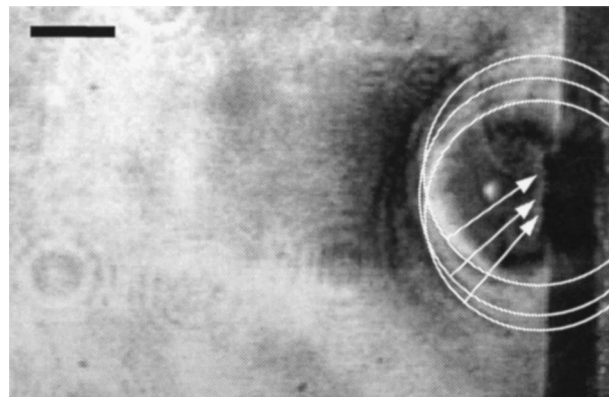


FIG. 5. Detailed consideration of Fig. 4(f). The arrows indicate emission centers of the shock waves. The horizontal bar in the image measures 1 mm.

ume, i.e., the bubble achieves minimum volume at approximately  $218 \mu\text{s}$  after initial breakdown. Several shock fronts are visible to the left of the bubble. Three of these fronts are traced in Fig. 5. A brief calculation shows that they compare to the individual spikes on the transducer plot  $1 \mu\text{s}$  before. A question we ask ourselves is why the minimum volume shock wave effects are not particularly prominent on the transducer trace, i.e., emitted shock waves are visible in the schlieren pictures, but why are they weak? Indeed examining Fig. 2 the pressure spike (A) emitted at initial breakdown completely dominates the spikes associated with the splash impact and minimum volume (B). We speculate that the result of the splash impact and the subsequent flow setup must impede the collapse process, reducing the compression at minimum volume, which in turn will give rise to weak shock wave emission. For this stand-off parameter value, the liquid film between the bubble and boundary is thinner, and therefore the resultant splash is expected to be less energetic, but it still occurs and consequently impedes the bubble collapse process, resulting in weak shock wave emission at minimum volume. The splash effect probably also disrupts the symmetry of the collapse as indicated by the multiple emissions.

Also visible in Figs. 4(f) and 5 are three dark regions. The dark regions above and below the bubble (marked by small arrows) seem to be a development of the dark region effect seen in Fig. 4(e), now bigger and distinctly curved. The dark regions are distinguished by the schlieren photography, as these regions have a different density to the surrounding fluid. Density changes occur either due to pressure or temperature changes. We speculate that because of their position, i.e., the top and bottom of the bubble, that the curved dark regions are associated with the splash impact and the resultant fluid flow.

The last photograph taken in this sequence [Fig. 4(g)] coincides with the peak of the slow pressure development as recorded by the transducer. As this photograph was only taken just over  $1 \mu\text{s}$  after Fig. 4(f), shock fronts can still be seen moving toward the left-hand side of the photograph. Also visible in this image is a large dark area between a large shock front on the left and the reexpanding cavitation bubble. We conclude that this larger region has now undergone a density change, and is therefore made visible by the schlieren imaging, in comparison to the surrounding fluid for



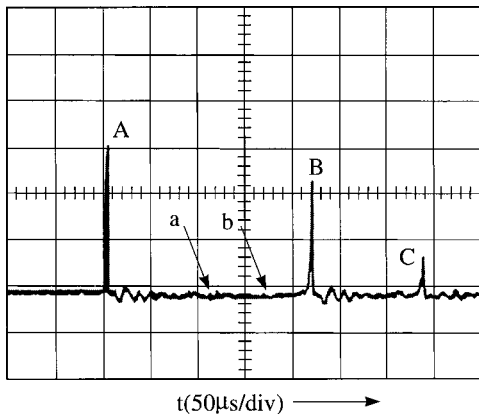


FIG. 6. Transducer plot at low resolution for  $\gamma=0.54$ .

this time in the bubble evolution process. The bubble volume is also larger in this picture, indicating that reexpansion has begun. The slow pressure development is associated with fluid flow effects which are greatly increased by the splash impact.

Figure 6 displays the transducer output at low temporal resolution for a time span which includes the first two oscillation cycles for a laser-generated cavitation bubble with the same maximum bubble radius as before ( $R_{\max}=1.09$  mm) but positioned 0.12 mm closer to the solid boundary (i.e.,  $\gamma=0.54$ ). The initial breakdown pressure spike (A) still has the largest amplitude, but the pressure spike (B) which is associated with effects around the first minimum volume is significantly larger. The maximum stress here causes a pressure increase approximately twice that in the larger stand-off parameter case ( $\gamma=0.65$ ). The time expanded transducer output around spike (B) is displayed in Fig. 7. The units on the abscissa are the same as in Fig. 3, while on the ordinate the signal is magnified two times with respect to Fig. 6. On comparing this pressure plot with Fig. 3, it is now found that one pressure spike dominates the results, although a longer lasting but much smaller amplitude pressure is recorded to the left of the spike. In all, it is estimated that the transducer records a positive pressure for approximately 17  $\mu\text{s}$ .

Again, to understand this pressure development, we present seven Schlieren pictures (Fig. 8). The first two im-

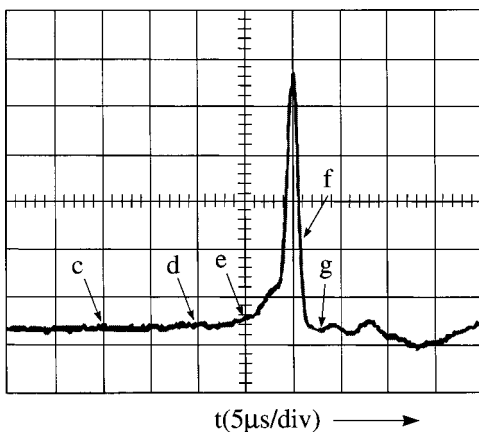


FIG. 7. Transducer plot of the highly resolved pressure peak due to the first collapse of the cavitation bubble for  $\gamma=0.54$ .

ages are taken at times prior to Fig. 7 (see Fig. 6), while the others correspond to the times as marked on the transducer plot. In Fig. 8(a), it is unclear whether there is a liquid layer between the right-hand bubble surface and the solid boundary, although in Fig. 8(b) the bubble does appear to be attached to the boundary. As the bubble collapses it assumes similar shapes to those observed in Fig. 4, except at the top and bottom of the bubble, nearest the boundary. In this case, these parts of the bubble appear to be distorted farther away from the bubble center along the boundary. This effect has also been observed by Benjamin and Ellis,<sup>7</sup> Schiffers *et al.*,<sup>19</sup> and Philipp and Lauterborn.<sup>18</sup> Comparing Fig. 8(d) with the transducer plot, it appears that this image was taken approximately 5  $\mu\text{s}$  after the jet impacts the boundary, although the pressure increase as recorded by the transducer initially is very slight. The transducer indicates a steeper pressure rise beginning around the time photograph 8(e) was taken. In the top half of this image the marked laser path is again visible, indicating flow through the bubble. Approximately 4  $\mu\text{s}$  later, the transducer plot rises considerably to a peak. Figure 8(f) shows the bubble 1–2  $\mu\text{s}$  after this peak. A collection of shock waves can be seen to the left of the bubble, although due to their closeness to one another, it is difficult to distinguish them. It is concluded from these observations that the main pressure spike is due to this multiple shock wave emission.

For this value of stand-off parameter, it is felt that the shock wave emission is solely due to the bubble achieving minimum volume. The schlieren pictures seem to indicate that when the liquid jet threads the bubble there is no liquid layer between the bubble surface and the boundary, or if a liquid film exists, then it is too thin to interfere with any resultant flows and no splash phenomena are observed. For  $\gamma=0.5/0.6$ , Philipp and Lauterborn<sup>18</sup> suggested that the noticeable elongation of the bubble along the boundary during the collapse phase results in increasingly larger jet diameters. Through this mechanism they predicted that the inner diameter of the toroidal bubble becomes larger while the outer surface remains relatively fixed, i.e., the bubble is now in effect collapsing from within. They assumed that the liquid jet sets up a vortex flow, which maintains the symmetry of the bubble volume compression, resulting in a violent collapse. From their pictures, they predicted that the bubble torus then disintegrates into many tiny bubbles along a ring, which collapse almost simultaneously, each emitting a shock wave. From our observations, it appears that since splash effects no longer appear to occur (or at least be significant), there is nothing to disrupt the symmetry of the proposed vortex flow of Philipp and Lauterborn<sup>18</sup> (or any other form of flow that might be set up) and consequently we observe a relatively “clean” minimum volume shock wave emission as they suggested. With regard to Philipp and Lauterborn’s assertion that many tiny bubbles are formed after minimum volume is achieved, it is unclear from the transducer plot whether this is the case or not, although a series of shock fronts are observed in the Schlieren Fig. 8(f), as opposed to just one.

In Fig. 8(f), between the shock fronts and the solid boundary, a dark region is also observed. Again, this indi-

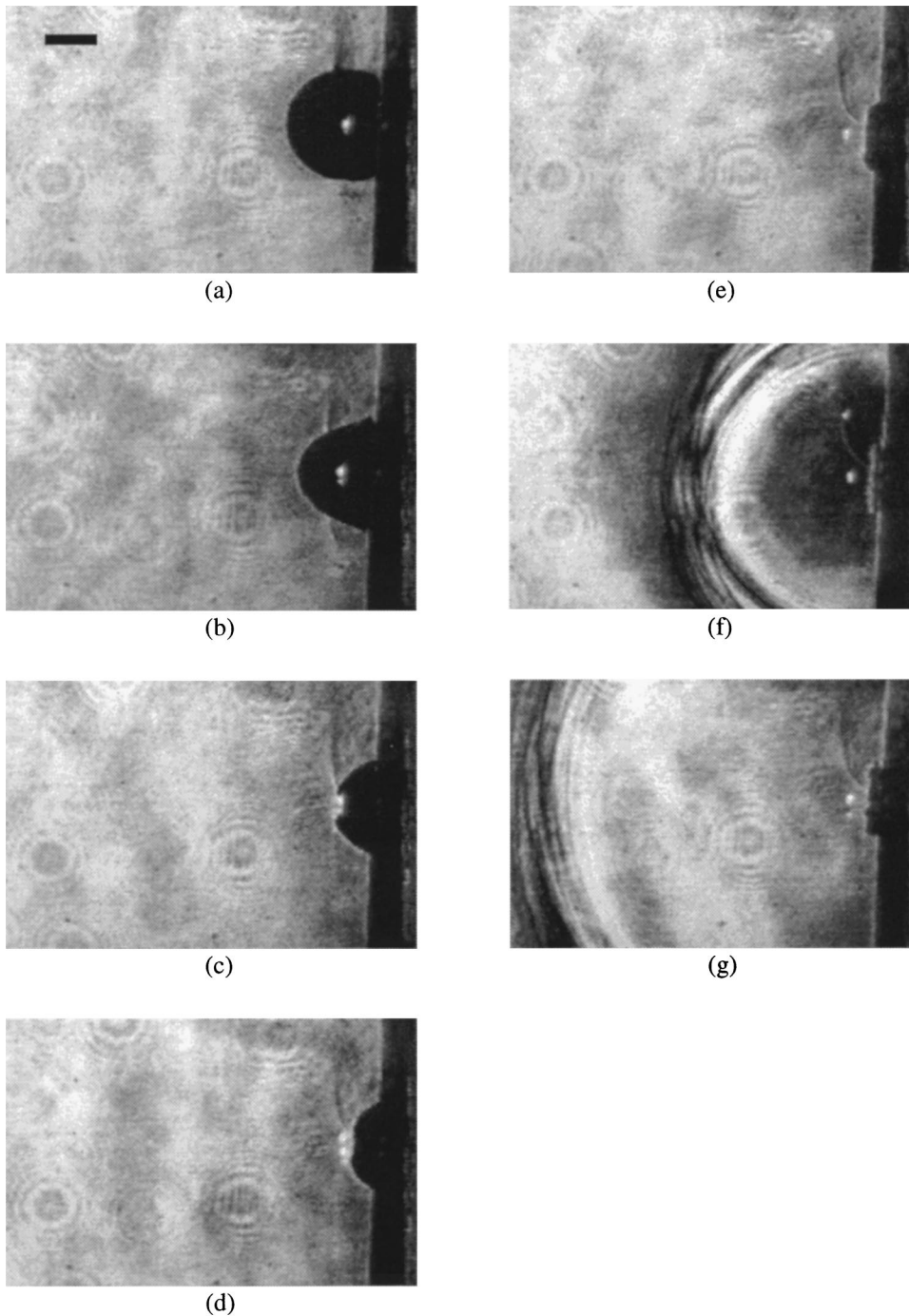


FIG. 8. Schlieren images of an oscillating cavitation bubble in front of a solid boundary for  $\gamma=0.54$ , at times  $t \cdot$  (a) 110, (b) 170, (c) 200, (d) 210, (e) 215, (f) 221, (g) 223  $\mu\text{s}$ . The horizontal bar in image (a) measures 1 mm.

icates the fluid in this region has a different density or density gradient to that in the outer lying fluid. As stated before, this is due to either pressure or temperature changes. The last photograph in the sequence [Fig. 8(g)] was taken when the pressure (voltage) as recorded by the transducer has dropped to a value similar to that before the jet impacted the boundary. The shock fronts have now moved to the extreme left of the figure while the dark region observed in the previous photograph has now gone. The bubble (or bubbles) in this picture is clearly expanding.

### III. CONCLUSIONS

In this paper we have experimentally investigated laser-generated cavitation bubbles created close to a solid bound-

ary for two values of stand-off parameter  $\gamma=0.65$  and  $0.54$ . Using a combination of schlieren photography and a thin film transducer placed on the surface of the solid we have recorded bubble shapes and induced pressures in both the solid and the surrounding fluid. Concentrating on time intervals around the first minimum volume, for the larger stand-off parameter value ( $\gamma=0.65$ ) the solid is subjected (as indicated by the transducer trace) to several distinct pressure stresses of varying duration and intensity. It is found that, due to the presence of a thin liquid layer between the collapsing bubble and the boundary, the splash phenomena is still valid for this stand-off parameter value. The impact of the splash on the bubble surface results in shock wave emission and subsequent disruption in the bubble shape. The con-

tinued collapse of the bubble is no longer symmetric and hindered by flows set up by the splash impact. As a result, only weak shock waves are emitted at minimum volume with different parts of the bubble reaching minimum volume at different times. The longer lasting pressure stress is due to fluid flow effects, which are greatly enhanced by the splash impact. For the smaller stand-off parameter value  $\gamma=0.54$ , the Schlieren pictures indicate that there is no observable liquid layer between the bubble and boundary during collapse, and splash effects are found to be no longer significant. The collapse process is relatively symmetric in this case and the compression of the bubble contents is such that strong shock waves are emitted at minimum volume.

It is interesting to note that changing the stand-off parameter by only a small amount, i.e., 0.65 to 0.54, results in a completely different pressure signal at the boundary around the time of the first minimum volume, demonstrating how sensitive a function of distance the collapse process can be in certain ranges.

### ACKNOWLEDGMENTS

This work was sponsored in part by grants from ESPRC and the DERA.

- <sup>1</sup>M. S. Plesset and A. Prosperetti, "Bubble dynamics and cavitation," *Annu. Rev. Fluid Mech.* **9**, 145–185 (1977).
- <sup>2</sup>K. A. Mörch, *Erosion*, edited by C. H. Preece (Academic, London, 1979), pp. 309–353.
- <sup>3</sup>J. R. Blake and D. C. Gibson, "Cavitation bubbles near boundaries," *Annu. Rev. Fluid Mech.* **19**, 99–123 (1987).
- <sup>4</sup>T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1994).
- <sup>5</sup>M. Kornfeld and L. Suvorov, "On the destructive action of cavitation," *J. Appl. Phys.* **15**, 495–506 (1944).

- <sup>6</sup>C. F. Naudé and A. T. Ellis, "On the mechanism of cavitation by non-hemispherical cavities collapsing in contact with a solid boundary," *Trans. Am. Soc. Mech. Eng. J. Basic Eng.* **83**, 648 (1961).
- <sup>7</sup>T. B. Benjamin and A. T. Ellis, "The collapse of cavitation bubbles and the pressures thereby produced against solid boundaries," *Philos. Trans. R. Soc. London, Ser. A* **260**, 221–240 (1966).
- <sup>8</sup>M. S. Plesset and R. B. Chapman, "Collapse of an initially spherical vapour cavity in the neighborhood of a solid boundary," *J. Fluid Mech.* **47**, 283–290 (1971).
- <sup>9</sup>W. Lauterborn and H. Bolle, "Experimental investigations of a cavitation-bubble collapse in the neighborhood of a solid boundary," *J. Fluid Mech.* **72**, 391–399 (1975).
- <sup>10</sup>L. A. Crum, "Surface oscillations and jet development in pulsating bubbles," *J. Phys. Colloq.* **40**, C8–225 (1979).
- <sup>11</sup>Y. Tomita and A. Shima, "Mechanisms of impulsive pressure generation and damage pit formation by the bubble collapse," *J. Fluid Mech.* **169**, 535–564 (1986).
- <sup>12</sup>J. R. Blake, B. B. Taib, and G. Doherty, "Transient cavities near boundaries. Part 1. Rigid boundary," *J. Fluid Mech.* **170**, 474–497 (1986).
- <sup>13</sup>A. Vogel, W. Lauterborn, and R. Timm, "Optical and acoustic investigations of the dynamics of laser-produced cavitation bubbles near a solid boundary," *J. Fluid Mech.* **206**, 299–338 (1989).
- <sup>14</sup>S. Zhang, J. H. Duncan, and G. L. Chahine, "The final stage of the collapse of a cavitation bubble near a rigid wall," *J. Fluid Mech.* **257**, 147–181 (1993).
- <sup>15</sup>A. Shima, K. Takayama, and Y. Tomita, "Mechanisms of the bubble collapse near a solid wall and the induced impact pressure generation," *Rep. Inst. High Speed Mech.* **48**, No. 367, 77–97 (1984).
- <sup>16</sup>S. J. Shaw, Y. H. Jin, W. P. Schiffers, and D. C. Emmony, "The interaction of a single laser-generated cavity in water with a solid surface," *J. Acoust. Soc. Am.* **99**, 2811–2824 (1996).
- <sup>17</sup>R. P. Tong, W. P. Schiffers, S. J. Shaw, J. R. Blake, and D. C. Emmony, "The role of 'splashing' in the collapse of a laser-generated cavity near a rigid boundary," *J. Fluid Mech.* (in press).
- <sup>18</sup>A. Philipp and W. Lauterborn, "Cavitation erosion by single laser-produced bubbles," *J. Fluid Mech.* **361**, 75–116 (1998).
- <sup>19</sup>W. P. Schiffers, D. C. Emmony, and S. J. Shaw, "Visualization of the fluid flow field around a laser generated oscillating bubble," *Phys. Fluids* **9**, 3201–3208 (1997).

# Field measurements of sonic boom penetration into the ocean

R. A. Sohn

*Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543*

F. Vernon, J. A. Hildebrand, and S. C. Webb

*Scripps Institution of Oceanography, La Jolla, California 92093-0205*

(Received 21 October 1999; accepted for publication 3 March 2000)

Six sonic booms, generated by F-4 aircraft under steady flight at a range of altitudes (610–6100 m) and Mach numbers (1.07–1.26), were measured just above the air/sea interface, and at five depths in the water column. The measurements were made with a vertical hydrophone array suspended from a small spar buoy at the sea surface, and telemetered to a nearby research vessel. The sonic boom pressure amplitude decays exponentially with depth, and the signal fades into the ambient noise field by 30–50 m, depending on the strength of the boom at the sea surface. Low-frequency components of the boom waveform penetrate significantly deeper than high frequencies. Frequencies greater than 20 Hz are difficult to observe at depths greater than about 10 m. Underwater sonic boom pressure measurements exhibit excellent agreement with predictions from analytical theory, despite the assumption of a flat air/sea interface. Significant scattering of the sonic boom signal by the rough ocean surface is not detected. Real ocean conditions appear to exert a negligible effect on the penetration of sonic booms into the ocean unless steady vehicle speeds exceed Mach 3, when the boom incidence angle is sufficient to cause scattering on realistic open ocean surfaces. © 2000 Acoustical Society of America. [S0001-4966(00)03106-4]

PACS numbers: 43.28.Mw, 43.30.Nb, 43.30.Hw [DLB]

## INTRODUCTION

Objects traveling faster than the speed of sound generate shock waves that result in impulsive acoustic signatures known as sonic booms. The typical acoustic signature of a sonic boom is the “N-wave” [Fig. 1(a)], which is characterized by sharp pressure jumps at the front and back of the waveform, with a slow pressure drop in between. It has been recognized from the early days of supersonic flight that sonic booms generate undesirable environmental impacts over populated areas,<sup>1</sup> primarily because of startle response to the shock wave pressure rise, and low-frequency building response (i.e., vibration, rattle).

The undesirable acoustic qualities of sonic booms led to legislation in the U.S. (and most countries internationally) forbidding supersonic flight and the generation of sonic booms over land, except in designated military corridors. As a result, most sonic booms are currently generated over the ocean. Sources of sonic booms over water include the Concorde, which flies routinely between Paris and New York, and rocket launches associated with satellite deployments.

The restriction of supersonic flight to air spaces over water has refocused sonic boom environmental impact research to the marine habitat, and to marine mammals, in particular. While the characteristics of sonic booms in air are well understood and supported by a vast body of research (e.g., Carlson and Maglieri,<sup>2</sup> Darden<sup>3</sup>), data constraining the penetration of sonic booms into the ocean, and the characteristics of boom pressure signatures underwater, are scarce. The original theory for the propagation of sonic booms across the air/sea interface was developed by Sawyers,<sup>4</sup> and by Cook.<sup>5</sup> For level flight, booms generated by objects traveling at speeds less than that of sound in water (1500 m/s, or Mach ~4.4) create an evanescent wavefield in the water

column, decaying exponentially with depth. The decay is wavelength dependent, with short wavelengths (high frequencies) attenuating faster than long wavelengths (low frequencies).

The Sawyers and Cook theories were validated with laboratory experiments involving spherical blasts<sup>6</sup> and small, high-speed projectiles.<sup>7</sup> Early attempts to validate the theory with field experiments, however, were unsuccessful. Young<sup>8</sup> and Urlick<sup>9</sup> attempted to quantify the penetration of sonic booms in the ocean (in separate experiments) by measuring boom pressure signatures immediately above, and at several depths below, the air/sea interface. Underwater sonic boom pressure measurements from these experiments exhibit different decay rates with respect to depth, and neither matches the analytical theory or laboratory data. Urlick's results deviated enough to cause him to question the validity of the evanescent wave theory for sonic booms in water.

Disagreement between the field data and the analytical theory introduced some uncertainty regarding the validity of the theory and its underlying assumptions in real world as opposed to laboratory conditions. In particular, the theories of Sawyers<sup>4</sup> and Cook<sup>5</sup> both assume a perfectly flat ocean surface, but the ocean surface is continually perturbed by ocean waves. The possible effects of a realistic ocean roughness on the penetration of booms into the water was recently investigated using numerical methods by Rochat and Sparrow,<sup>10</sup> and Cheng and Lee,<sup>11</sup> with each arriving at different conclusions. Rochat and Sparrow concluded that roughness has a negligible effect, with underwater pressure level variations from a flat interface of 1 decibel or less. In contrast, Cheng and Lee concluded that the sea surface roughness exerts a first-order effect on boom penetration, particularly at large depths and low frequencies. At issue is the magnitude of the scattered component of sonic boom

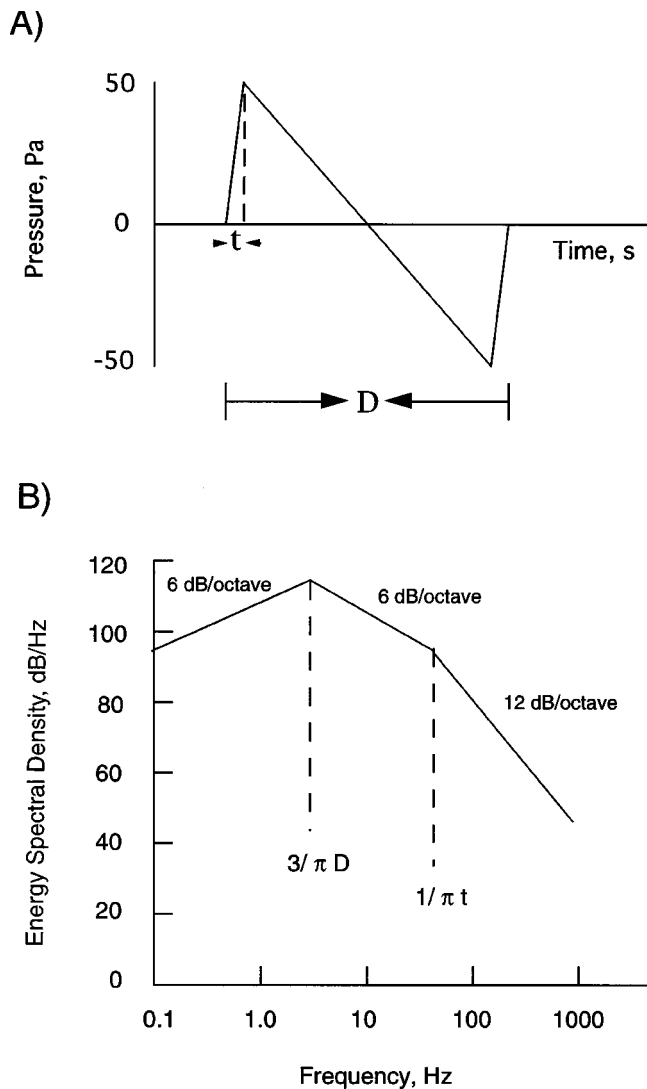


FIG. 1. Characteristics of a typical “N-wave” sonic boom measured on the ground. (a) Simple *N*-wave time series. The straight-line approximation is parameterized by a “rise time” (*t*) which is the time from the onset of the boom to maximum pressure, and the total duration (*D*) of the waveform. Rise times typically range from 2–20 ms, and durations are typically 100–400 ms. (b) Theoretical energy spectrum of an *N*-wave boom with a rise time of 8 ms and a duration of 350 ms.

energy in the water column, and its proportion to the evanescent signal.

The lack of consistency between the numerical studies,<sup>10,11</sup> and disagreement between the analytic theory<sup>4,5</sup> with the field experiments,<sup>8,9</sup> underscores the need for reliable measurements of sonic booms underwater to serve as a benchmark for the validation of theoretical models, and to provide a foundation for environmental impact assessments. The early data of Young and Urick suffer from the technological limitations of their day. Specifically, the data acquisition systems employed in the experiments did not have adequate low-frequency response, and the pressure measurements are likely contaminated by the interaction of the sonic boom with the mechanical systems used to suspend the hydrophones in the water column. A sonic boom measured by Urick<sup>9</sup> is shown in Fig. 2. The pressure signatures bear little resemblance to an *N*-wave, primarily because the measurement system lacked the low-frequency response to measure

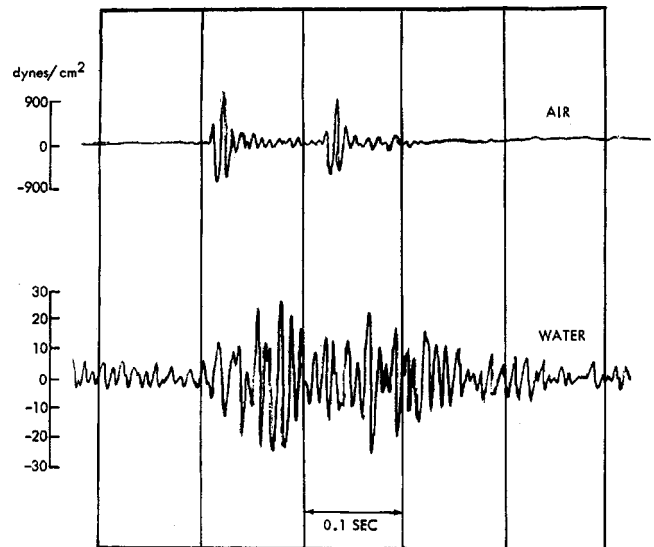


FIG. 2. Sonic boom measured by Urick (Ref. 7) in the air and in the water at a depth of ~15 m. The effective pass-band of these measurements is 75–150 Hz.

the slow pressure decay between the fore and aft shocks. In addition, the ringing observed in the water column measurement indicates that the data are contaminated by mechanical interactions with the suspension system.

Instrumentation has improved dramatically since the experiments of Young and Urick, and modern systems are capable of making high-fidelity measurements of sonic booms underwater. For example, a Concorde boom was serendipitously recorded in 1996 by a hydrophone array off Nova Scotia,<sup>12</sup> and the underwater boom waveform contains the low-frequency components missing in the Young and Urick measurements. However, correlation of this measurement with theoretical results is difficult because the boom waveform was not measured in air, and because the underwater waveform is curiously complicated by ringing that may have resulted from the excitation of a low-frequency seismic mode in the shallow seabed.

In this paper we present the results of a field study that provides the first simultaneous, high-fidelity measurements of sonic booms in the air and ocean. We take advantage of modern instrumentation systems to extend the frequency range of the measurements down to a few Hz, and use data telemetry from a small spar buoy to avoid contaminating the incoming sonic boom as it crosses the air/sea interface. We measured six sonic booms at five depths in the water (down to 112 m), and just above the air/sea interface. We find that the pressure signatures we measured are in excellent agreement with the analytical theory of Cook<sup>5</sup> (as implemented by Sparrow and Ferguson<sup>13</sup>) down to ~40–50 m, where the signal is lost in the ambient field.

## FIELD PROGRAM

The field experiment was conducted on May 11–12, 1999 in the East Cortez Basin (32.2°N, 118.7°W), approximately 140 km W-SW of San Diego, CA. The location provided deep water (1600 m) within a military air space. Sur-

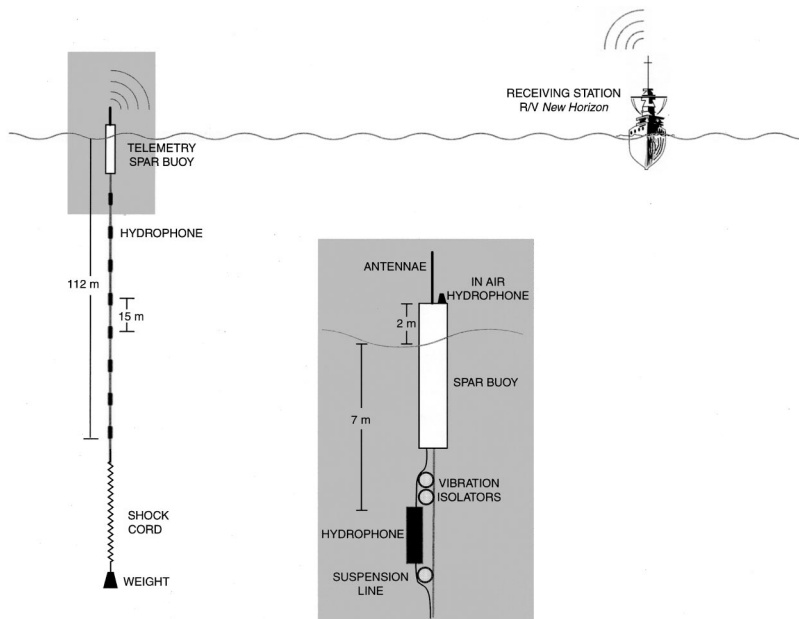


FIG. 3. Schematic illustration of instrumentation employed in this experiment. Details of the near surface package and the suspension system are shown in inset.

face support for the experiment was provided by the R/V NEW HORIZON. Flight support was provided by the Naval Air Station at Pt. Mugu, CA.

During the course of the experiment, the sea-state and weather were essentially constant. The NOAA environmental buoy near Catalina Island (33.75°N, 119.08°W) reported a significant wave height and period of 1.2 m and 8 s, respectively, an air temperature of 14 °C, and an average wind speed of 2 m/s. Wind and swell estimates made on station aboard the R/V NEW HORIZON were somewhat greater, with prevailing winds of 4.5–9 m/s (10–20 kts) and a swell of 2–3 m at a predominant period of 8–10 s.

The conditions during the experiment fall into the Beaufort Force 4–5 category. Wind and waves generally grew during the day. By afternoon scattered white caps were present, but breaking waves were not observed.

### Instrumentation

The instrument package consisted of a 115-m vertical hydrophone array suspended from a small spar buoy (6.5-m tall, 0.4-m diameter) at the sea surface (Fig. 3). The vertical array contained nine hydrophone elements (Benthos/Aquadyne AQ-1 cartridges), with one phone mounted in air on top of the buoy (~2 m above the water), and eight phones at 15-m intervals in the water column. The shallowest and deepest phones were located at 7 and 112 m depth, respectively. The hydrophones' signals were sampled at 500 Hz and analog bandpass filtered. The in-air hydrophone provided a flat response from 1–60 Hz, a 10-dB/decade roll-off between 60 and 150 Hz, and a ~50 dB/decade roll-off above 150 Hz. The water hydrophones had a flat response from 3–200 Hz, with a steep roll-off (~70 dB/decade) above 200 Hz.

Annular vibration isolators with a nominal resonant frequency of ~1 Hz were utilized to decouple the vibration of the suspension line from the hydrophone elements (see Fig. 3, inset). A 20-kg weight was attached to the bottom of the hydrophone array with a length of shock cord to maintain a

vertical profile in the water column, and to move the resonant frequency of the suspension system (~0.1 Hz) outside the frequency range of the measurements. A GPS receiver was mounted on top of the buoy to provide position updates once a minute, with a nominal error of 100 m.

The hydrophone data were digitized at the buoy, and then telemetered in real-time to a receiving station aboard the R/V NEW HORIZON. A buoy telemetry system was employed to avoid contaminating the sonic boom as it crossed the air/sea interface. During the supersonic passes the R/V NEW HORIZON stood off several kilometers from the buoy with the engines idling and the propellers de-clutched. The bandwidth of the telemetry system permitted the digitization of six data channels (out of nine hydrophone stations) at a 500-Hz sample rate. Different channel configurations were employed for the first and second day of the experiment (see Results, below).

Measurements of the in-air sonic booms generated during the experiment were also made by personnel from NASA Dryden using the SABER system.<sup>14</sup> These measurements are not reported here, but provided redundancy should the in-air sensor on the data buoy have failed. The SABER measurements (with a sampling rate of 10 kHz) were also used to examine the high-frequency characteristics of the booms entering the water column.

### Flight plan

Six supersonic passes were made with U.S. Navy F-4 aircraft over the two days of the experiment. The overflight altitude was varied from 610–6100 m (2000–20000 ft) to provide a range of boom pressures (96–530 Pa, or 2–11 psf) at the air/sea interface. Aircraft speeds ranged from Mach 1.07–1.26, corresponding to the aircraft's best speed at each altitude.

After transiting from Pt. Mugu to the experimental site, the aircraft established radio contact with the R/V NEW HORIZON and were given an updated target position. Once a

**Pass 1, Mach 1.07, 610 m (2000 ft)**

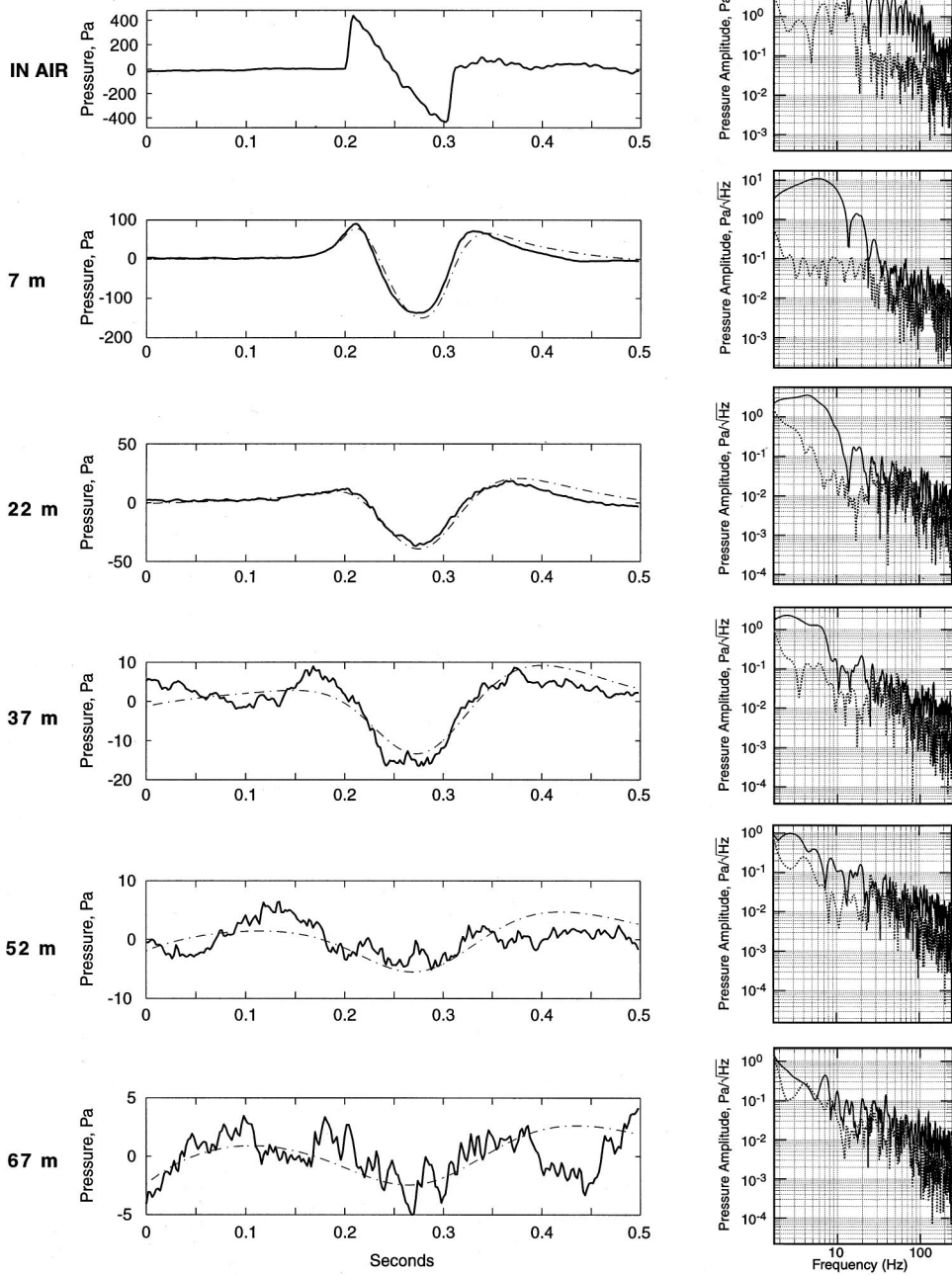


FIG. 4. Pressure measurements and theoretical predictions from Pass 1, Mach 1.07 at 610-m altitude. Time series (left side) and amplitude spectra (right side) are shown for each measurement. Sensor depth is shown on the far left side. For time series plots, data are shown as a solid line and theoretical predictions (based on in-air measurement at top) are shown as a dashed line. For spectral plots, boom amplitude spectra are shown as a solid line and ambient noise spectra are shown as a dashed line. Note that the pressure scale (y-axis) used for the time series plots varies with depth.

supersonic run was underway, the test pilots noted the speed, altitude, and heading of the aircraft, along with a single position (latitude/longitude) at the beginning of the run. The aircraft position and magnetic heading (accurate to within 3 deg) at the beginning of the run were used to estimate a flight track for each supersonic pass.

Since the aircraft did not have GPS data loggers, it is impossible to know the exact lateral distance between the flight track and the data buoy. Based on straight-line flight track estimates, all of the tracks pass over the buoy within the tolerance of the estimates, except for one (Pass 2). The straight-line estimate for Pass 2 runs ~1 km west of the buoy. This is significant given that the horizontal error is only slightly less than the aircraft altitude (1.5 km). However, no aircraft had an automatic heading-hold system, and

therefore the actual flight tracks (for all the supersonic runs) may have varied from the straight-line estimates.

**Results**

Time series and spectral plots of each sonic boom measured during the experiment (Pass 1–6) are shown in Figs. 4–9. Data from six hydrophones were recorded during each run. Measurements made during the first day of the experiment (Pass 3, 5, 6; Figs. 6, 8, 9) recorded hydrophone data in-air, and at 7, 22, 37, 82, and 112 m beneath the sea surface. Measurements made during the second day of the experiment (Pass 1, 2, 4; Figs. 4, 5, 7) recorded data in-air, and at 7, 22, 37, 52, and 67 m beneath the sea surface. The shift toward shallower depths on the second day was made after it

**Pass 2, Mach 1.15, 1525 m (5,000 ft)**

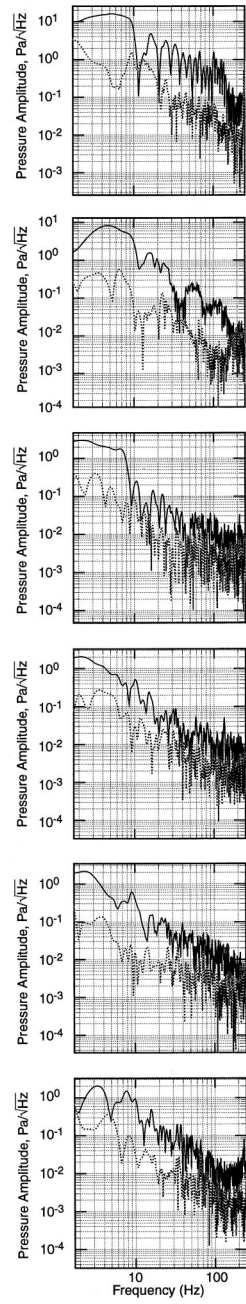
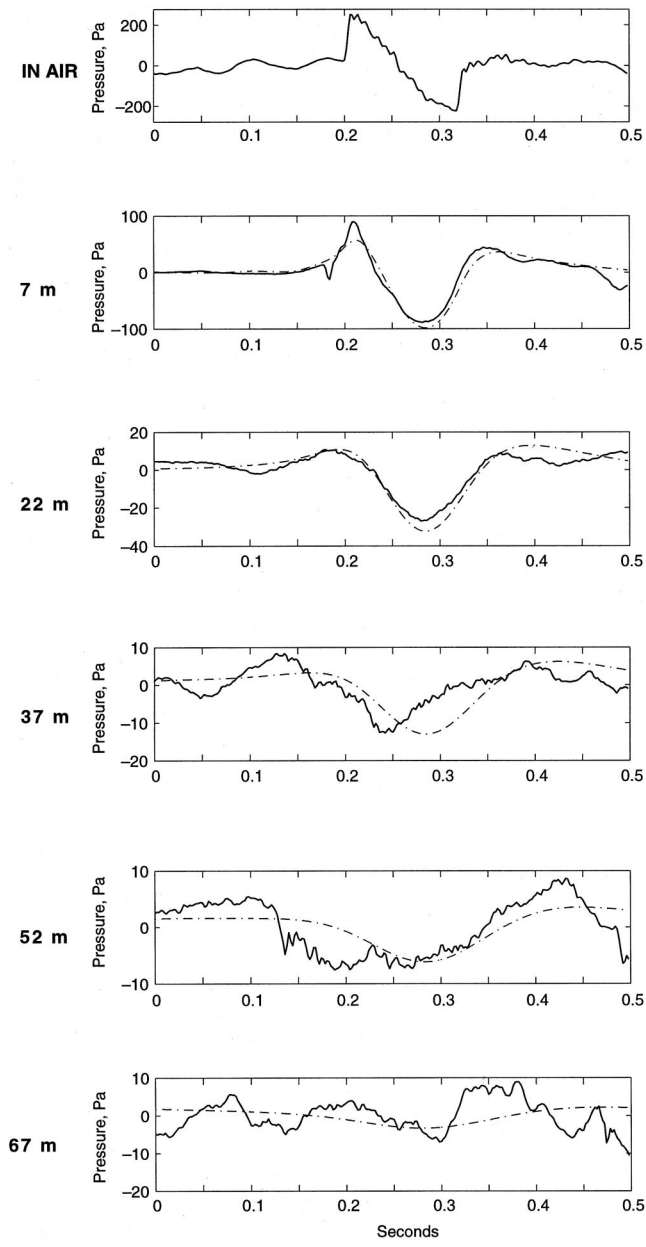


FIG. 5. Pressure measurements and theoretical predictions from Pass 2, Mach 1.15 at 1525-m altitude. Plots as in Fig. 4.

was realized that the booms were failing to generate detectable signals near the bottom of the hydrophone array.

The time series data for the in-water measurements are compared with the linear, analytical theory of Sparrow and Ferguson,<sup>13</sup> which is based on the work of Cook.<sup>5</sup> The method assumes a flat air/sea interface, and no interaction with the seafloor (deep water), but allows for arbitrary boom wave shapes. To generate the theoretical pressure signatures in the water column, the spectrum of the in-air pressure signature was calculated, weighted by an exponential decay with respect to wavelength and depth, and then transformed back into the time domain. A Blackman window was applied to the in-air data segment before calculating the fast Fourier transform (FFT) to reduce Gibbs phenomenon associated with truncating an infinite series. The theoretical waveforms shown in Figs. 4–9 were bandpass filtered (3–200 Hz) to mimic the analog circuitry of the in-water hydrophones. This

filtering slightly distorts the theoretical wave shapes, but is required to provide an equal comparison to the measurements.

The amplitude spectrum of the boom waveform and the ambient noise field at each hydrophone channel are shown on the right side of Figs. 4–9. The amplitude spectrum of the ambient noise field represents the FFT of a randomly selected segment of data immediately preceding the boom arrival. A Gaussian window was applied to both the boom pressure signature and the ambient noise segment prior to estimation of the amplitude spectrum.

**Discussion**

The primary experimental objective of this work was to make high-fidelity measurements of sonic booms at the air/sea interface and at several depths in the water column. We



**Pass 3, Mach 1.17, 2135 m (7,000 ft)**

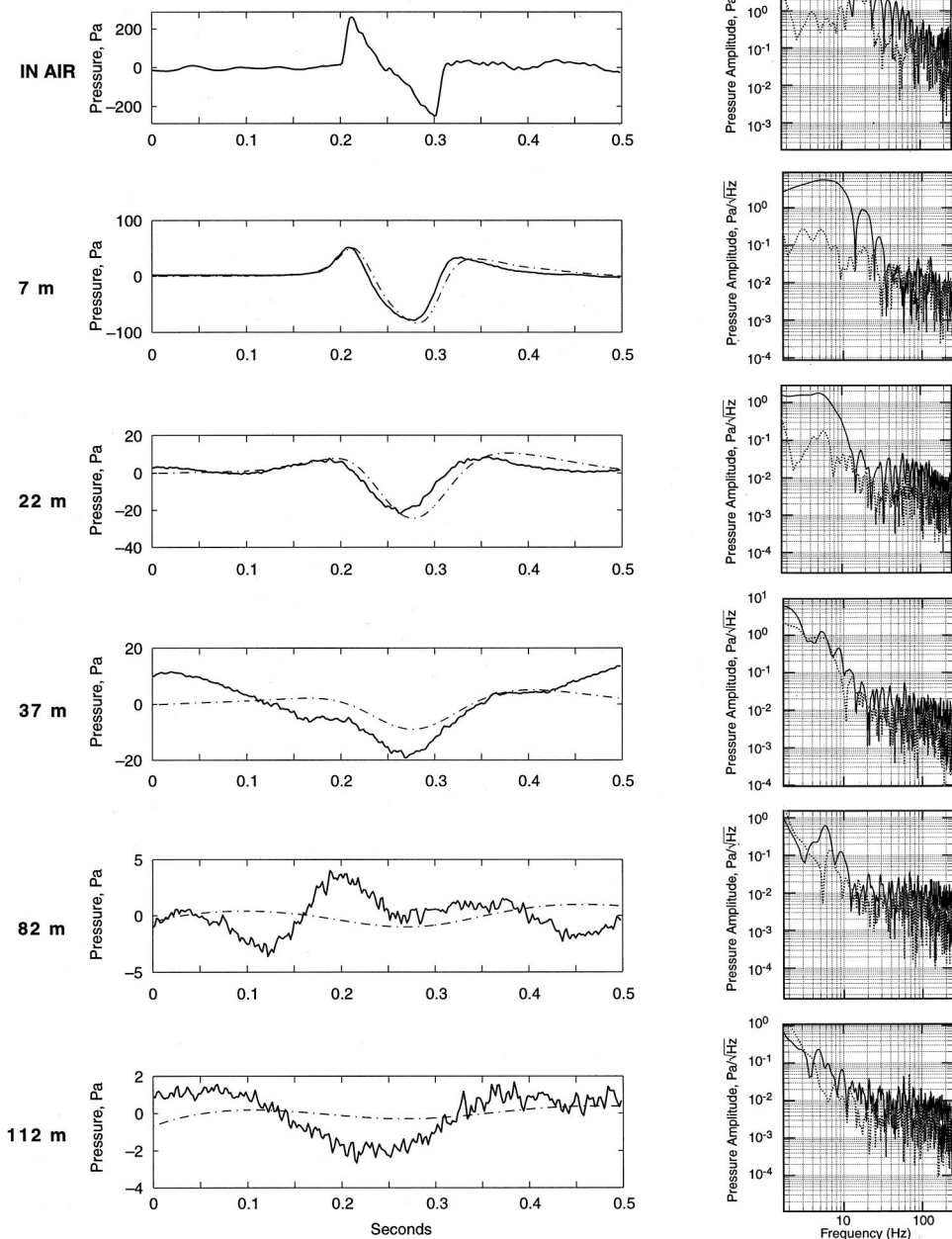


FIG. 6. Pressure measurements and theoretical predictions from Pass 3, Mach 1.17 at 2135-m altitude. Plots as in Fig. 4.

begin the discussion by assessing the extent to which this objective was met. We then compare our measurements to theoretical predictions, and discuss the implications of the similarities/differences for the validation of the theory under real ocean conditions. We conclude with a brief review of some remaining issues and unanswered questions regarding the penetration of sonic booms into the ocean.

**Fidelity of sonic boom pressure measurements**

We begin by examining the fidelity of the in-air measurement, which is important considering that it is used as a source function for the theoretical predictions of boom pressures underwater. All of the in-air pressure signatures in this experiment are characterized by fairly simple *N*-waves, and the amplitude spectra of the in-air signals have the expected shape, with two separate corner frequencies corresponding to

the boom duration and rise time [compare top right panels of Figs. 4–9 with Fig. 1(b)]. The low-pass filter applied to the in-air pressure data (60 Hz corner frequency) removes any contributions from reflected phases at the microphone. It appears that our in-air measurements adequately characterize the sonic boom impinging the air/sea interface above the vertical hydrophone array, especially at low frequencies, which are of primary importance to this study.

Several of the supersonic passes made during the experiment were at fairly low altitudes, and under these conditions individual shocks from the various aerodynamic features (e.g., nose, wings, cockpit) would not be expected to have coalesced into single bow and tail shocks (e.g., Hayes<sup>15</sup>). Uncoalesced shocks create extra spikes in the acoustic signature. These extra spikes generate relatively high-frequency pressure perturbations that are removed by the low-pass filters, and in these cases we expect that the simple *N*-waves

Pass 4, Mach 1.13, 2865 m (9400 ft)

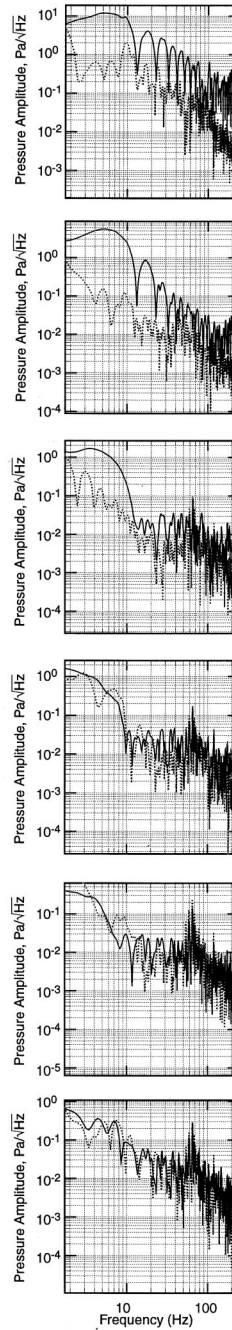
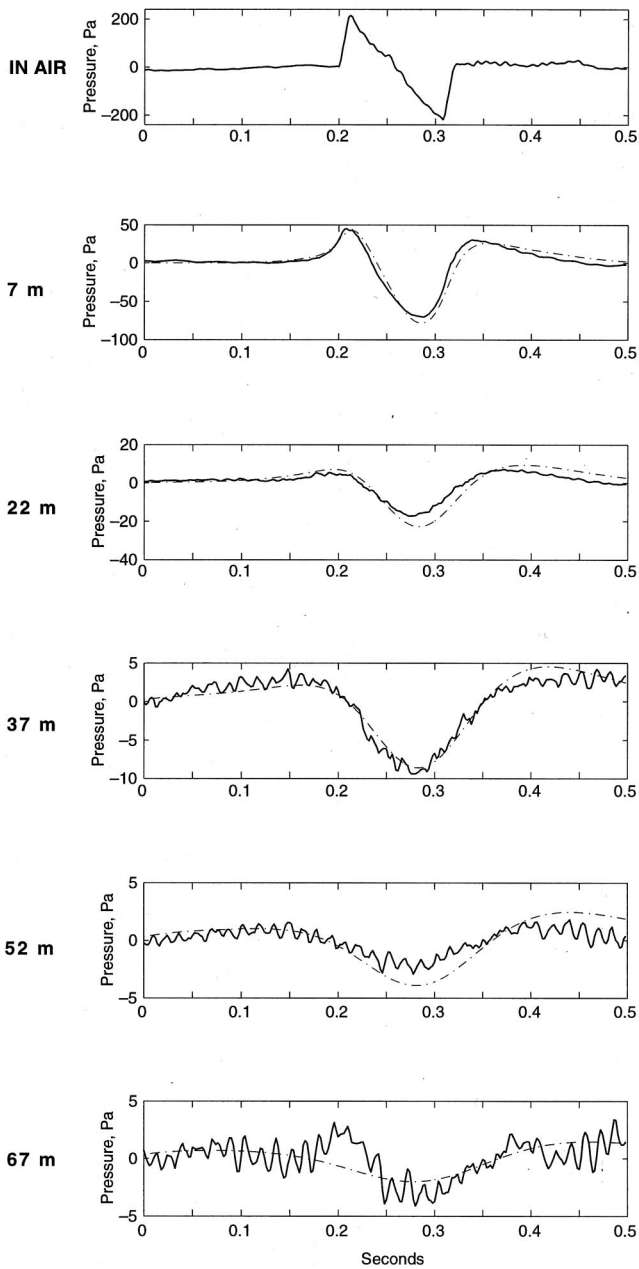


FIG. 7. Pressure measurements and theoretical predictions from Pass 4, Mach 1.13 at 2865-m altitude. Plots as in Fig. 4.

rendered by the in-air sensor do not perfectly represent the actual booms at the sea surface. Indeed, the SABER measurements made aboard the R/V NEW HORIZON for the low-altitude passes contain spikes embedded in the *N*-wave signature (although these measurements also contain spikes from reflections off the ship's superstructure). It will be seen in the following section that the failure to capture high-frequency spikes in the in-air measurement is not a significant shortcoming since these features are almost immediately removed from the evanescent wavefield beneath the sea surface.

The principle concern for the underwater boom measurements is to keep noise levels on the individual sensors as close to ambient as possible. Achieving low noise levels on a hydrophone array suspended at shallow depths is difficult because the hydrophones are mechanically linked to the mo-

tion of the surface wavefield. The suspension system decouples the hydrophones from the jerking of the array by motions of the buoy and from strum (e.g., Sotirin and Hildebrand<sup>16</sup>) induced by current flowing past the array.

Inspection of the ambient noise pressure spectra in Figs. 4–9 indicates that our attempts to minimize noise levels on the hydrophone array using vibration isolators and shock cord (Fig. 3) were fairly successful. The pressure variance of the ambient field on individual hydrophone elements was typically less than 100 Pa<sup>2</sup> in the relevant band from 3–200 Hz. This corresponds to a dynamic head of less than 1 mm of water, root-mean-square (rms). Ambient noise pressure variance on the deepest phones is especially small, with typical values of 2–5 Pa<sup>2</sup>. In practical terms this resulted in excellent signal-to-noise levels for boom measurements down to about 40–50-m depth. At this depth the amplitude of the

**Pass 5, Mach 1.21, 4570 m (15,000 ft)**

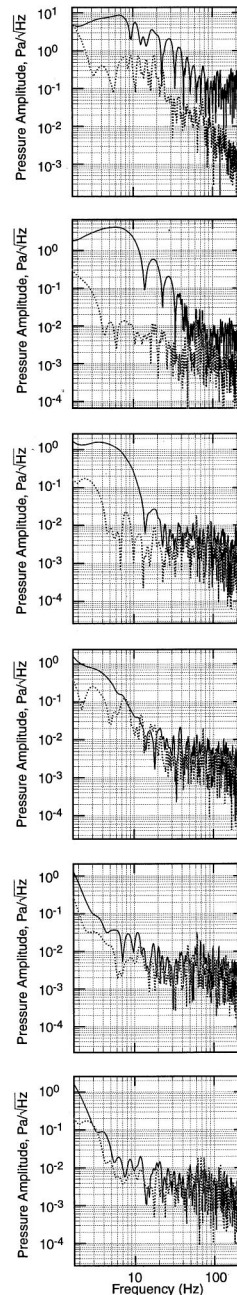
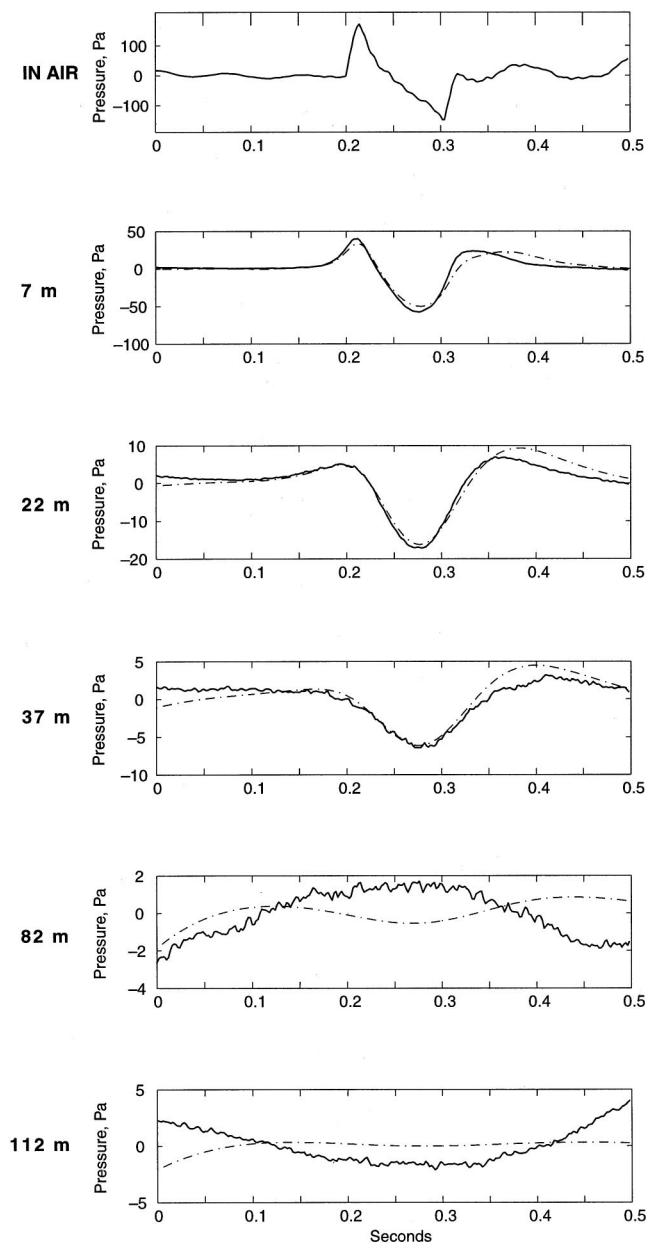


FIG. 8. Pressure measurements and theoretical predictions from Pass 5, Mach 1.21 at 4570-m altitude. Plots as in Fig. 4.

boom pressure signal is equal to or less than the ambient levels on the hydrophone array.

Additionally, the underwater boom signatures do not contain any ringing as do those measured by Urick<sup>9</sup> or Desharnais and Chapman.<sup>12</sup> Nor does the acoustic field contain any measurable perturbation from the surface buoy. This demonstrates that by using a small diameter spar buoy as a surface mooring for the data acquisition system we avoided contaminating the boom waveform with mechanical coupling down the suspension line. By conducting the experiment in deep water, we also appear to have avoided interaction with the seabed.

We conclude that the pressure measurements made during the course of this experiment provide accurate renderings of the sonic boom wavefields generated at the instrument array, especially at low frequencies.

**Agreement between data and theory**

The agreement between our data and the analytical method of Sparrow and Ferguson,<sup>13</sup> which is based on the theory of Cook,<sup>5</sup> can be observed by comparing the solid and dashed lines in the time series plots of Figs. 4–9. Data and theory are in agreement at all depths and for all booms within the limitations of the signal-to-noise ratio. The signal is above the noise to depths of 37 m for all booms, and to greater depths for lower altitude flights with stronger booms.

Examination of spectral attenuation provides additional insight into the agreement between our data and linear theory. The evanescent decay of a sonic boom underwater<sup>4,5</sup> scales as  $e^{-k_0\mu z}$ , where  $k_0$  is the wavelength in air divided by Mach number,  $z$  is depth, and  $\mu = \sqrt{1 - M^2/W^2}$ , where  $M$  is Mach number and  $W$  is the ratio of sound speed in air to

**Pass 6, Mach 1.26, 6100 m (20,000 ft)**

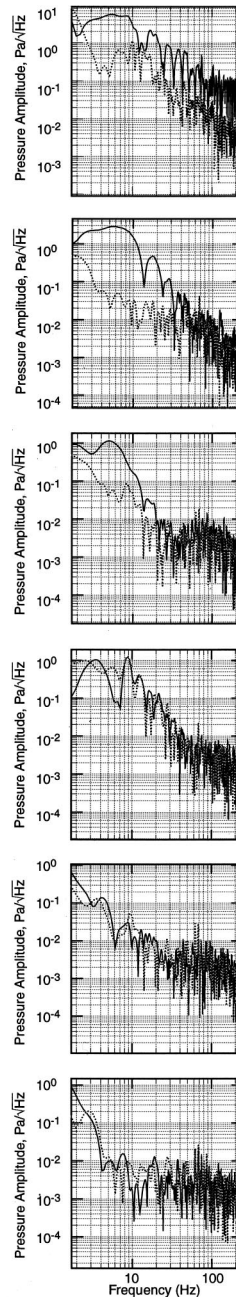
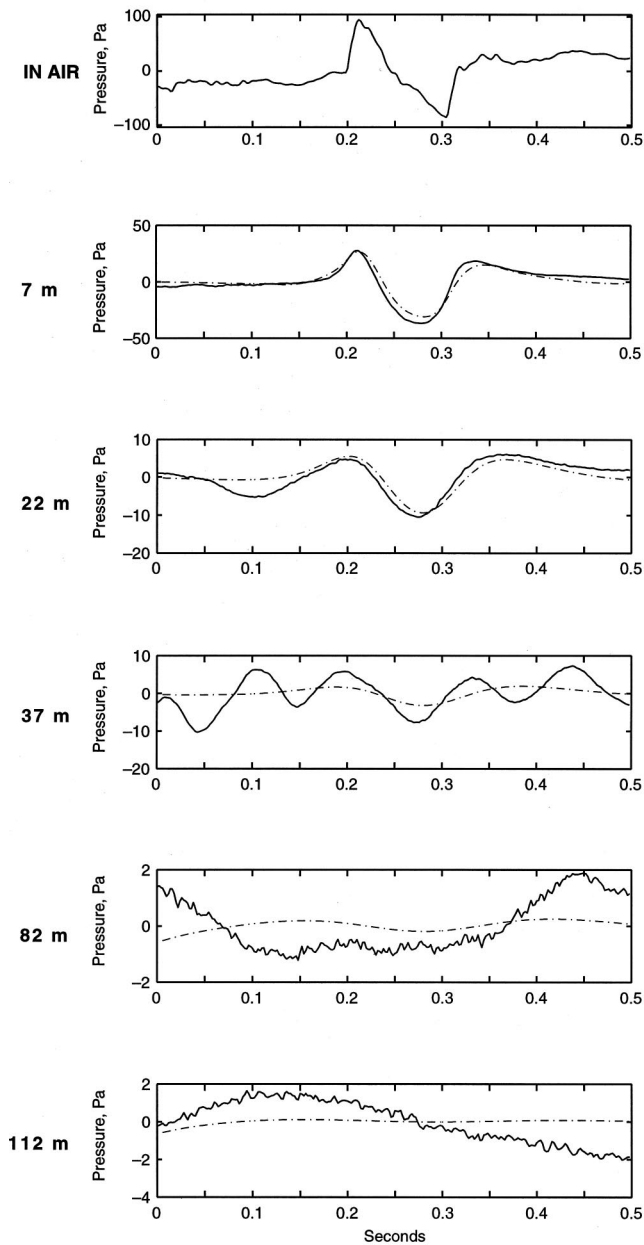


FIG. 9. Pressure measurements and theoretical predictions from Pass 6, Mach 1.26 at 6100-m altitude. Plots as in Fig. 4.

water. The correlation between this theoretical expression and our Pass 1 measurements (strongest boom) is shown in Fig. 10. The data follow the theoretical decay curves until they approach the noise floor, at which point ambient noise overwhelms the signal.

The agreement between the predicted waveform and the signal measured at the deeper hydrophones precludes the existence of a scattered component of the sonic boom signal propagating as an acoustic wave in the water with an amplitude greater than 4 Pa peak to peak. The largest sonic boom measured in air had a peak-to-peak amplitude of 800 Pa, demonstrating that the scattered component in the water was no more than 0.5% of the incident boom amplitude.

In order for an airborne acoustic wave to enter the ocean it must have a grazing angle of at least  $77^\circ$  (from the horizontal). The grazing angles of booms generated during this experiment (i.e., Mach cone angle) range from  $\sim 20\text{--}30^\circ$ .

This means that, in the high-frequency approximation, ocean waves had to be  $50\text{--}60^\circ$  steep to generate a meaningful scattered component and cause significant deviations from a flat interface assumption. However, waves in the open ocean break if their steepness exceeds  $\sim 8^\circ$  (Stokes<sup>17</sup>). Indeed, observations of wave steepness in the open ocean from a variety of sea states give values of  $0.5\text{--}6^\circ$  (Khandekar<sup>18</sup>).

Given the low supersonic speeds, and hence low grazing angles of the booms generated during our experiment, our failure to measure a significant scattered component of the boom signal underwater is not surprising. Indeed, we only expect a significant scattered component if the boom incidence angle is within a few degrees of  $77^\circ$ , or just below the angle required for acoustic transmission. If we allow for a maximum wave steepness of  $6^\circ$ , then we expect a detectable scattered component at incidence angles of  $71^\circ$ , corresponding to a vehicle speed of Mach  $\sim 3$ .

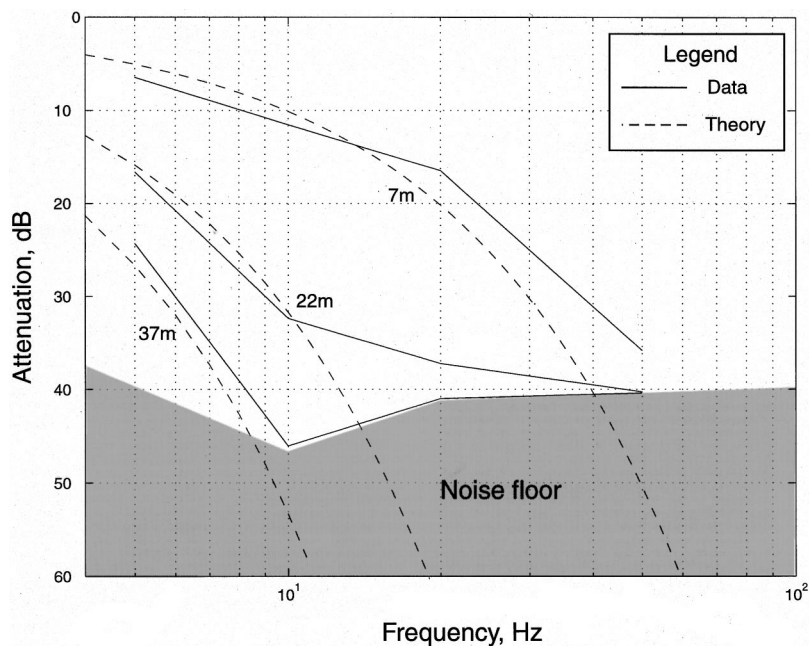


FIG. 10. Evanescent attenuation of sonic booms underwater. Pressure data from the top three hydrophones from Pass 1 (Fig. 4) are compared with linear theory (Refs. 4 and 5). Mean spectral noise levels at 22-m depth are shown for comparison. Attenuation is referenced to the spectral levels of the in-air measurement, such that the noise floor would move up on this plot for the smaller booms of Passes 2–6.

Thus the scattered boom signal is expected to be negligible until vehicle speeds reach Mach  $\sim 3$ . At Mach  $\sim 3$  very rough sea states have the potential to scatter significant amounts of boom energy into the water column. Between Mach 3 and Mach 4.4 the magnitude of the scattered signal will increase with vehicle speed and sea state. Above Mach 4.4 standard acoustic transmission theory applies.

### Remaining issues

The results of this experiment demonstrate that the penetration of sonic booms into deep water from level aircraft flight at velocities significantly less than 1500 m/s (Mach  $\sim 4.4$ , the speed of sound in water) can be accurately predicted with analytical theory. We found that the presence of a ‘real’ surface wavefield at the air/sea interface did not cause any observable differences between the data and the theory. Thus there is now uniform agreement between the original theories of Sawyers<sup>4</sup> and Cook,<sup>5</sup> laboratory tests,<sup>6,7</sup> the numerical method of Rochat and Sparrow,<sup>10</sup> and the field results of our experiment. As a result we consider the first-order physics of the penetration of sonic booms across the air/sea interface to be well understood and validated.

There are three special cases of sonic boom penetration into the ocean that were not addressed in this experiment: penetration into shallow water, penetration from booms propagating at speeds greater than Mach  $\sim 3$ , and penetration from booms generated during unsteady flight (maneuvers). Penetration of booms into shallow water is a phenomenon that will almost certainly require experimental data owing to the difficulties associated with incorporating a realistic seafloor into numerical computations.<sup>19</sup> In addition, a single experiment may not suffice in this regard since the shallow-water problem may be site-specific owing to the diversity of seabed compositions found offshore the United States (and globally). For example, the continental margins of the eastern U.S. have a different composition, and hence different

geoacoustic characteristics, than the continental slope of the western coast.

As discussed above, significant amounts of scattered energy from booms are expected if vehicle speeds exceed Mach 3, and if they exceed Mach 4.4 then sonic boom penetration into the water is governed by standard plane wave transmission theory. These high speed scenarios are associated with much more efficient boom penetration into the water, and may generate underwater pressure levels substantially larger than those measured in this experiment. However, few vehicles in existence today travel faster than Mach 3, and those that do tend to do so at very high altitudes (e.g., space shuttle reentries). Consequently, these types of booms are both rare and of low amplitude. Therefore, it is not clear whether or not this particular scenario warrants a concerted experimental program.

Flight maneuvers have the potential to modify the penetration of sonic booms into the ocean by changing the angle of incidence of the booms at the air/sea interface. In this case maneuvers are broadly interpreted to include any unsteady flight operations, including climb, descent, and acceleration. We are concerned not so much with classic focusing effects,<sup>20</sup> but rather with phase matching of the acoustic signal at the air/sea interface. Under the proper conditions, an object traveling at relatively low supersonic speeds (a 29° dive at Mach 1.5, for example) can generate a boom pressure field that phase matches along the horizontal air/sea interface. This is physically equivalent to the conditions resulting from steady flight at speeds greater than Mach 4.4 as described in the previous paragraph. A survey of the potential for routine rocket launches and aircraft operations to maneuver and generate phase-matched booms over the ocean is beyond the scope of this work. However, if the flight track parameters for a given mission are known, the phase velocity of the boom pressure field at the air/sea interface can be

easily checked. A field experiment to measure the penetration of phase-matched booms from maneuvering objects would be significantly more complicated than that conducted in this work, but might be warranted, for example, if routine rocket launches generate phase-matched booms.

## CONCLUSIONS

We measured six sonic booms from Navy F-4 aircraft under steady flight with sensors located just above the air/sea interface and at five depths in the upper 115 m of the water column. Boom pressures exhibit a frequency-dependent decay with depth, with low frequencies penetrating significantly farther than high frequencies. All of the boom pressure signals measured in this experiment decay to ambient levels in all frequency bands by 40–50 m. Boom waveforms measured at individual depths in the water column exhibit excellent agreement with analytical theory that assumes a flat air/sea interface. At supersonic speeds significantly less than Mach 3, we conclude that the ocean wavefield does not significantly affect boom penetration into the ocean, and that analytical theory (e.g., Sawyers,<sup>4</sup> Cook,<sup>5</sup> Sparrow and Ferguson<sup>13</sup>) is a valid tool to estimate underwater sonic boom pressures. In particular, these theories can be confidently used to estimate potential environmental impacts of sonic booms underwater.

## ACKNOWLEDGMENTS

We thank Jacques Lemire and Glenn Offield for superb engineering support during the instrumentation development and field experiment, Bill Gaines for facilitating the experimental logistics and clearances, Jo Griffith for graphics, and the test pilots at the Naval Air Station Pt. Mugu for their flight support during the experiment. We thank Vic Sparrow for helpful discussions, and two anonymous reviewers for suggestions that have improved the manuscript. This work was funded by the NASA Langley Research Center (Technical Monitor, Dr. Kevin Shepherd).

- <sup>1</sup>P. N. Borsky, "Community reactions to sonic booms in the Oklahoma City area," AMRI-TR-65-37, U.S. Air Force (Feb. 1965).
- <sup>2</sup>H. W. Carlson and D. J. Maglieri, "Review of sonic-boom generation theory and prediction methods," J. Acoust. Soc. Am. **51**, 675–685 (1972).
- <sup>3</sup>"Status of sonic boom methodology and understanding," NASA Conference Publication 3027 edited by C. M. Darden (1988).
- <sup>4</sup>K. Sawyers, "Underwater sound pressure from sonic booms," J. Acoust. Soc. Am. **44**, 523–524 (1968).
- <sup>5</sup>R. K. Cook, "Penetration of a sonic boom into water," J. Acoust. Soc. Am. **47**, 1430–1436 (1970).
- <sup>6</sup>J. F. Waters and R. E. Glass, "Penetration of sonic boom energy into the ocean: An experimental simulation," Hydrospace Research Corp. Final Report on Contract FA70WAI-185, HRC TR 288 (June 1970), available from NTIS/DTIC as AD 711 963.
- <sup>7</sup>P. Intriери and G. Malcolm, "Ballistic range investigation of sonic-boom overpressures in water," AIAA J. **11**, 510–516 (1973).
- <sup>8</sup>R. W. Young, "Penetration of sonic booms into the ocean," J. Acoust. Soc. Am. **44**, 392 (1968).
- <sup>9</sup>R. J. Urick, "Sonic booms in the sea," Naval Ordnance Laboratory Technical Report NOLTR 71-30 (1971).
- <sup>10</sup>J. L. Rochat and V. W. Sparrow, "Two-dimensional focusing of sonic boom noise penetrating an air-water interface," AIAA J. **35**(1), 35–39 (1997).
- <sup>11</sup>H. K. Cheng and C. J. Lee, "A theory of sonic boom noise penetration under a wavy ocean," AIAA Paper 98-2958, 2nd AIAA Theoretical Fluid Mechanics Meeting, Albuquerque, NM (1998).
- <sup>12</sup>F. Desharnais and D. M. F. Chapman, "Underwater measurements of a sonic boom," Proc. Oceans 97 MTS/IEEE, **592–596** (1997).
- <sup>13</sup>V. W. Sparrow and T. Ferguson, "Penetration of shaped sonic boom noise into a flat ocean," AIAA Paper 97-0486, 35th Aerospace Sciences Meeting and Exhibit, Reno, NV (1997).
- <sup>14</sup>S. R. Norris, E. A. Haering, Jr., and J. E. Murray, "Ground-based sensors for the SR-71 sonic boom propagation experiment," NASA Dryden Flight Research Center, NASA-TM-104310 September 1995.
- <sup>15</sup>W. D. Hayes, "Brief review of the basic theory," *Sonic Boom Research*, edited by A. R. Seebass in NASA SP-147 (April 1967), pp. 3–7.
- <sup>16</sup>B. J. Sotirin and J. A. Hildebrand, "Large aperture digital acoustic array," IEEE J. Ocean Eng. **13**, 271–13,281 (1988).
- <sup>17</sup>G. G. Stokes, "Supplement to a paper on the theory of oscillatory waves," *Mathematical and Physical Papers* (Cambridge University Press, Cambridge, 1880), pp. 314–326.
- <sup>18</sup>M. L. Khandekar, *Operational Analysis and Prediction of Ocean Wind Waves* (Springer-Verlag, New York, 1989), pp. 214.
- <sup>19</sup>V. W. Sparrow, "Determination of aircraft sonic boom noise penetration into seas, bays, and lakes for environmental assessment," Pennsylvania State University Final Technical Report on Contract F41624-96-1-0006, February 1998.
- <sup>20</sup>D. L. Lansing and D. J. Maglieri, "Comparison of measured and calculated sonic boom ground patterns due to several different aircraft maneuvers," NASA TN D-2730 (1965).

# Decay of large underwater bubble oscillations

B. Edward McDonald<sup>a)</sup> and Charles Holland  
*Saclant Undersea Research Centre, La Spezia, Italy*

(Received 9 December 1998; revised 9 August 1999; accepted 31 January 2000)

Pressure–time series from breathing-mode oscillation of large (centimeter scale or larger) underwater bubbles reveal much higher decay rates than can be explained using viscous, thermal, or radiative mechanisms which apply to microbubbles. It is shown that if one assumes energy transfer to shape oscillations (surface capillary waves) of large amplitude in subharmonic resonance with the breathing mode [M. S. Longuet-Higgins, *J. Acoust. Soc. Am.* **91**, 1414 (1992)], then the shape oscillations can drive fluid motions outside the bubble capable of exciting turbulent instabilities. Application of an appropriate eddy viscosity from mixing-length theory to the viscous decay mechanism appears to offer a credible explanation for the observed large decay rates. An analysis is given to show that energy is transferred from the breathing mode to surface capillaries fast enough to make the proposed decay mechanism viable.

[S0001-4966(00)01405-3]

PACS numbers: 43.28.Py, 43.30.Jx [DLB]

## INTRODUCTION

Ocean acoustic research is sometimes carried out using explosions or implosions, both of which result in a large (a few to several cm radius) bubble which pulsates at a frequency governed by depth and bubble parameters. After the large bubble is created, the resulting pressure–time series at a distant hydrophone soon settles down to an apparently linear oscillation resembling a strongly damped sinusoid. In fact, the damping tends to be about one factor of  $e$  per oscillation cycle regardless of whether the source is an imploding lightbulb or nearly a kilogram of TNT. We will review some of the data leading to this conclusion and then seek to determine within an idealized model how the energy loss rate of a large bubble can be so great and so nearly independent of bubble size and depth.

Implosions of glass spheres<sup>1,2</sup> or lightbulbs<sup>3,4</sup> have been investigated with increasing interest as acoustic sources for probing the underwater environment, in part because of regulatory restrictions and dangers inherent in the use of explosives in shallow-water environments. Small implosive sources may be used in the vicinity of receiving arrays with minimal concern about equipment damage. Lightbulbs provide an inexpensive and safe implosion device when combined with a breaking mechanism to trigger the implosion. A useful discussion of several lightbulb implosion experiments has recently been given by Heard *et al.*<sup>3</sup> While useful relations exist<sup>3,4</sup> for predicting the peak level and breathing frequency, no current theory adequately predicts the oscillation decay rate for large bubbles.

Figure 1 gives examples of lightbulb implosion pressure–time series taken by SACLANTCEN<sup>4</sup> at depths of 36 and 91 m. Although pressure–time series from implosion experiments are subject to variability due to the nonsphericity of the source and interaction among water, gas, and glass fragments,<sup>2</sup> signatures of the type shown in Fig. 1 are quite typical of such implosions at depths down to approximately

300 m. In fact, Fig. 2 of Ref. 3 displays a lightbulb implosion pressure–time series for a depth of 18 m which is remarkably similar in shape to Fig. 1(a) and (c). One notable feature of the time series in Fig. 1 is the short decay time of the acoustic pressure, which is on the order of the bubble oscillation period.

At depths greater than about 300 m, lightbulbs implode spontaneously because of the hydrostatic pressure. For implosions at depths greater than about 300 m, glass spheres can be constructed with enough strength to withstand the pressure until the breaking mechanism is triggered. At depths much greater than 300 m, a different type of time series is found. At depths ranging between approximately 2500 and 3500 m, Orr and Schoenberg<sup>1</sup> recorded implosion pressure–time series consisting of one large spike with hardly any oscillatory tendency. As we will show below, a decrease in number of oscillations with implosion depth is consistent with acoustic radiation as being the dominant decay mechanism at great depth.

The literature applicable to the decay of bubble oscillations<sup>5–9</sup> is primarily oriented toward small (subcentimeter) bubbles. The gas volume<sup>3</sup> inside a 100-W lightbulb is approximately 150 cm<sup>3</sup> and would correspond to a sphere of radius 3.3 cm at atmospheric pressure. Theoretical work for the implosion and oscillation of bubbles of this approximate size is scarce. Perhaps the reason for this is that such large bubbles in the ocean are rare, rise quickly to the surface, and are subject to breaking up.

Although the current work began as an attempt to understand the oscillation of implosion bubbles, we came to realize that our results also apply to some features of bubble oscillations produced by underwater explosions [e.g., SUS (signal underwater sound) charges]. Chapman<sup>10</sup> has given a thorough and useful review of the properties of bubble pulses following SUS charge detonations. Results of that review are incorporated into our comparisons between theory and data below.

After reviewing the dominant decay mechanisms (ther-

<sup>a)</sup>Electronic mail: mcdonald@saclantc.nato.int

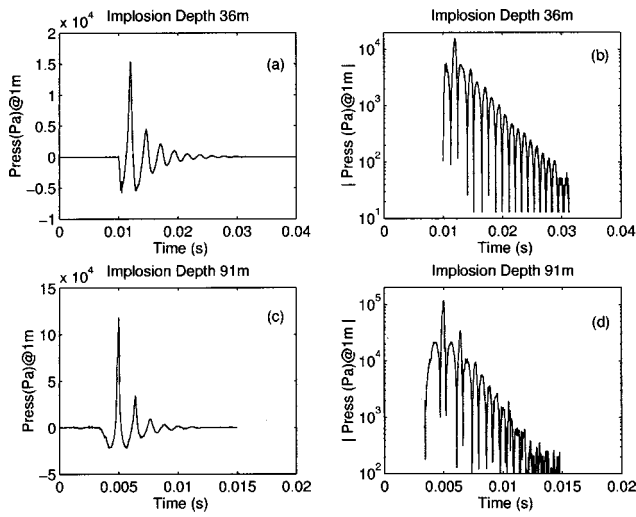


FIG. 1. Pressure–time series (Ref. 4) at 1-m radius for lightbulb implosions at depths 36 and 91 m. Left panels use linear scale, while log scale for absolute value of the pressure is shown in right panels to demonstrate linear decay.

mal, viscous, and radiative) normally invoked for small bubble oscillations (and which are inadequate to explain the damping of centimeter-scale bubble oscillations at depths less than about 300 m), we will propose a simple model which appears consistent with the observed decay of oscillations in lightbulb-sized bubbles. The model invokes turbulence in the fluid just outside the bubble excited by shape oscillations<sup>9</sup> which are in subharmonic resonance with the bubble’s breathing mode. The model yields a predicted ratio of decay time to oscillation period which is independent of bubble radius and depth, in approximate agreement with data for depths less than about 300 m.

It is beyond the scope of the present paper to give an in-depth review of the often-difficult theory of fluid turbulence. A few basic results from turbulence theory are, however, cited from the fluid dynamics literature. It is our hope to encourage further investigation into the possible role of turbulence in large-bubble decay by showing that a few simple and fairly robust results from turbulence theory lead straightforwardly to a theoretical explanation for decay rates much larger than linear theories of bubble oscillation predict.

## I. BUBBLE OSCILLATION DAMPING CONSTANTS

When nonlinearities in bubble implosion/ explosion have disappeared, breathing-mode oscillations of spherical bubbles have time dependence  $\exp(-\beta_{\text{TH}} - \beta_R - \beta_V - i\omega_b)t$ , where the breathing frequency is

$$\omega_b = a^{-1} \sqrt{\frac{3\gamma p}{\rho}}, \quad (1)$$

$a$  is the equilibrium bubble radius,  $\gamma$  is the ratio of specific heats of the gas (1.4 for air),  $p$  is the ambient pressure, and  $\rho$  is the fluid density. The linear nature of the oscillation is evident after the first one or two oscillation periods in Fig. 1, and persists for several oscillations thereafter. We are concerned with bubbles of sufficient size to be regarded as adiabatic, so that  $a \propto p^{-1/3\gamma}$  for a lightbulb implosion. The damp-

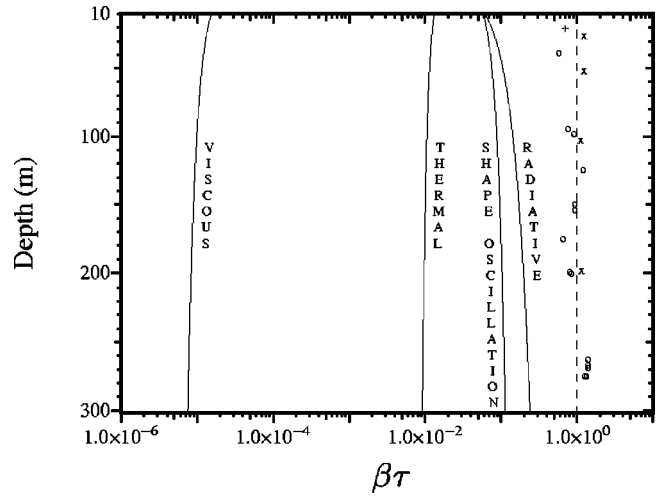


FIG. 2. Curves: Decay exponent  $\beta\tau$  for various mechanisms as a function of depth for oscillation of a bubble of radius 3 cm at pressure 1 atm taken to depth.  $\tau$  is the bubble oscillation period  $\tau = 2\pi/\omega_b$ . Dashed line: Fluid turbulence decay from present model with  $\kappa = 0.5$  in Eq. (14). Data points: + from Ref. 3, Fig. 2; o Saclantcen data (Ref. 4); x from ratio of second to third bubble pulse peak pressure for SUS charges (Ref. 10).

ing constants are as follows:<sup>5,9</sup> Thermal damping ( $\omega_b/2\pi < 10$  kHz):

$$\frac{\beta_{\text{TH}}}{\omega_b} = \frac{3(\gamma-1)}{2a} \left( \frac{D}{2\omega_b} \right)^{1/2}; \quad (2)$$

Radiative damping:

$$\frac{\beta_R}{\omega_b} = \frac{1}{2c} \left( \frac{3\gamma p}{\rho} \right)^{1/2}; \quad (3)$$

Viscous damping:

$$\beta_V = \frac{2\nu}{a^2}. \quad (4)$$

Constants introduced in (2)–(4) are the thermal diffusivity of air,  $D \approx 0.2$  cm<sup>2</sup>/s, the water sound speed  $c = 1500$  m/s, and the kinematic viscosity of water  $\nu \approx 0.01$  cm<sup>2</sup>/s. The damping rates per oscillation period due to thermal, viscous, and radiative effects for a lightbulb-sized bubble  $a_0 = 3$  cm of gas at pressure 1 atmosphere imploded at depth are shown in Fig. 2. The shape oscillation damping rate is discussed below. Also plotted in Fig. 2 are data for bubbles created by lightbulb implosions and bubble pulses (excluding the initial shock wave) from 820-g TNT SUS charges.<sup>10</sup> Remarkably, the decay per oscillation of the bubble created by 820-g TNT SUS charges is nearly indistinguishable from that for lightbulb implosion bubbles, suggesting that some sort of scale-independent mechanism may be at work.

It is clear from Fig. 2 that the sum of the theoretical decay rates for a nominal 3-cm radius bubble taken adiabatically to depth is far too small to explain the observed damping rates, of order one  $e$ -fold per oscillation period. Thus, we are motivated to consider other mechanisms that would explain the observed decay rate. We first considered nonlinearity in a purely radially oscillating bubble. The numerical integration carried out by Prosperetti and Lezzi<sup>11</sup> retained dominant nonlinearities for a bubble collapse from a radius 4 times its pressure equilibrium radius. At the end of the first



oscillation period, the bubble had lost approximately 74 percent of its initial energy. When these numbers are interpreted as an underwater implosion, the effective depth is approximately 2700 m, and the resultant oscillation decay turns out to be well explained by linear acoustic radiation. This statement is justified as follows:

For a lightbulb of internal pressure  $0.8 \text{ atm}^3$  imploding at a depth of 2700 m, the linear decay in amplitude  $\exp(-\beta_R \tau)$  due to radiation [Eq. (3)] is approximately 0.5. Near equilibrium, a submerged bubble's internal energy increases approximately quadratically with the difference between bubble radius and its equilibrium value. Thus, Eq. (3) would predict that each oscillation period results in a diminution of energy by a factor of approximately 0.25. The 74% energy loss found in Ref. 11 appears mostly linear in nature for a scaled implosion depth of 2700 m. Since radiative and nonlinear losses decrease with decreasing implosion depth, the mechanism proposed by Ref. 11 will not account for the observed energy loss for shallow (less than 300-m depth) implosions.

Another possible rapid energy loss mechanism by large bubbles is that of shape oscillations in subharmonic resonance with the breathing mode.<sup>9</sup> Longuet-Higgins had considered this mechanism to explain large damping rates observed in subcentimeter scale bubbles. Shape oscillations may be larger in amplitude than the breathing-mode oscillations because they do not change the bubble's volume. The normal modes of shape oscillations are such that the radial displacements of the bubble surface are spherical surface harmonics,<sup>9</sup> and are thus proportional to  $P_n(\cos \theta)$ ,  $P_n(x)$  being the Legendre polynomial of order  $n$ , and  $\theta$  the polar angle with respect to some arbitrary direction. The surface harmonics consist of standing capillary waves on a sphere with frequency<sup>9</sup>

$$\sigma_n^2 = (n-1)(n+1)(n+2) \frac{T}{\rho a^3}, \quad (5)$$

where  $T=73 \text{ dyne/cm}$  is the surface tension of water. The subharmonic resonance condition invoked by Longuet-Higgins is

$$2\sigma_n = \omega_b. \quad (6)$$

Substituting Eqs. (1) and (5) into (6) yields the resonance condition

$$(n-1)(n+1)(n+2) = a \frac{3\gamma p}{4T}. \quad (7)$$

We can take the large  $n$  limit in (7) as justified by substitution of minimal parameters  $a=1 \text{ cm}$ ,  $p=1 \text{ atm}$ , yielding a value of  $1.437 \times 10^4$  for the right-hand side, and  $n \approx 24$ . Lightbulb implosions will always lead to larger values of  $a$  and  $p$ , so the resonance condition becomes

$$n^3 \approx a \frac{3\gamma p}{4T}. \quad (8)$$

Even though  $n$  is constrained to integer values, the resonance is a very broad one due to the rapid decay of oscillations. This means that it is likely for large  $n$  that several harmonics near the value given in (8) will be excited.

According to Longuet-Higgins,<sup>9</sup> the dominant loss mechanism for shape oscillations is viscosity, resulting in a decay constant

$$\beta_{nV} = (n+2)(2n+1)v/a^2 \approx \frac{2\nu n^2}{a^2}. \quad (9)$$

The decay per oscillation period due to shape oscillations is shown with the other mechanisms in Fig. 1, and is again too small to explain the damping of order one  $e$ -fold per oscillation at these depths for centimeter scale bubbles.

## II. A MODEL INVOKING EDDY VISCOSITY

We hypothesize that fluid instabilities similar to those involved in turbulence may be generated by the shape oscillations of large bubbles, with a result of high energy decay rates. The mechanism we propose is that during its oscillation, the surface of the bubble develops fingering instabilities that manifest themselves as high-order shape oscillations. The high spatial variation of the resulting fluid velocity then drives smaller scale instabilities reminiscent of turbulence. Our proposed mechanism is analogous to a multiscale instability<sup>12</sup> proposed successfully to explain the onset of small-scale irregularities in the earth's ionosphere.

One of the most basic ways to model turbulent-like effects in a fluid is to use an eddy viscosity  $\nu_e$  from mixing length theory<sup>13</sup>

$$\nu_e \propto \langle \bar{v} \lambda \rangle, \quad (10)$$

where  $\bar{v}$  is the velocity of a parcel of fluid relative to the average flow velocity in its vicinity,  $\lambda$  is the mixing length or characteristic distance the parcel travels before dissolving into the background fluid, and the brackets  $\langle \rangle$  indicate an ensemble average. Eddy viscosity is one form of so-called anomalous transport<sup>14</sup> in which turbulent fluid motions lead to dissipative effects qualitatively similar to molecular dissipation (e.g., viscosity) but quantitatively many orders of magnitude larger. In computations of ocean fluid processes, it is customary to employ an eddy viscosity coefficient tuned to bring about the closest agreement between computed results and observations. Specific choices depend upon the length and time scales of interest, and the degree to which fluid instabilities may be operable. A recent calculation of nonlinear oceanic internal waves<sup>15</sup> used a value of  $10 \text{ m}^2/\text{s}$ , which is  $10^7$  times the kinematic viscosity of water.

Results developed in the theory of turbulence<sup>13</sup> are usually applied to describe the time-averaged evolution of an unstable fluid. The averaging period is taken to be much longer than the time scale of the dominant flow instabilities. The presence of an oscillating flow in our case does not prevent us from invoking the mixing length theoretical result of Eq. (10), since we will show *a posteriori* in Fig. 4 below that for the fluid surrounding a large oscillating bubble, fluid instability operates on a time scale much shorter than the bubble oscillation period.

If large shape oscillations are occurring (with large meaning that maximum radial displacement is a non-negligible fraction of the oscillation wavelength), it is rea-

sonable to take  $\lambda$  in Eq. (10) proportional to the shape oscillation wavelength, and  $\bar{v}$  proportional to the capillary wave speed. Thus,

$$\lambda \propto \frac{a}{n}, \quad (11)$$

and

$$\bar{v} \propto \omega_b \lambda. \quad (12)$$

The proportionality constants in (10)–(12) are of order unity. Thus, we hypothesize that

$$\nu_e = \kappa \omega_b \frac{a^2}{n^2}, \quad (13)$$

where  $\kappa$  is a dimensionless constant of order unity.

If we adapt the decay constant (9) to include  $\nu_e$  rather than the molecular value  $\nu$ , we find

$$\beta_{ne} \approx \frac{2\nu_e n^2}{a^2} = 2\kappa\omega_b, \quad (14)$$

where  $\beta_{ne}$  stands for the eddy transport decay rate of surface harmonic  $n$ . The decay rate per oscillation period for this hypothesized mechanism is

$$2\pi\beta_{ne}/\omega_b = 2\kappa = \text{const.} \quad (15)$$

The above mechanism with  $\kappa \approx 0.5$  would then lead to approximately one  $e$ -fold per oscillation, independent of bubble radius and depth (as long as  $n$  is large, which will apply for centimeter scale bubbles) in agreement with the time series of Fig. 1. As shown in Fig. 2, the radiative contribution to the decay increases as the square root of depth, so that at some large depth radiative damping would dominate and further enhance the total decay rate. This seems to be consistent with the results of Orr and Schoenberg<sup>1</sup> indicating very strong damping and almost no oscillation tendency at implosion depths of approximately 2500 m and greater.

If fluid instabilities are indeed operative in damping the oscillation of large bubbles, one should consider whether the traditional criterion for turbulence is satisfied; namely that the Reynolds number for the fluid motion be large compared to some value determined by the geometry of the system. In a system with characteristic length and velocity scales  $L$  and  $v$ , the Reynolds number is  $vL/\nu$ , where  $\nu$  is the molecular kinematic viscosity. For shape oscillations of mode  $n$ , the characteristic length scale is taken to be the average capillary wavelength,  $L = 2\pi a/n$ , and the velocity scale is taken to be  $v = \omega_b L$ . The Reynolds number is thus

$$R_e = \frac{\omega_b}{\nu} \left( \frac{2\pi a}{n} \right)^2. \quad (16)$$

The dependence upon depth of  $R_e$  for shape oscillations of the nominal 3-cm bubble taken to depth is shown in Fig. 3.  $R_e$  is of order  $10^4$ , certainly large enough to indicate fluid instability. Its depth dependence may be shown from (16) to be proportional to  $p^{-1/6-1/9\gamma}$ , where  $1/6+1/9\gamma = 0.246$ . We have shown in the dotted line of Fig. 3 that a fit of the form  $p^{-1/4}$  is quite accurate for the depth dependence of  $R_e$ .

The large values of  $R_e$  for depths less than about 300 m and its gradual decrease with depth support our hypothesis

that at small to moderate depths turbulence may be active in damping oscillations of large bubbles, while at great depth radiative damping becomes dominant.

### III. ENERGY TRANSFER: RAYLEIGH-TAYLOR INSTABILITY

In order for the proposed mechanism to explain the rapid damping of the breathing mode, there must be a mechanism for transferring energy to shape oscillations on a time scale shorter than the oscillation period. We will present some theoretical arguments to the effect that energy is indeed transferred rapidly to capillary waves during the outward acceleration phases of bubble oscillation. The mechanism is the Rayleigh–Taylor instability,<sup>16,17</sup> familiar as the operative instability mechanism when a heavy fluid overlays a lighter one. In such a configuration an initially plane interface between the fluids develops ripples, and then fingers of the two fluids migrate into each other as the heavier fluid attempts to sink and the lighter one attempts to rise. In the case of an oscillating bubble, gravity is replaced by the outward acceleration of the bubble wall during half the oscillation period. Although this acceleration is time dependent, we will obtain order of magnitude estimates for instability growth time using a constant root-mean-squared (rms) value for the acceleration.

For capillary wavelengths much smaller than the bubble radius, we may use results for a plane interface between air and water with  $x$  along the interface and  $y$  normal to it. For disturbances of the form  $\exp(ikx - k|y| + \Gamma t)$ , the instability growth rate including surface tension but neglecting the density of air when added to or subtracted from that of water is<sup>18</sup>

$$\Gamma^2 = Ak - \frac{T}{\rho} k^3, \quad (17)$$

where  $A$  is the acceleration of the interface taken to be positive in the direction of the water side of the interface. One should notice that during a contraction (negative  $A$ ) Eq. (17) leads to imaginary growth rate (propagating rather than growing capillary waves). Thus, the alternation of the sign of

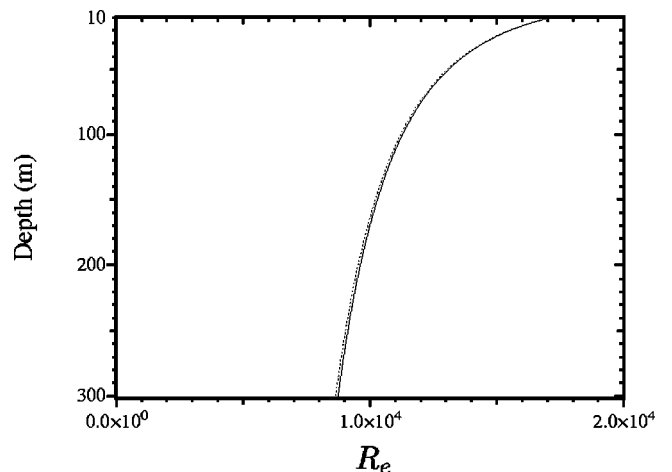


FIG. 3. Reynolds number [Eq. (16)] based on capillary wavelength and phase speed of shape oscillations of a 3-cm bubble taken to depth. The dotted line is a fit to  $p^{-1/4}$ .

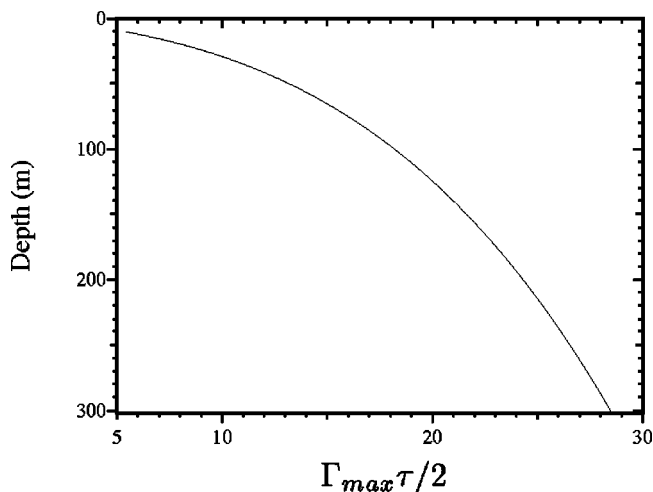


FIG. 4. The number of growth times for the most unstable Rayleigh–Taylor mode during the first outward acceleration half-cycle following a lightbulb implosion.

$A$  leads to alternating periods of capillary growth and propagation, but not to capillary decay.

The growth rate of the most unstable wave is obtained from (17) by maximizing with respect to  $k$ . Of the two roots of the resulting quadratic for  $k$ , only one is positive. We know that since  $A$ ,  $T$ , and  $\rho$  on the right-hand side of are all positive,  $\Gamma^2$  reaches a single maximum for  $k > 0$  and then becomes negative for large  $k$ . Substitution of the positive quadratic root of the extremum condition into (17) gives

$$\Gamma_{\max}^2 = \frac{2}{3} A^{3/2} \sqrt{\frac{\rho}{3T}}. \quad (18)$$

For  $A$  we take the value  $a_0(2\sqrt{2})^{-1}(1 - (p_0/p)^{1/3\gamma})\omega_b^2$ . This is a rough estimate of the rms acceleration during an imploded lightbulb's first outward acceleration phase. Substituting lightbulb parameters cited above, we may use Eqs. (1) and (18) to estimate the number of instability growth times  $\Gamma_{\max}\tau/2$  during the outward acceleration half cycle, where as above  $\tau = 2\pi/\omega_b$ . The result, displayed as a function of depth in Fig. 4, shows that the number of instability growth times in the first outward acceleration half cycle is considerably greater than unity for depths greater than about 10 m. The implication is that activation of the surface capillary ripples occurs on a time scale shorter than the bubble oscillation period.

#### IV. SUMMARY

We have shown that a possible candidate for the decay mechanism of large underwater bubble oscillations at modest depths (less than about 300 m) is fluid turbulence outside the bubble forced by shape oscillations in subharmonic resonance with the bubble's breathing mode. A simple model was proposed in which the decay times for shape oscillation conform to the observations at modest depths when eddy

viscosity (13) is used in (9), with one free parameter  $\kappa = 0.5$ . In order to make a more rigorous argument for this mechanism, one needs to examine in detail the subharmonic resonance as discussed by Longuet-Higgins<sup>9</sup> as a means of transferring energy quickly from the breathing mode to the shape oscillations. We have made a step in this direction by showing in the analysis leading to Fig. 4 that the Rayleigh–Taylor fingering instability has ample time to generate parasitic capillary waves during the first outward acceleration following a lightbulb implosion.

#### ACKNOWLEDGMENT

Work supported by the North Atlantic Treaty Organization.

- <sup>1</sup>M. Orr and M. Schoenberg, "Acoustic signatures from deep water implosions of spherical cavities," *J. Acoust. Soc. Am.* **59**, 1155 (1976).
- <sup>2</sup>K. Vokurka, "Oscillations of bubbles generated by imploding and exploding hollow glass spheres in liquids," *Acustica* **68**, 231–240 (1989).
- <sup>3</sup>G. J. Heard, M. McDonald, N. R. Chapman, and L. Jashke, "Underwater light bulb implosions: A useful acoustic source," in *Proceedings of IEEE Oceans '97, 6–9 October 1997, Halifax, N. S.* (Marine Technology Society, Halifax, N.S., 1997).
- <sup>4</sup>B. E. McDonald, Charles W. Holland, and Luigi Troiano, "Scaling laws for implosive ocean acoustic sources," in *Proceedings of the 4th European Conference on Underwater Acoustics, 21–24 September 1998*, edited by A. Alippi and G. B. Cannelli (CNR-IDAC, Rome, 1998), pp. 95–100.
- <sup>5</sup>H. Medwin and C. S. Clay, *Fundamentals of Acoustic Oceanography* (Academic New York, 1998), Chap. 8.
- <sup>6</sup>A. Prosperetti, "Application of the subharmonic threshold to the measurement of the damping of oscillating gas bubbles," *J. Acoust. Soc. Am.* **61**, 11–16 (1977).
- <sup>7</sup>A. Prosperetti, "Thermal effects and damping mechanisms in the forced radial oscillations of gas bubbles in liquids," *J. Acoust. Soc. Am.* **61**, 17–27 (1977).
- <sup>8</sup>A. I. Eller, "Damping constants of pulsating bubbles," *J. Acoust. Soc. Am.* **47**, 1469–1470 (1970).
- <sup>9</sup>M. S. Longuet-Higgins, "Nonlinear damping of bubble oscillations by resonant interaction," *J. Acoust. Soc. Am.* **91**, 1414–22 (1992).
- <sup>10</sup>N. R. Chapman, "Measurement of the waveform parameters of shallow explosive charges," *J. Acoust. Soc. Am.* **78**, 672–681 (1985).
- <sup>11</sup>A. Prosperetti and A. Lezzi, "Bubble dynamics in a compressible liquid. I. First-order theory," *J. Fluid Mech.* **168**, 457 (1986).
- <sup>12</sup>R. N. Sudan, J. Akinrimisi, and D. T. Farley, "Generation of small scale irregularities in the equatorial electrojet," *J. Geophys. Res.* **78**, 240–248 (1973).
- <sup>13</sup>L. D. Landau and E. M. Lifschitz, *Fluid Mechanics* (Pergamon, London, 1959), Sec. 35.
- <sup>14</sup>B. E. McDonald, S. L. Ossakow, S. T. Zalesak, and N. J. Zabusky, "Scale sizes and lifetimes of  $F$  region plasma cloud striations as determined by the condition of marginal stability," *J. Geophys. Res.* **86**, 5775 (1981).
- <sup>15</sup>P. Brandt, A. Rubino, W. Alpers, and J. O. Backhaus, "Internal waves in the Strait of Messina studied by a numerical model and synthetic aperture radar images from the ERS 1/2 satellites," *J. Phys. Oceanogr.* **27**, 647–663 (1997).
- <sup>16</sup>D. L. Book, "Convective stabilization of the Rayleigh–Taylor instability with self-consistent treatment of boundary conditions," *Plasma Phys.* **34**, 737–747 (1992).
- <sup>17</sup>J. P. Dahlburg, J. H. Gardner, G. S. Doolen, and S. W. Haan, "The effect of shape in the three-dimensional Rayleigh–Taylor instability. I. Single mode perturbations," *Phys. Fluids B* **5**, 571–584 (1993).
- <sup>18</sup>H. Lamb, *Hydrodynamics*, 6th ed. (Dover, New York, 1945), Chap. IX.

# Regularized matched-mode processing for source localization

Nicole E. Collison<sup>a)</sup> and Stan E. Dosso<sup>b)</sup>

*School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia V8W 3P6, Canada*

(Received 27 September 1999; accepted for publication 16 February 2000)

This paper develops a new approach to matched-mode processing (MMP) for ocean acoustic source localization. MMP consists of decomposing far-field acoustic data measured at an array of sensors to obtain the excitations of the propagating modes, then matching these with modeled replica excitations computed for a grid of possible source locations. However, modal decomposition can be ill-posed and unstable if the sensor array does not provide an adequate spatial sampling of the acoustic field (i.e., the problem is underdetermined). For such cases, standard decomposition methods yield minimum-norm solutions that are biased towards zero. Although these methods provide a mathematical solution (i.e., a stable solution that fits the data), they may not represent the most physically meaningful solution. The new approach of regularized matched-mode processing (RMMP) carries out an independent modal decomposition prior to comparison with the replica excitations for each grid point, using the replica itself as the *a priori* estimate in a regularized inversion. For grid points at or near the source location, this should provide a more physically meaningful decomposition; at other points, the procedure provides a stable inversion. In this paper, RMMP is compared to standard MMP and matched-field processing for a series of realistic synthetic test cases, including a variety of noise levels and sensor array configurations, as well as the effects of environmental mismatch. © 2000 Acoustical Society of America. [S0001-4966(00)04305-8]

PACS numbers: 43.30.Bp, 43.60.Pt [DLB]

## INTRODUCTION

The problem of localizing an acoustic source in the ocean has received a great deal of attention in recent years. Matched-field processing (MFP) localization matches acoustic pressure fields measured at an array of sensors with modeled replica fields computed for a grid of possible source locations using a numerical propagation model (e.g., Refs. 1–4). The grid point with the highest match is taken to be the estimated source position. Matched-mode processing (MMP) consists of first decomposing the measured fields to obtain the excitations or amplitudes of the constituent propagating modes (e.g., Refs. 4–15); the source location can then be estimated by matching these excitations with a grid of modeled replica excitations.

A potential advantage of MMP is that subsets of the complete mode set can be considered to address problems such as environmental mismatch.<sup>11</sup> For example, in cases where seabed geoacoustic properties are poorly known, the matching can be applied only to low-order modes that interact minimally with the bottom. MFP methods, which consider the total field, cannot be adapted in this manner.

A potential disadvantage of MMP involves the modal decomposition itself, which represents a linear inverse problem that can be ill-posed and unstable (small errors on the data lead to large errors on the solution). In particular, MMP can give poor results when the sensor array provides an inadequate spatial sampling of the acoustic field leading to an underdetermined inverse problem. If the array contains fewer sensors than modes, the highest-order mode functions are

undersampled, causing spatial aliasing, and the decomposition for these modes is singular. Alternatively, if the array aperture spans too small a fraction of the water column, the lower-order modes are poorly sampled, leading to an ill-conditioned inverse problem. Singular value spectra and resolution matrices computed for the modal inverse problem can provide insight into how well individual modes are resolved in a particular scenario.<sup>16</sup>

For a transient source at very long range in deep water, Heaney and Kuperman<sup>15</sup> developed a time-windowed mode filter that overcame the underdeterminacy of modal decomposition by windowing the data in the time domain in such a way that only a well-determined subset of the modes was present in each time window. However, this paper will consider methods of directly solving an underdetermined modal inverse problem.

Inversion methods typically applied to underdetermined modal decompositions (reviewed in Sec. I) are based (explicitly or implicitly) on determining the minimum-norm or “smallest” solution, i.e., the solutions are biased toward zero. Although these methods provide a mathematical solution (a stable solution that fits the data), they may not represent the most physically meaningful solution. This paper develops and tests a new approach, referred to as regularized matched-mode processing (RMMP), which applies particularly to underdetermined problems. RMMP carries out an independent modal decomposition prior to comparison with the replica excitations for each grid point, using the replica itself as the *a priori* estimate in a general regularized formulation of the inverse problem.<sup>16–19</sup> This should provide a more meaningful modal decomposition for grid points at or near the actual source location; at other points, the procedure provides a stable inversion, although it may not be particu-

<sup>a)</sup>Current address: Defense Research Establishment Atlantic, 9 Grove St., PO Box 1012, Dartmouth, NS B2Y 3Z7, Canada.

<sup>b)</sup>Electronic mail: sdosso@uvic.ca

larly physical. At every point of the search grid, RMMP produces the optimal match possible between the decomposed and replica modal excitations, subject to fitting the acoustic data. The conjecture examined here is that the optimal match generally favors the actual source position over other points of the search grid. In this paper, RMMP source localization is illustrated and compared to standard MMP and MFP for a series of synthetic shallow-water test cases involving a variety of sensor array configurations (including undersampled and short-aperture arrays), and several signal-to-noise levels. In addition, the ability to adapt RMMP to environmental mismatch is illustrated for a case where seabed geoacoustic properties are poorly known.

The remainder of this paper is organized as follows. Section IA reviews applications of linear inverse theory to modal decomposition. Section IB develops the new approach of RMMP. Section II A compares source localization results for RMMP, MMP, and MFP for a known ocean environment, while Sec. IIB considers the three localization methods in the presence of environmental mismatch. Finally, Sec. III summarizes the paper.

## I. THEORY

### A. Modal decomposition

In the far field, the normal-mode model for the acoustic pressure signal  $s$  at depth  $z$  due to a source at range  $r$  and depth  $z_s$  can be written<sup>20</sup>

$$s(z) = b \sum_{m=1}^M \phi_m(z) \phi_m^*(z_s) \frac{e^{ik_m r}}{\sqrt{k_m r}}, \quad (1)$$

where  $\phi_m$  and  $k_m$  represents the  $m$ th mode function and wave number, respectively,  $M$  is the total number of propagating modes, and  $b = e^{i\pi/4} \sqrt{2\pi/\rho(z_s)}$  where  $\rho$  is density (\* represents complex conjugation). The mode functions are eigenfunction solutions of the depth-separated wave equation, and form an orthonormal set according to<sup>20</sup>

$$\int_0^\infty \phi_m(z) \phi_n^*(z) / \rho(z) dz = \begin{cases} 0 & \text{if } m \neq n, \\ 1 & \text{if } m = n. \end{cases} \quad (2)$$

If the signal is recorded at a vertical line array (VLA) of  $N$  sensors, (1) can be written as a linear matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{s}, \quad (3)$$

where  $\mathbf{s}$  is the signal vector

$$\mathbf{s} = [s(z_1), \dots, s(z_N)]^T, \quad (4)$$

$\mathbf{A}$  is an  $N \times M$  matrix with columns consisting of the sampled mode functions

$$\mathbf{A} = \begin{bmatrix} \phi_1(z_1) & \cdots & \phi_M(z_1) \\ \vdots & \ddots & \vdots \\ \phi_1(z_N) & \cdots & \phi_M(z_N) \end{bmatrix}, \quad (5)$$

and  $\mathbf{x}$  is a vector of the mode excitations or amplitudes at the receivers

$$\mathbf{x} = b \left[ \phi_1(z_s) \frac{\exp[ik_1 r]}{\sqrt{k_1 r}}, \dots, \phi_M(z_s) \frac{\exp[ik_M r]}{\sqrt{k_M r}} \right]^T \quad (6)$$

( $T$  denotes transpose). Note that  $\mathbf{x}$  contains all the information regarding source location. In practical cases, the acoustic signal  $\mathbf{s}$  is contaminated with noise  $\mathbf{n}$ , so (3) becomes

$$\mathbf{A}\tilde{\mathbf{x}} = \mathbf{s} + \mathbf{n} \equiv \mathbf{p}, \quad (7)$$

where  $\mathbf{p}$  is the vector of measured acoustic pressures (signal plus noise)

$$\mathbf{p} = [p(z_1), \dots, p(z_N)]^T. \quad (8)$$

Modal decomposition (also referred to as mode filtering) consists of solving (7) for an estimate  $\tilde{\mathbf{x}}$  of the true mode excitations  $\mathbf{x}$ , which represents a discrete, linear inverse problem.

According to (2), if the mode functions  $\phi_m(z)$  are adequately sampled over their entire extent (i.e., to a depth where they have decayed to essentially zero) and density is close to unity, the matrix  $\mathbf{A}$  will be approximately orthogonal (Hermitian), i.e.,

$$\mathbf{A}^\dagger \mathbf{A} \approx \mathbf{I}, \quad (9)$$

where  $\mathbf{I}$  is the identity matrix and  $^\dagger$  denotes conjugate transpose. The simplest approach to modal decomposition is based on modal orthogonality and applies  $\mathbf{A}^\dagger$  as a generalized inverse, known as the sampled mode-shapes (SMS) mode filter, to (7) to obtain<sup>5,6</sup>

$$\tilde{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{p}. \quad (10)$$

Buck *et al.*<sup>14</sup> note that the SMS generalized inverse is optimal for detecting a single mode in spatially white noise, but is nonoptimal for more than one mode when the spatial sampling does not preserve mode orthogonality. They suggest normalizing  $\mathbf{A}^\dagger$  by applying a weighting matrix  $\mathbf{W}$  so that the diagonal elements of  $\mathbf{W}\mathbf{A}^\dagger \mathbf{A}$  are unity, where

$$\mathbf{W} = \text{diag}[\|\phi_1\|^{-2}, \dots, \|\phi_M\|^{-2}], \quad (11)$$

and  $\|\phi_j\|^2 = \int_0^\infty |\phi_j(z)|^2 / \rho(z) dz$ . Unfortunately, in practical cases, an adequate sampling of the modes to ensure orthogonality is often not possible, particularly in cases of an acoustically penetrable seabed where the mode functions extend into the bottom. In these cases, the SMS inverse can perform poorly.

To consider the modal inverse problem (7) in more detail, assume that the noise  $n_i$  on measurement  $p_i$  is due to an independent, Gaussian-distributed random process with zero mean and standard deviation  $\sigma_i$ . The least-squares (maximum-likelihood) solution is determined by minimizing the  $\chi^2$  misfit<sup>16-19</sup>

$$\chi^2 = |\mathbf{G}(\mathbf{A}\tilde{\mathbf{x}} - \mathbf{p})|^2 = (\mathbf{A}\tilde{\mathbf{x}} - \mathbf{p})^\dagger \mathbf{G}^\dagger \mathbf{G} (\mathbf{A}\tilde{\mathbf{x}} - \mathbf{p}), \quad (12)$$

where the matrix  $\mathbf{G} = \text{diag}[1/\sigma_1, \dots, 1/\sigma_N]$  weights the data equations according to their uncertainties (for correlated noise,  $\mathbf{G}^\dagger \mathbf{G}$  is replaced by the inverse of the data covariance matrix). Setting  $\partial \chi^2 / \partial \tilde{\mathbf{x}} = 0$  for a minimum yields the solution

$$\tilde{\mathbf{x}} = [\mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{G} \mathbf{A}]^{-1} \mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{G} \mathbf{p}. \quad (13)$$

Note that if  $\mathbf{A}$  is orthogonal and all data have the same uncertainty ( $\mathbf{G} = \sigma \mathbf{I}$ ), the least-squares solution (13) reduces to the SMS solution (10). The least-squares approach provides

an unbiased estimate of the true solution (i.e., the solution to the noise-free problem) provided the matrix in square brackets is nonsingular. In addition to the possibility of singularity, the matrix can be ill-conditioned, leading to an unstable inversion. In modal decomposition, singularity and ill-conditioning are a common result of an inadequate sensor array. If the array contains fewer sensors than modes, the higher-order mode functions will be spatially aliased, and the modal decomposition is singular; if the array spans too small a fraction of the water column, the lower-order modes will be poorly sampled, leading to an ill-conditioned inverse problem (illustrated in Sec. II).

A common approach for solving ill-posed inversions is based on singular-value decomposition (SVD).<sup>18</sup> The SVD of  $\mathbf{GA}$  is given by

$$\mathbf{GA} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^\dagger, \quad (14)$$

where  $\mathbf{U}$  and  $\mathbf{V}$  are  $N \times M$  and  $M \times M$  matrices, respectively, whose columns are orthonormal eigenvectors, and  $\mathbf{\Lambda}$  is an  $M \times M$  diagonal matrix of singular values ordered according to decreasing magnitude. Using (14) in (13), the SVD solution to the linear inverse problem is given by

$$\tilde{\mathbf{x}} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}^\dagger\mathbf{Gp} = \sum_{m=1}^M \mathbf{v}_m \mathbf{u}_m^\dagger \mathbf{Gp} / \lambda_m, \quad (15)$$

where the  $\mathbf{v}_m$  and  $\mathbf{u}_m$  represent the  $m$ th columns of  $\mathbf{V}$  and  $\mathbf{U}$ , respectively. Equation (15) is not defined if one or more  $\lambda_m$  are zero, which corresponds to a singular matrix in (13). An ill-conditioned matrix is characterized by one or more  $\lambda_m$  being very small. Equation (15) indicates that small  $\lambda_m$  can cause the errors on the data  $\mathbf{p}$  to have a greatly magnified effect on the solution  $\tilde{\mathbf{x}}$  (i.e., small  $\lambda_m$  cause instability). Yang<sup>8</sup> and Voronovich *et al.*<sup>12</sup> applied SVD to modal decomposition, but decomposed  $\mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{GA}$  instead of  $\mathbf{GA}$ . However, the SVD of  $\mathbf{GA}$  is recommended, because its singular values are the square roots of the singular values of  $\mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{GA}$ , which are more numerically stable for small singular values.<sup>18</sup> To stabilize singular or ill-posed inversions, the reciprocals of zero or small singular values in (15) can be set to zero, thereby removing their effect on the inversion. If  $M - Q$  small singular values are omitted, (15) becomes

$$\tilde{\mathbf{x}} = \sum_{m=1}^Q \mathbf{v}_m \mathbf{u}_m^\dagger \mathbf{Gp} / \lambda_m, \quad Q \leq M. \quad (16)$$

How small  $\lambda_m$  needs to be in order to be omitted from the inversion can be somewhat arbitrary. Omitting singular values generally decreases the solution variance at the expense of a decrease in parameter resolution. Some care is required to obtain an optimal tradeoff between solution accuracy and resolution. In an Arctic application of MMP, Yang<sup>8</sup> noted that the singular values divided naturally into two groups, with the singular values in one group being several orders of magnitude larger than those in the other. He proposed that the large singular values correspond to eigenvectors that span the mode ‘‘signal space’’ while the eigenvectors associated with the small singular values span the mode ‘‘noise space,’’ providing a physical basis for neglecting the small singular values. However, when the singular values do not

divide naturally into groups, deciding which ones to omit is more difficult. An objective approach, applied in geophysical inverse theory, is to require that the data be fit to a statistically appropriate level, for example, to omit small  $\lambda_m$  until the  $\chi^2$  misfit is approximately equal to its expected value,  $\chi^2 \approx \langle \chi^2 \rangle$ .<sup>21,22</sup> For  $N$  real data with Gaussian errors  $\langle \chi^2 \rangle = N$ ; for  $N$  complex data with errors on real and imaginary parts  $\langle \chi^2 \rangle = 2N$  (assumed in this paper). Note that the formulation in (16) applies to both over- and underdetermined inversions, although for  $Q < M$ , the solution is not unbiased. Removing singular values stabilizes the inversion by implicitly determining the ‘‘smallest’’ solution in the sense that  $|\tilde{\mathbf{x}}|$  is biased to be as close to zero as possible.<sup>18</sup>

Another approach to ill-posed inversions is zeroth-order or smallest regularization, which is based on minimizing an objective function  $\Psi$  that combines a term representing the data mismatch and a term representing the magnitude of the solution<sup>18</sup>

$$\Psi = |\mathbf{G}(\mathbf{A}\tilde{\mathbf{x}} - \mathbf{p})|^2 + \theta |\mathbf{H}\tilde{\mathbf{x}}|^2. \quad (17)$$

In (17),  $\mathbf{H}$  represents an arbitrary weighting, known as the regularization matrix, and  $\theta$  is a tradeoff parameter that controls the relative importance of the two terms in the minimization. Minimizing  $\Psi$  with respect to  $\tilde{\mathbf{x}}$  leads to

$$\tilde{\mathbf{x}} = [\mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{GA} + \theta \mathbf{H}^\dagger \mathbf{H}]^{-1} \mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{Gp}. \quad (18)$$

For the simple case of an identity weighting,  $\mathbf{H} = \mathbf{I}$ , (18) can be expressed as

$$\tilde{\mathbf{x}} = \sum_{m=1}^M \mathbf{v}_m \mathbf{u}_m^\dagger \mathbf{Gp} \lambda_m / (\lambda_m^2 + \theta). \quad (19)$$

The factor  $\lambda_m / (\lambda_m^2 + \theta)$  in (19) replaces  $1/\lambda_m$  in (15). For an appropriate choice of  $\theta$  in (19), this factor is small for small  $\lambda_m$  (rather than large), and a stable inversion is achieved. The tradeoff parameter  $\theta$  can be set so as to just stabilize the inversion; alternatively,  $\theta$  can be set to achieve the expected  $\chi^2$  misfit to the data (see the Appendix). Timonov<sup>13</sup> and Buck *et al.*<sup>14</sup> considered a smallest regularized inversion using the identity weighting. Voronovich *et al.*<sup>12</sup> considered two forms of smallest regularization, one with  $\mathbf{H} = \mathbf{I}$ , and the other with  $\mathbf{H} = \text{diag}[|\phi_1(z_s)|^{-1}, \dots, |\phi_M(z_s)|^{-1}]$  so as to concentrate the minimization on the mode amplitudes whose excitations at the source are small.

In the case where there are fewer sensors than modes, the modal decomposition is underdetermined, with  $M - N$  zero singular values indicating spatially aliased modes. The singularity can be overcome using SVD or smallest regularization, as above. However, Yang<sup>8,11</sup> suggested that an alternative approach is to omit the aliased high-order modes from the formulation (effectively, set  $M = N$ ) and then invert the resulting square mode matrix. This method only inverts the properly sampled low-order modes and removes the singularity. However, omitting the aliased modes from the inversion effectively delegates them to be part of the noise, thereby reducing the signal-to-noise ratio (SNR) by an unknown amount.

Finally, a useful measure of how well individual mode excitations can be resolved for a particular scenario is given by the resolution matrix<sup>16</sup> associated with the linear inverse

problem (7). Let  $\mathbf{A}^{-g}$  represent the generalized inverse associated with any of the inversion methods outlined in this section. Noting that  $\langle \mathbf{p} \rangle = \mathbf{s} = \mathbf{A}(\tilde{\mathbf{x}})$  from (7), it follows that for the noise-free case  $\tilde{\mathbf{x}} = \mathbf{A}^{-g}\mathbf{s} = \mathbf{A}^{-g}\mathbf{A}(\tilde{\mathbf{x}})$ , or

$$\tilde{\mathbf{x}} = \mathbf{R}(\tilde{\mathbf{x}}), \quad (20)$$

where  $\mathbf{R} \equiv \mathbf{A}^{-g}\mathbf{A}$  is defined to be the resolution matrix. If  $\mathbf{R} = \mathbf{I}$ , then each parameter of the solution is perfectly resolved. However, if  $\mathbf{R}$  has significant nonzero off-diagonal terms, the parameters of  $\tilde{\mathbf{x}}$  represent weighted averages of the expected parameters, and cannot be individually resolved. Resolution matrices for several different VLA configurations are illustrated in Sec. II.

## B. Regularized matched-mode processing

The previous section described SVD and smallest regularization methods which allow modal decomposition to be applied to ill-posed inversions resulting from inadequate sensor arrays. However, both approaches yield the smallest solution (minimum  $|\tilde{\mathbf{x}}|$ ) for a given level of misfit. In essence, these methods apply a trivial *a priori* estimate of zero modal excitations to the decomposition. Although these methods produce a mathematical solution, they do not necessarily represent the most physically meaningful solution, as there is no physical reason to bias  $\tilde{\mathbf{x}}$  towards zero.<sup>14</sup>

A more general approach to regularized inversion applies a nontrivial *a priori* estimate  $\hat{\mathbf{x}}$  by minimizing an objective function  $\Psi$  that combines the data mismatch and a regularizing term incorporating the prior estimate<sup>16-19</sup>

$$\Psi = |\mathbf{G}(\mathbf{A}\tilde{\mathbf{x}} - \mathbf{p})|^2 + \theta |\mathbf{H}(\tilde{\mathbf{x}} - \hat{\mathbf{x}})|^2. \quad (21)$$

In (21),  $\mathbf{H}$  again represents an arbitrary weighting matrix, and  $\theta$  is a tradeoff parameter that controls the relative importance of the data misfit and *a priori* estimate in the minimization. Minimizing  $\Psi$  with respect to  $\tilde{\mathbf{x}}$  leads to

$$\tilde{\mathbf{x}} = \hat{\mathbf{x}} + [\mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{G} \mathbf{A} + \theta \mathbf{H}^\dagger \mathbf{H}]^{-1} \mathbf{A}^\dagger \mathbf{G}^\dagger \mathbf{G}(\mathbf{p} - \mathbf{A}\hat{\mathbf{x}}). \quad (22)$$

An objective choice for  $\theta$  is the value that produces a  $\chi^2$  data misfit equal to the expected value, effectively applying the *a priori* information subject to achieving a statistically meaningful fit to the data. The methodology of efficiently determining  $\theta$  is given in the Appendix.

In applying regularized inversion to modal decomposition, the obvious question is how to obtain a meaningful *a priori* estimate  $\hat{\mathbf{x}}$ . The idea underlying the new approach of RMMP developed here is to make use of the modeled replica mode excitations that are calculated for the MMP search grid. In MMP source localization, the replica excitations for each point of the search grid are computed via a normal-mode forward model, which provides a unique, stable, noise-free solution. Thus, if the source is actually located at a particular grid point, the replica excitations computed for that point provide an ideal *a priori* estimate for the modal decomposition. This observation provides the basis for RMMP: an independent regularized inversion of the measured acoustic fields for the mode excitations is carried out prior to matching with the replica excitations for each grid point, using the replica itself as the prior estimate  $\hat{\mathbf{x}}$ . For grid points at or near the actual source location, using the corresponding replica as

an *a priori* estimate and minimizing  $|\mathbf{H}(\tilde{\mathbf{x}} - \hat{\mathbf{x}})|$  should provide a more meaningful solution to the modal decomposition inverse problem than minimizing  $|\tilde{\mathbf{x}}|$  or  $|\mathbf{H}\tilde{\mathbf{x}}|$  (as per SVD or smallest-regularized inversions, respectively). For grid points away from the source location, this procedure still provides a stable inversion, although the regularization may not be particularly physical. The regularization matrix  $\mathbf{H}$  can be taken to be the identity matrix, or  $\mathbf{H}$  can represent a weighting designed to concentrate the regularization on specific modes (considered in Sec. II). At every point of the search grid, RMMP produces the maximum match possible (subject to fitting the acoustic data) between the replica and the constructed modal excitations. The conjecture examined in this paper is that the procedure of maximizing all matches generally favors the source position over other points of the search grid. Since RMMP requires an independent modal inversion prior to matching excitations at each grid point, it is important that these inversions be carried out in an efficient manner, as described in the Appendix.

The use of a nontrivial *a priori* estimate  $\hat{\mathbf{x}}$  in regularized inversion requires a consideration of the treatment of magnitude and phase information. Matched-field and standard matched-mode approaches to source localization are based on matching the relative magnitude and phase between the sensors of an array, since the absolute magnitude and phase of the source are typically unknown. This is commonly done using the Bartlett (linear) processor

$$B = \frac{|\mathbf{d}^\dagger \mathbf{q}(r, z)|^2}{|\mathbf{d}|^2 |\mathbf{q}(r, z)|^2}, \quad (23)$$

where  $\mathbf{d}$  represents either the measured field or its constituent mode excitations, and  $\mathbf{q}(r, z)$  represents either the replica fields or replica excitations for matched-field and matched-mode methods, respectively. The Bartlett processor is based on the projection of the modeled replicas onto the measured data (or its modal components), and as such, depends only on the relative magnitude and phase from sensor to sensor over the array. However, for a nonzero  $\hat{\mathbf{x}}$ , the regularization objective function (21), which is based on  $l_2$  difference norms, is potentially dependent on absolute magnitude and phase. Thus, the prior estimate must be normalized (in both magnitude and phase) to remove this dependence. An appropriate normalization can be derived as follows.

By considering (21), it is apparent that for the data  $\mathbf{p}$  and the *a priori* estimate  $\hat{\mathbf{x}}$  to be consistent,  $\mathbf{p}$  and  $\hat{\mathbf{p}} \equiv \mathbf{A}\hat{\mathbf{x}}$  must have the same magnitude and phase. (In practical applications, the data and prior estimate will generally not be completely consistent; nonetheless, normalizations based on this goal remove the arbitrary nature of the magnitude and phase of  $\mathbf{p}$  and  $\hat{\mathbf{x}}$ .) For the following discussion, let  $\hat{\mathbf{x}}'$  represent the prior estimate (computed using a normal mode model) before normalization, and  $\hat{\mathbf{x}}$  represent the normalized value used in (21) and (22). An appropriate magnitude normalization is to scale  $\hat{\mathbf{x}}'$  so that  $\mathbf{A}\hat{\mathbf{x}}' = \hat{\mathbf{p}}'$  has the same magnitude as  $\mathbf{p}$ . The required phase normalization can be accomplished by rotating  $\hat{\mathbf{x}}'$  so that the phases of  $\mathbf{p}$  and  $\hat{\mathbf{p}}$  are as similar as possible. Let the phases of the  $j$ th element of  $\mathbf{p}$  and  $\mathbf{A}\hat{\mathbf{x}}'$  be given by  $\psi_j$  and  $\hat{\psi}_j$ , respectively. The rotation  $\beta$  that mini-

minimizes the squared phase difference across the array  $\Delta = \sum_{j=1}^N (\psi_j - \hat{\psi}_j - \beta)^2$ , found by setting  $\partial\Delta/\partial\beta=0$ , is given by

$$\beta = \frac{1}{N} \sum_{j=1}^N (\psi_j - \hat{\psi}_j) \quad (24)$$

(note that this phase rotation simply corresponds to the average phase difference over the array). Thus, the normalization

$$\hat{\mathbf{x}} = \frac{|\mathbf{p}|}{|\mathbf{A}\hat{\mathbf{x}}'|} e^{i\beta\hat{\mathbf{x}}'} \quad (25)$$

produces the *a priori* estimate  $\hat{\mathbf{x}}$  that is as consistent as possible (in magnitude and phase) with the acoustic data  $\mathbf{p}$ . This has the effect of removing absolute magnitude and phase information from the inversion, while retaining relative (sensor-to-sensor) information. The phase rotation derived here minimizes the phase difference over all sensors in the array, and should be less sensitive to noise than basing the rotation on any single sensor, as is commonly done. An alternative normalization can be derived by seeking the complex scalar  $\tau$  that minimizes  $|\tau\hat{\mathbf{p}}' - \mathbf{p}|^2$ , which has the solution  $\tau = \hat{\mathbf{p}}'^{\dagger}\mathbf{p}/|\hat{\mathbf{p}}'|^2$ ; however, this procedure allows undesirable cases such as  $\hat{\mathbf{p}} = \mathbf{0}$  and is not followed here.

## II. EXAMPLES

### A. Localization in a known environment

For the purposes of this paper, the ultimate test for a new approach to modal decomposition, such as RMMP, is how well it performs in source localization. This and the following section compare RMMP, MMP, and MFP localization results for a series of realistic synthetic test cases. The MMP algorithm is based on a smallest-regularized modal decomposition that achieves the expected  $\chi^2$  misfit and omits aliased modes (as described in Sec. II A), since a previous study of the various MMP techniques indicated this approach yielded the best results on similar test cases.<sup>23,24</sup> The RMMP algorithm implements the theory developed in Sec. I B and the Appendix, retaining aliased modes. In this section, an identity regularization matrix is applied for both MMP and RMMP. For all three methods, the match between data (acoustic fields or modal excitations) and replicas is quantified using the Bartlett processor (23).

The test cases are based on the shallow-water environment illustrated in Fig. 1. The environment consists of a 300-m water column with a typical Northeast Pacific continental shelf sound-speed profile<sup>25</sup> overlying a 50-m-thick sediment layer and semi-infinite basement. The sediment layer has a compressional speed of  $c_p = 1650$  m/s, shear speed of  $c_s = 300$  m/s, density of  $\rho = 1.6$  g/cm<sup>3</sup>, and compressional and shear attenuations of  $\alpha = 0.3$  dB/ $\lambda$  (where  $\lambda$  is the acoustic wavelength). The basement has  $c_p = 2000$  m/s,  $c_s = 800$  m/s,  $\rho = 2.1$  g/cm<sup>3</sup>, and  $\alpha = 0.5$  dB/ $\lambda$ . This environment supports 12 propagating modes at a source frequency of 40 Hz, as shown in Fig. 2. Synthetic acoustic data were computed for this environment for a source at  $(r, z) = (6 \text{ km}, 100 \text{ m})$  using the wave-number integration propaga-

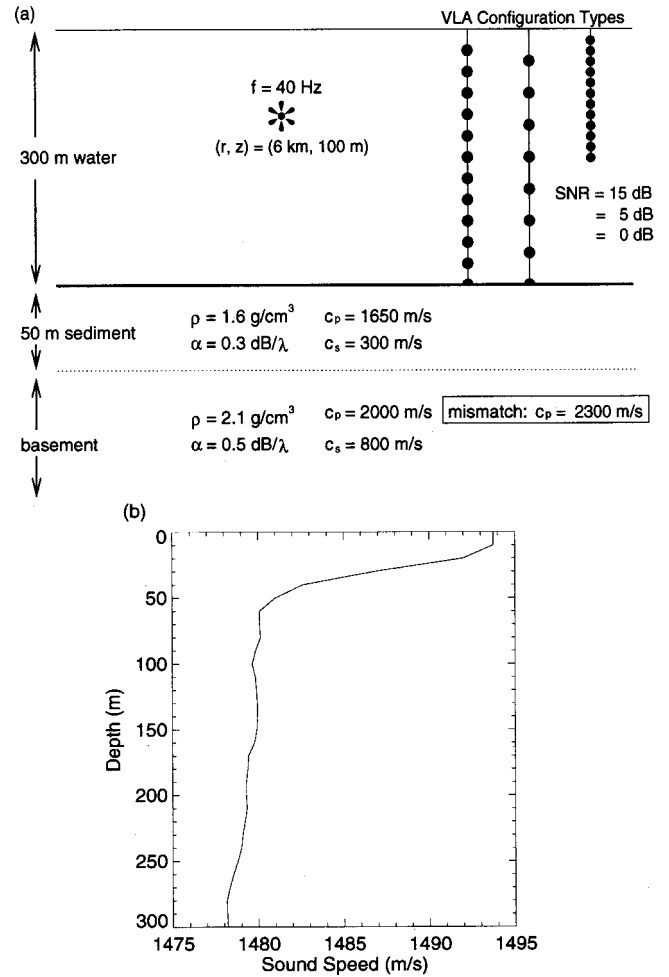


FIG. 1. (a) Schematic diagram of the shallow-water environment assumed for the test cases (parameters defined in the text). The mismatch in basement compressional speed is considered in Sec. II B. (b) Ocean sound-speed profile for test cases.

tion model SAFARI,<sup>26</sup> and subsequently adding random (Gaussian) noise to a designated SNR. The replica acoustic fields (for MFP) and replica mode excitations (for MMP and RMMP) were computed using the normal-mode model ORCA.<sup>27</sup> In this section, cases are considered in which the replicas were computed with exact knowledge of the environmental parameters. Section II B considers the case of localization with environmental mismatch in the basement properties.

In compiling the source localization results, three different SNRs of 15, 5, and 0 dB were considered, where the (per sensor) SNR is defined

$$\text{SNR} = 10 \log_{10} \left[ \frac{|s|^2}{|\mathbf{n}|^2} \right] \quad (26)$$

(note that no averaging of the acoustic fields was applied). Source localization results were also compiled for a number of different VLA configurations that sample the 12 modes in various ways: (i) well-sampled, with 12 sensors equally spaced over the entire water column; (ii) undersampled, with 6–11 sensors equally spaced over the water column; and (iii) short-aperture, with 12 sensors spanning the top 0.5–0.9 of the water column.



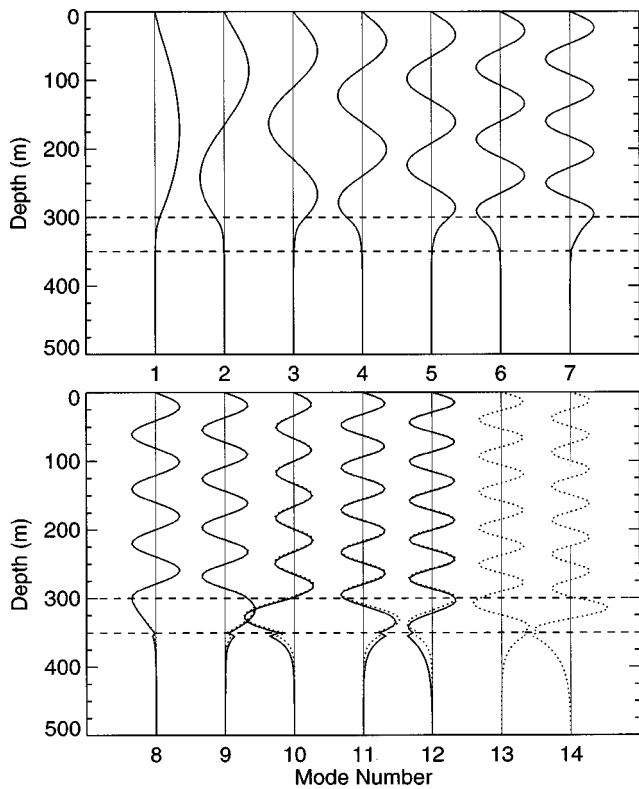


FIG. 2. Mode functions for the true environment (solid lines), and for the mismatched environment (dotted lines). Dashed lines denote the water-sediment and sediment-basement interfaces.

Before examining the localization results, it is interesting to consider the singular-value spectra and modal resolution for the various array configurations. Singular-value spectra of the mode matrices are shown in Fig. 3 for the well-sampled and undersampled cases. For the well-sampled case, Fig. 3(a) shows that all singular values are within one order of magnitude, indicating that sufficient linearly independent information is available for all modes, and hence the modal inversion is well-conditioned and stable. Figure 3(b)–(g) show that each time the number of sensors is reduced by one, a zero singular value occurs, indicating aliased modes and a singular modal inverse problem. Figure 4 shows singular value spectra for the short-aperture cases. In this case, no zero singular values occur. However, as the aperture of the array is reduced, the number of small singular values increases and the magnitude of the small singular values decreases, indicating correspondingly less linearly independent information and an ill-conditioned modal inversion. Figure 5 shows modal resolution matrices for selected cases, computed for the smallest regularized solution at SNR=15 dB, retaining all modes in each case. Figure 5(a) shows that the well-sampled case provides excellent resolution of the excitations of the lower-order modes, although the higher-order mode excitations are somewhat less well resolved, due to the penetration of the corresponding mode functions into the seabed (Fig. 2). It is interesting to note that mode 10, the most poorly resolved mode in Fig. 5(a), also has the largest amplitude in the sub-bottom in Fig. 2. Figure 5(b) shows the resolution matrix for the undersampled case with eight sensors. In this case, the seven highest-order modes are aliased,

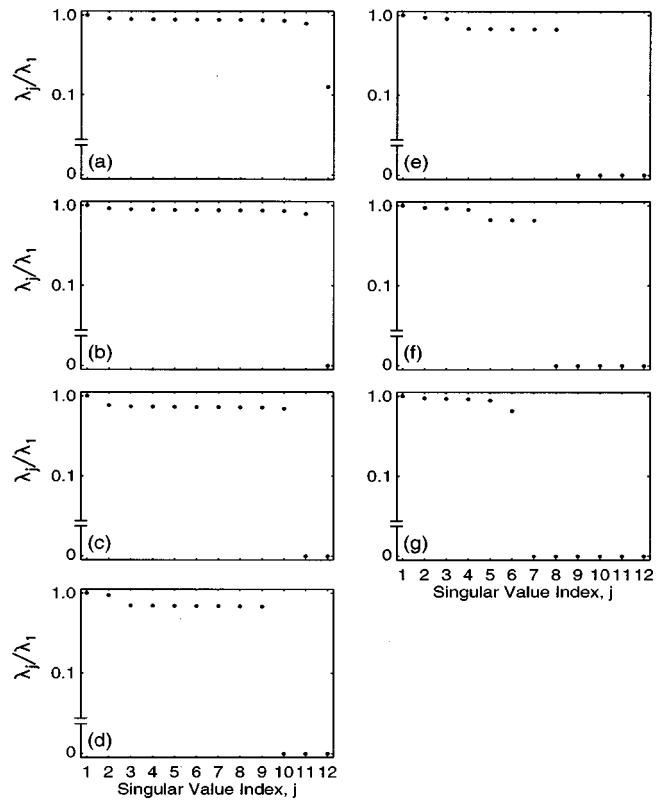


FIG. 3. Singular-value spectra (normalized by the largest value  $\lambda_1$ ) of the mode matrix for VLAs consisting of 12–6 sensors in (a)–(g), respectively. Note the discontinuity on the vertical axes.

resulting in complete ambiguity between the corresponding mode excitations. Figure 5(c) shows the resolution matrix for a VLA spanning the top 0.7 of the water column. In comparing Fig. 5(c) and (a), it is apparent that the short aperture array leads to a general decrease in resolution, particularly of

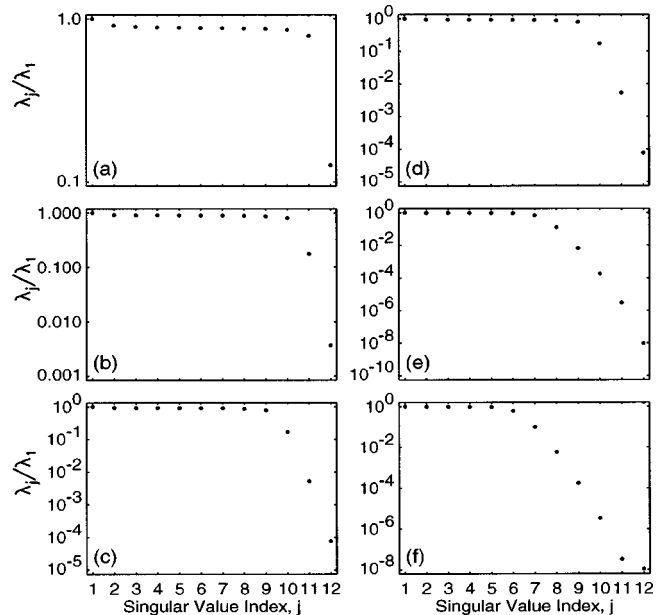


FIG. 4. Singular-value spectra (normalized by the largest value  $\lambda_1$ ) of the mode matrix for VLAs spanning 1.0–0.5 of the water column are shown in (a)–(f), respectively.

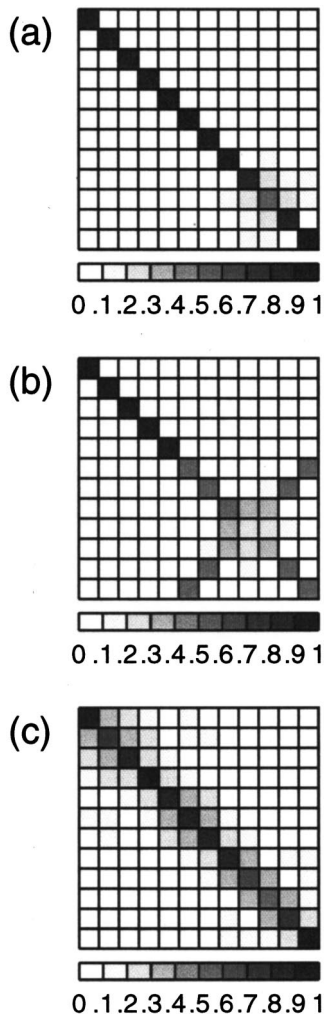


FIG. 5. Resolution matrices (normalized by largest element) for smallest-regularized solution and (a) well-sampled case, (b) undersampled case with eight-sensor VLA, and (c) short-aperture case with VLA spanning 0.7 of the water column.

the lower-order mode excitations, although the resolution matrix is still diagonally dominant.

An example of source localization for this test case is given in Fig. 6, which shows ambiguity surfaces for RMMP, MMP, and MFP for the well-sampled case [Fig. 6(a)–(c)] for an undersampled case with eight sensors [Fig. 6(d)–(e)], and for a short-aperture case with a VLA spanning the top 0.7 of the water column [Fig. 6(f)–(g)]. In each case, the SNR=5 dB. The search grid employed extended from 0–12 km in range with a range increment of 100 m, and from 0–300 m in depth with a depth increment of 10 m (i.e., a total of 3600 grid points). The estimated source location corresponds to the grid point at which the Bartlett match was a maximum. For the well-sampled cases, all three methods unambiguously localize the source at the correct position of (6 km, 100 m). For the undersampled cases, the highest match occurs at a strong sidelobe at approximately (9.5 km, 225 m) for all three methods. For the short-aperture case, a more complicated sidelobe pattern is evident, and RMMP and MFP localize the source correctly, while MMP does not. It is interesting to note that for all methods, the sidelobe structure of the ambiguity surfaces is similar; however, RMMP obtains higher matches at all positions as it represents an optimization process.

The localization results in Fig. 6 are for a single realization of random noise on the data. To obtain meaningful general comparisons between the three localization algorithms, 100 independent acoustic data sets  $\mathbf{p}$  were computed by adding 100 different random noise realizations  $\mathbf{n}$  to the acoustic signal  $\mathbf{s}$  for a source located at  $(r, z) = (6 \text{ km}, 100 \text{ m})$ . The relative performance of the three localization algorithms is quantified by the estimated probability of correct localization  $\hat{P}$ , taken to be the fraction of times (for the 100 noisy data sets) that the estimated source location is within a suitably small region about the true source location, defined here by  $\pm 200 \text{ m}$  in range and  $\pm 10 \text{ m}$  in depth (Tabrikian *et al.*<sup>28</sup> and

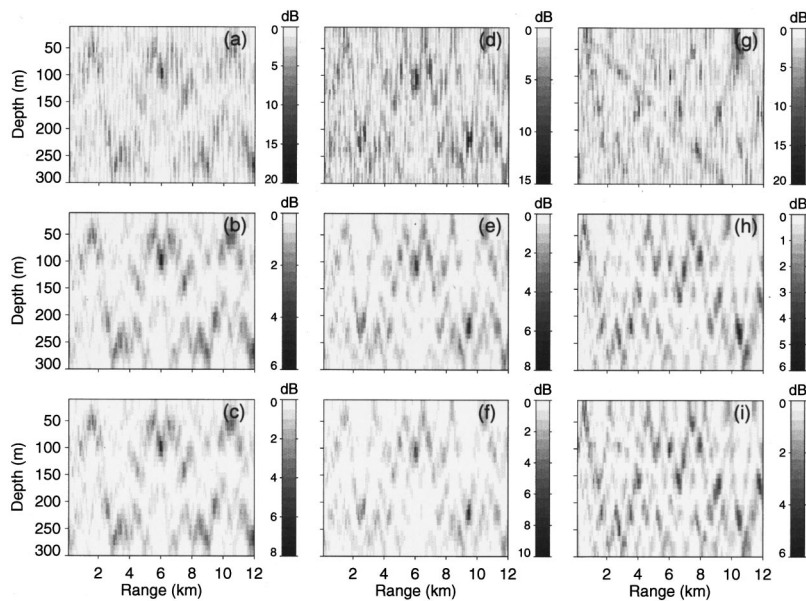


FIG. 6. Ambiguity surfaces for RMMP (top row), MMP (center row), and MFP (bottom row) for the well-sampled case (left column), undersampled case of an eight-sensor VLA (center column), short-array case with aperture 0.7 of the water column (right column). In each case the SNR=5 dB.

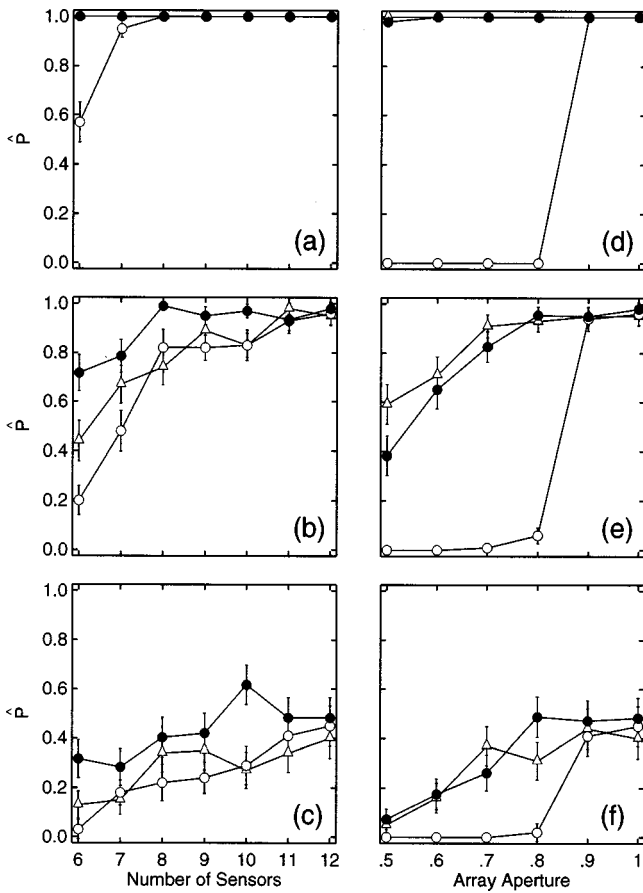


FIG. 7. Estimated probability of correct localization  $\hat{P}$  with known environmental parameters for MFP (triangles), MMP (open circles), and RMMP (filled circles). (a)–(c) give results for undersampled cases at SNRs of 15, 5, and 0 dB, respectively; (d)–(f) give results for short-aperture cases at the same SNRs. Error bars denote 90% confidence intervals.

Michalopoulou<sup>29</sup> used similar criteria). Note that since RMMP produces the optimal match (with respect to the noise) at each grid point, it is possible for one or more perfect matches to occur in an ambiguity surface. In cases where more than one perfect match occurs, the localization is assigned a fractional value depending on how many of the perfect matches are located within the acceptable region (e.g., two-thirds when two of three perfect matches are within  $\pm 200$ -m range,  $\pm 10$ -m depth of the true location). Ninety-percent confidence intervals for the probability of correct localization are computed using<sup>30</sup>

$$\hat{P} \pm 1.645 \sqrt{\hat{P}(1-\hat{P})/n}, \quad (27)$$

where  $n = 100$  for the 100 noise realizations.

Figure 7 shows the probability of correct localization results for RMMP, MMP, and MFP obtained for the known environment, with error bars denoting 90% confidence limits. Figure 7(a)–(c) shows  $\hat{P}$  for the well-sampled case and undersampled cases consisting of 6–11 sensors for SNRs of 15, 5, and 0 dB. Figure 7(d)–(e) shows  $\hat{P}$  for the well-sampled case and for short-array cases consisting of 0.5–0.9 of the water column for the same SNRs. In all cases, the probabilities of correct localization decrease with SNR. Figure 7(a) shows that at SNR=15 dB with 7–12 sensors, all methods yield a high probability of correct source localiza-

tion; however, with a 6-sensor VLA, MMP has a substantially lower probability than RMMP and MFP. Figure 7(b) shows that at SNR=5 dB, all methods localize quite well for eight or more sensors. For 6–11 sensors, RMMP has consistently higher  $\hat{P}$  values than the other two methods, with substantial improvements for  $\leq 8$  sensors. Figure 7(c) shows that for the undersampled cases at SNR=0 dB, RMMP consistently localizes at a higher level than MFP and MMP. For array apertures of 0.9–1.0 of the water column, Fig. 7(d)–(f) shows that all methods have similar probabilities of localization. Figure 7(e) shows that at SNR=5 dB, MFP produces slightly better overall localization results than RMMP, while at SNR=0 dB [Fig. 7(f)] RMMP produces slightly better results than MFP. For apertures of 0.5–0.8, MMP performs very poorly at all SNRs for the short-aperture test cases.

The probability of correct localization results shown in Fig. 7 compares the relative success of RMMP, MMP, and MFP for 100 noisy data sets and a number of VLA configurations. Another way to illustrate the performance of the various methods for selected cases is using probability ambiguity surfaces (PAS), which show the estimated source positions for each of the 100 localizations. Figure 8(a)–(c) show PAS for a six-sensor VLA for RMMP, MMP, and MFP, respectively, at SNR=5 dB. These figures show that RMMP has the highest probability of correct localization [as in Fig. 7(b)], and that in each case, the majority of the incorrect localizations for all methods concentrates in similar regions which correspond to sidelobes of the individual ambiguity surfaces shown in Fig. 6. Figure 8(d)–(f) show PAS for a 0.6-aperture VLA for the three methods. It is interesting to note that the poor performance of MMP for this case is due almost entirely to incorrect localizations at a single strong sidelobe.

## B. Localization with environmental mismatch

To this point, it has been assumed that the properties of the ocean environment are known perfectly for all localization examples. Of course, this is never the case in practical applications. In particular, seabed geoacoustic properties (especially basement parameters) are difficult and expensive to measure, and are often only approximately known. Errors in the environmental parameters, referred to as environmental mismatch, affect both matched-field and matched-mode methods. In MFP, the modeled replica fields are not exactly correct. In matched-mode methods, both the modal decomposition and the replica modal excitations are inexact. However, in some cases, matched-mode methods can be adapted to environmental mismatch by omitting the modes that are most strongly affected by the mismatch. MFP, which considers the total field, cannot be adapted in this way.

The synthetic acoustic data considered in this section were created using the ocean environment in shown in Fig. 1; however, all replicas are computed for an environment that included a mismatch in the basement compressional speed of 300 m/s (i.e., using 2300 m/s instead of 2000 m/s). An acoustic source in this mismatched environment produces 14 (rather than 12) propagating modes, shown in Fig. 2; in

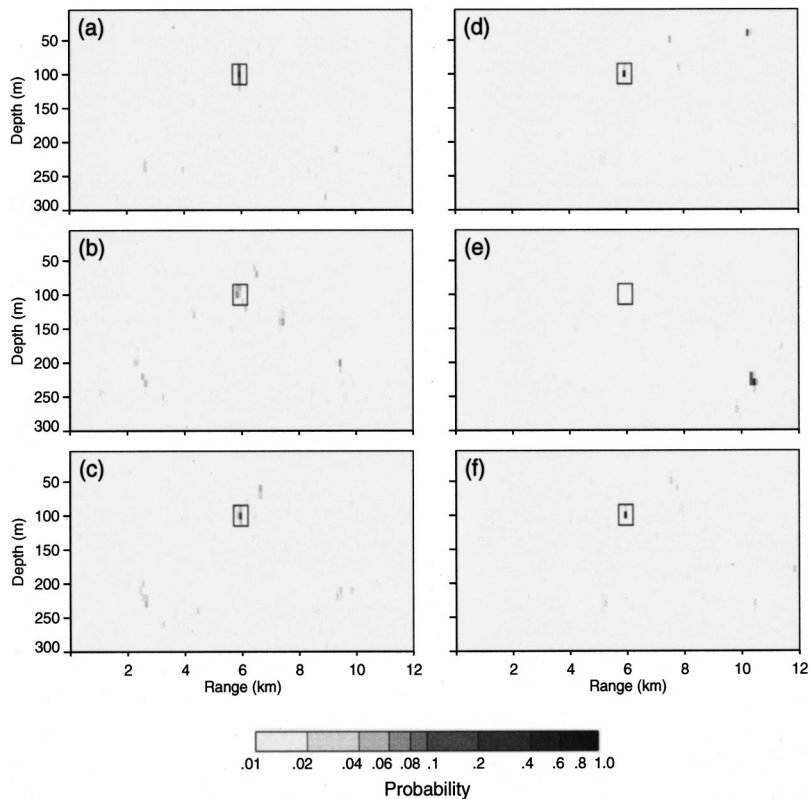


FIG. 8. PAS for a six-sensor VLA at SNR=5 dB with known environmental parameters using (a) RMMP, (b) MMP, and (c) MFP, corresponding results for a 0.6-aperture VLA are shown in (d)–(f). Color levels indicate the estimated probability of correct localization for each grid point, with black boxes delineating the area deemed a correct localization.

addition, the high-order modes which penetrate significantly into the basement are noticeably altered compared to the modes for the known environment.

As in the previous section, localization results for RMMP, MMP, and MFP are considered for three SNRs and for well-sampled, undersampled, and short-aperture cases. Figure 9 shows the probabilities of correct localization for these cases. In this figure, higher-order modes were not omitted from the Bartlett processor to ameliorate environmental mismatch; however, as before, MMP omitted aliased higher-order modes for the undersampled cases. Both MMP and RMMP use an identity regularization matrix. In comparing Fig. 9 with Fig. 7 (no mismatch), it is clear that the probability of correct localization decreases with environmental mismatch, in some cases by a substantial amount. Also, although the results of RMMP were generally superior to the MFP results in Fig. 7, the RMMP and MFP results are much more comparable in Fig. 9.

As previously mentioned, a potential advantage of the modal processors is the ability to concentrate on subsets of the complete mode set. In particular, if environmental mismatch of the geoacoustic parameters is suspected, it may make sense to omit the higher-order modes which interact most strongly with the seabed. Figure 10 shows localization results obtained by including only the eight lower-order modes in the Bartlett processor for RMMP and MMP (the MFP results in Fig. 10 are identical to those in Fig. 9). In Fig. 10, RMMP employed a diagonal regularization matrix with a value of 1.0 on the main diagonal for the first eight modes and 0.5 for the remaining higher-order modes to effectively concentrate the regularization on the modes that are retained. MMP used an identity regularization matrix. A comparison of Fig. 10 with Fig. 9 (all modes included) indi-

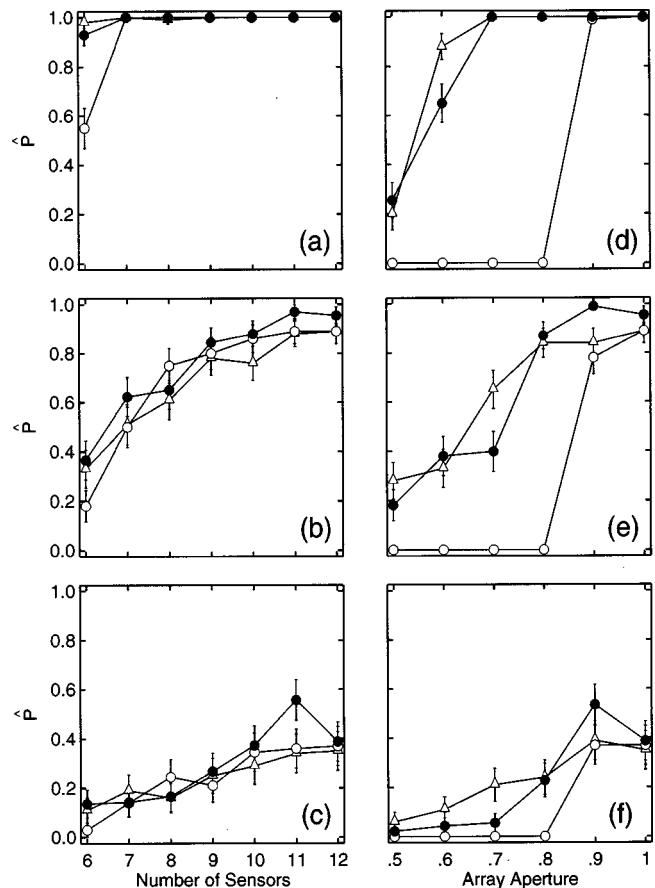


FIG. 9. Estimated probability of correct localization  $\hat{P}$  with environmental mismatch (higher-order modes retained) for MFP (triangles), MMP (open circles), and RMMP (filled circles). (a)–(c) give results for undersampled cases at SNRs of 15, 5, and 0 dB, respectively; (d)–(f) give results for short-aperture cases at the same SNRs. Error bars denote 90% confidence intervals.

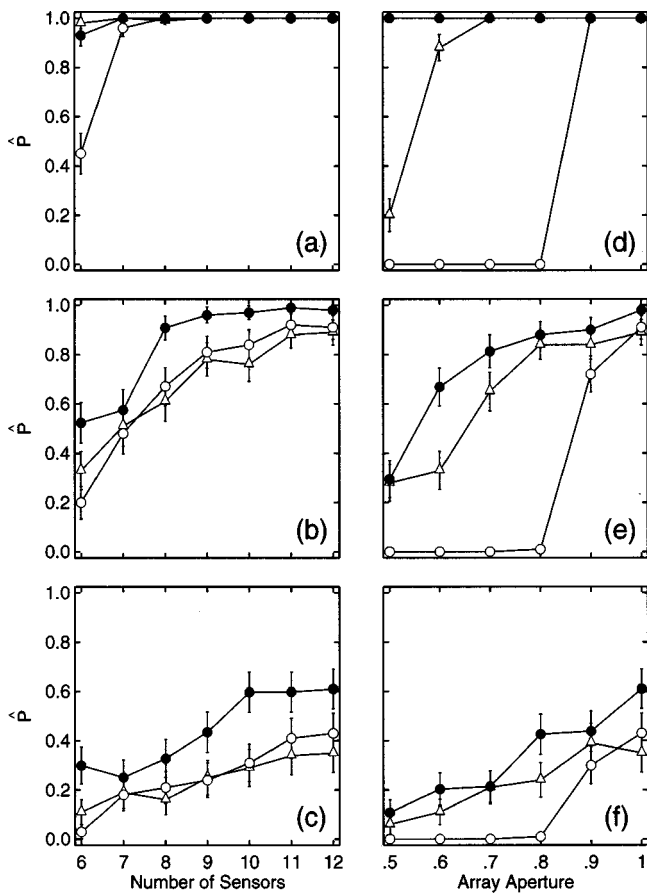


FIG. 10. Estimated probability of correct localization  $\hat{P}$  with environmental mismatch (higher-order modes omitted) for MFP (triangles), MMP (open circles), and RMMP (filled circles). (a)–(c) give results for undersampled cases at SNRs of 15, 5, and 0 dB, respectively; (d)–(f) give results for short-aperture cases at the same SNRs. Error bars denote 90% confidence intervals.

icates a significant improvement in RMMP localization when using only the well-determined modes. The difference is smaller for MMP (for the undersampled cases, this is likely due to the fact that aliased higher-order modes are omitted in both Fig. 9 and Fig. 10). For the undersampled cases at SNR=15 dB [Fig. 10(a)], all methods work well, with the exception of MMP for a six-sensor array. For the short-aperture cases at the same SNR, good results are obtained for MFP for apertures  $>0.6$ , while the RMMP results are uniformly excellent at all apertures. For both undersampled and short-aperture cases at SNR=5 and 0 dB [Fig. 10(b)–(c) (e)–(f)] RMMP consistently achieves a significantly higher probability of correct localization than MFP or MMP. The advantage obtained using RMMP compared to the other localization methods is substantially greater for the mismatched environment (omitting higher-order modes) than for the exact environment (retaining all modes). For instance, the localization results for the known environment in Fig. 7(e) show that for a 0.6-aperture VLA with SNR=5 dB, RMMP and MFP have probabilities of correct localization of  $P=0.65$  and  $P=0.71$ , respectively. However, for the same case in the mismatched environment, the localization results in Fig. 10(e) are  $P=0.65$  and  $P=0.33$  for RMMP and MFP, respectively. Figure 11 shows PASs for the six-sensor and 0.6-aperture cases for RMMP, MMP, and MFP. The major difference between this figure and the PASs obtained for the exact environment (Fig. 8) is the degradation in the MFP results.

### III. SUMMARY AND DISCUSSION

This paper developed and tested a new approach to modal decomposition and matched-mode processing for

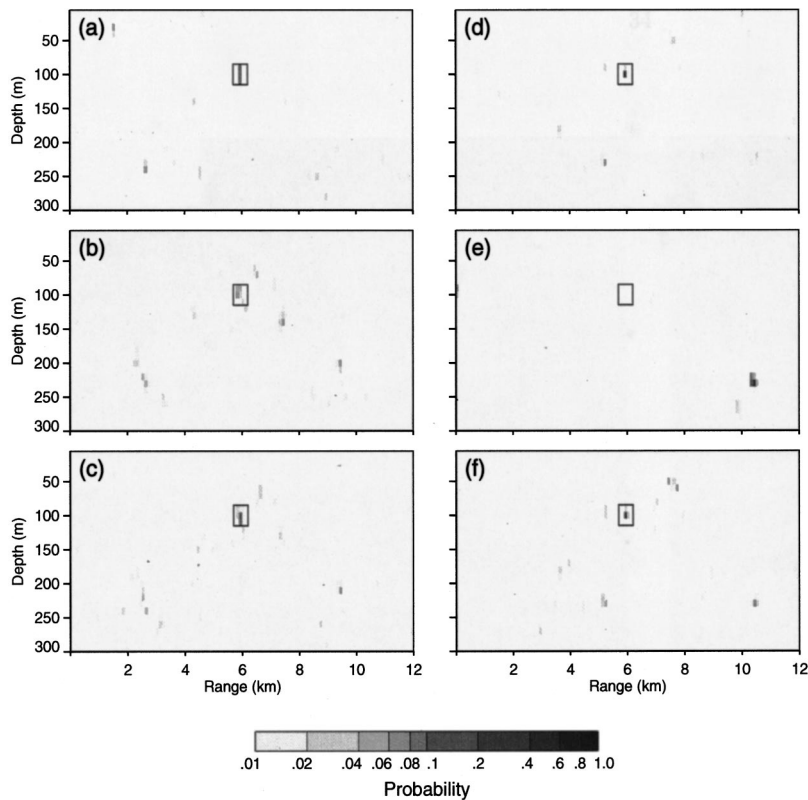


FIG. 11. PAS for a six-sensor VLA at SNR=5 dB with environmental mismatch (higher-order modes omitted) using (a) RMMP, (b) MMP, and (c) MFP, corresponding results for a 0.6-aperture VLA are shown in (d)–(f). Color levels indicate the estimated probability of correct localization for each grid point, with black boxes delineating the area deemed a correct localization.

source localization, with particular application to underdetermined problems. Regularized matched-mode processing carries out an independent inversion of the measured acoustic fields for the constituent modal excitations prior to comparison with the replica excitations at each point of a search grid, using the replica itself as an *a priori* estimate in a regularized inversion. For grid points at or near the source location, this should provide a more physical solution than standard minimum-norm inversion schemes which (explicitly or implicitly) apply an *a priori* estimate of zero. Away from the source position, a stable solution is obtained, even if the *a priori* estimate is not particularly physical. In effect, RMMP produces the optimal match between measured (decomposed) and replica excitations at every point of the search grid. It is observed here that optimizing the match generally favors the source location over other points. As RMMP requires the modal decomposition to be solved for each grid point (as opposed to only once prior to matching for standard MMP), an efficient inversion scheme was developed.

RMMP was compared to MMP and MFP for a series of a shallow-water test cases involving several noise levels and a variety of sensor-array configurations (well-sampled, undersampled, and short-aperture). It was found that RMMP produced generally better localization results in cases where the environmental parameters were exactly known. For cases involving environmental mismatch in the geoacoustic parameters, by omitting the highest-order modes that are most sensitive to bottom properties, RMMP produced substantially better localization results than the other two approaches. The optimal number of modes to omit is difficult to define precisely in general. If the range of possible environmental mismatch can be estimated, examining mode functions and wave numbers computed for end-case environments can provide insight into the number of modes that are essentially insensitive to the mismatch. Fortunately, the localization results do not appear to be overly sensitive to this factor: for the case presented here, similar localization results were obtained when 6–10 low-order modes were retained.

RMMP essentially represents an optimization process providing the best possible match, given the data uncertainties, at each point on the ambiguity surface. This was shown here to generally favor the true source location over other locations (i.e., provides a higher probability of correct localization). However, it should be noted that the resulting higher sidelobe levels could potentially lead to a higher probability of false alarms for low SNR cases.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful discussions with David Thomson, Ross Chapman, Michael Wilmut, and Cedric Zala. This work was supported by a National Science and Engineering Research Council (NSERC) research grant.

## APPENDIX: EFFICIENT COMPUTATION OF TRADEOFF PARAMETER

For a meaningful regularized modal decomposition, the tradeoff parameter  $\theta$  must be determined so that noisy data are fit to a statistically meaningful level, such as achieving

the expected  $\chi^2$  misfit. Determining the appropriate value of  $\theta$  can be the most time-consuming aspect of the inversion, and in RMMP, a new value of  $\theta$  must be determined for every grid point. Hence, it is important to devise as efficient a scheme as possible. The following approach was developed and found to be highly effective. This approach can be applied to both RMMP and standard smallest regularization; in the latter case,  $\hat{\mathbf{x}} = \mathbf{0}$ .

By defining

$$\bar{\mathbf{p}} = \mathbf{G}(\mathbf{p} - \mathbf{A}\hat{\mathbf{x}}), \quad \bar{\mathbf{A}} = \mathbf{G}\mathbf{A}\mathbf{H}^{-1}, \quad \bar{\mathbf{x}} = \mathbf{H}(\bar{\mathbf{x}} - \hat{\mathbf{x}}), \quad (\text{A1})$$

the regularized solutions (18) or (22) can be simplified to

$$\bar{\mathbf{x}} = [\bar{\mathbf{A}}^\dagger \bar{\mathbf{A}} + \theta \mathbf{I}]^{-1} \bar{\mathbf{A}}^\dagger \bar{\mathbf{p}}. \quad (\text{A2})$$

Substituting the SVD of  $\bar{\mathbf{A}}$

$$\bar{\mathbf{A}} = \bar{\mathbf{U}} \bar{\mathbf{\Lambda}} \bar{\mathbf{V}}^\dagger, \quad (\text{A3})$$

into (A2), the regularized solution becomes

$$\bar{\mathbf{x}} = \bar{\mathbf{V}} [\bar{\mathbf{\Lambda}}^2 + \theta \mathbf{I}]^{-1} \bar{\mathbf{\Lambda}} \bar{\mathbf{U}}^\dagger \bar{\mathbf{p}}. \quad (\text{A4})$$

Using the definitions in (A1), the  $\chi^2$  misfit (12) can be written as

$$\chi^2 = (\bar{\mathbf{A}}\bar{\mathbf{x}} - \bar{\mathbf{p}})^\dagger (\bar{\mathbf{A}}\bar{\mathbf{x}} - \bar{\mathbf{p}}). \quad (\text{A5})$$

Substituting (A3) and (A4) into (A5), the functional dependence of  $\chi^2$  on  $\theta$  is given by

$$\chi^2(\theta) = \bar{\mathbf{p}}^\dagger \bar{\mathbf{p}} - \bar{\mathbf{p}}^\dagger \bar{\mathbf{U}} [\bar{\mathbf{\Lambda}}^4 - 2\theta \bar{\mathbf{\Lambda}}^2] [\bar{\mathbf{\Lambda}}^2 + \theta \mathbf{I}]^{-2} \bar{\mathbf{U}}^\dagger \bar{\mathbf{p}}. \quad (\text{A6})$$

An efficient approach to determine the value of  $\theta$  that produces the desired  $\chi^2(\theta)$  to high precision is given by Newton's method. Expanding  $\chi^2(\theta)$  in a Taylor-series expansion to first order about an arbitrary starting value  $\theta_0$  leads to

$$\chi^2(\theta) = \chi^2(\theta_0) + \frac{\partial \chi^2(\theta_0)}{\partial \theta} (\theta - \theta_0). \quad (\text{A7})$$

Equation (A7) can be rearranged to solve for  $\theta$

$$\theta = \theta_0 + \left[ \frac{\partial \chi^2(\theta_0)}{\partial \theta} \right]^{-1} [\chi^2(\theta) - \chi^2(\theta_0)], \quad (\text{A8})$$

where the partial derivative in (A8) is evaluated, using (A6), to be

$$\frac{\partial \chi^2(\theta_0)}{\partial \theta} = 2\theta_0 \sum_{j=1}^M \frac{\bar{\lambda}_j^2}{(\bar{\lambda}_j^2 + \theta_0)^3} (\bar{\mathbf{U}}^\dagger \bar{\mathbf{p}})_j^2. \quad (\text{A9})$$

If  $\chi^2(\theta)$  is sufficiently close to the desired value, the procedure is complete. If not, the starting value is updated,  $\theta_0 \leftarrow \theta$ , and the procedure is repeated iteratively until a satisfactory value is obtained. Finally, from (A1) the solution is

$$\tilde{\mathbf{x}} = \mathbf{H}^{-1} \bar{\mathbf{x}}(\theta) + \hat{\mathbf{x}}. \quad (\text{A10})$$

Note that this approach requires  $\mathbf{H}$  to be a nonsingular matrix (e.g., it can have no zero rows or columns, so all elements of  $\tilde{\mathbf{x}}$  must be regularized).

The above procedure allows  $\theta$  to be determined using only a single application of SVD, i.e., repetitive computation of the matrix inversion in (18) or (22) for different values of  $\theta$  is avoided (note that the matrix inversion in (A6) involves

the trival case of a diagonal matrix). Although Newton's method converges quadratically near a solution, it requires an appropriate starting value  $\theta_0$ . Such a value can be determined using a bisection method. For RMMP,  $\theta_0$  is determined in this manner only for the first grid point. For all subsequent grid points, the  $\theta$  value corresponding to final solution for the previous grid point is used as the new starting value. For all inversions in this paper, the approach outline here required only one to three iterations of Newton's method to obtain the desired misfit to a precision of less than 1%.

- <sup>1</sup>H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," *J. Acoust. Soc. Am.* **59**, 368–373 (1976).
- <sup>2</sup>J. M. Ozard, "Matched field processing in shallow water for range, depth, and bearing estimates: Results of experiment and simulation," *J. Acoust. Soc. Am.* **86**, 744–753 (1989).
- <sup>3</sup>A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE J. Ocean Eng.* **18**, 401–424 (1993).
- <sup>4</sup>A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).
- <sup>5</sup>R. H. Ferris, "Comparison of measured and calculated normal-mode amplitude functions for acoustic waves in shallow water," *J. Acoust. Soc. Am.* **52**, 981–988 (1972).
- <sup>6</sup>C. T. Tindle "Measurements of the frequency dependence of normal modes," *J. Acoust. Soc. Am.* **64**, 1178–1185 (1978).
- <sup>7</sup>E. C. Shang, "Source depth estimation in waveguides," *J. Acoust. Soc. Am.* **77**, 1413–1418 (1985).
- <sup>8</sup>T. C. Yang, "A method of range and depth estimation by modal decomposition," *J. Acoust. Soc. Am.* **82**, 1736–1745 (1987).
- <sup>9</sup>G. R. Wilson, R. A. Koch, and P. J. Vidmar, "Matched mode localization," *J. Acoust. Soc. Am.* **84**, 310–320 (1988).
- <sup>10</sup>M. J. Hinich and E. J. Sullivan, "Maximum-likelihood passive localization using mode filtering," *J. Acoust. Soc. Am.* **85**, 214–219 (1989).
- <sup>11</sup>T. C. Yang, "Effectiveness of mode filtering: A comparison of matched-field and matched-mode processing," *J. Acoust. Soc. Am.* **87**, 2072–2084 (1990).
- <sup>12</sup>A. G. Voronovich *et al.*, "Comparative analysis of methods for the normal mode decomposition of a sound field in a waveguide: Numerical simulation and full-scale experiment," *Sov. Phys. Acoust.* **38**, 365–370 (1992).
- <sup>13</sup>A. Timonov, "Development of a spectral regularization modal decomposition algorithm," DREP Contractors Report 9601, Defence Research Establishment Pacific, Victoria, BC (1996).
- <sup>14</sup>J. R. Buck, J. C. Preisig, and K. E. Wage, "A unified framework for mode filtering and the maximum *a posteriori* mode filter," *J. Acoust. Soc. Am.* **103**, 1813–1824 (1998).
- <sup>15</sup>K. D. Heaney and W. A. Kuperman, "Very long-range source localization with a small vertical array," *J. Acoust. Soc. Am.* **104**, 2149–22159 (1998).
- <sup>16</sup>W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory* (Academic, New York, 1984).
- <sup>17</sup>C. van Schooneveld, "Inverse problems: A tutorial survey," in *Underwater Acoustic Data Processing*, edited by Y. T. Chan (Kluwer, The Netherlands, 1989), pp. 393–411.
- <sup>18</sup>W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Numerical Computing*, 2nd ed. (Cambridge University Press, New York, 1992).
- <sup>19</sup>S. E. Dosso, M. R. Fallat, B. J. Sotirin, and J. L. Newton, "Array element localization for horizontal arrays via Occam's inversion," *J. Acoust. Soc. Am.* **104**, 846–859 (1998).
- <sup>20</sup>G. V. Frisk, *Ocean and Seabed Acoustics* (Prentice-Hall, Englewood Cliffs, NJ, 1994).
- <sup>21</sup>D. W. Oldenburg, "One-dimensional inversion of natural source magnetotelluric observations," *Geophysics* **44**, 1218–1244 (1979).
- <sup>22</sup>D. W. Oldenburg, "An introduction to linear inverse theory," *IEEE Trans. Geosci. Remote Sens.* **GE-22**, 665–674 (1984).
- <sup>23</sup>N. E. Collison, "Regularized matched-mode processing for ocean acoustic source localization," M.Sc. thesis, University of Victoria, Victoria, Canada, 1999.
- <sup>24</sup>N. E. Collison and S. E. Dosso, "A comparison of modal decomposition algorithms for matched-mode processing," *Can. Acoust.* (submitted).
- <sup>25</sup>M. L. Taillefer, "Range dependent matched field source localization and tracking in shallow water in a continental slope region of the Northeast Pacific Ocean," M.Sc. thesis, University of Victoria, Victoria, Canada, 1998.
- <sup>26</sup>H. Schmidt, "SAFARI: Seismo-acoustic fast field algorithm for range-independent environments," SACLANT ASW Research Centre Rep. SR-113, La Spezia, Italy (1988).
- <sup>27</sup>E. K. Westwood, C. T. Tindle, and N. R. Chapman, "A normal mode model for acoustoelastic ocean environments," *J. Acoust. Soc. Am.* **100**, 3631–3645 (1996).
- <sup>28</sup>J. Tabrikian, J. L. Krolik, and H. Messer, "Robust maximum-likelihood source localization in an uncertain shallow-water waveguide," *J. Acoust. Soc. Am.* **101**, 241–249 (1997).
- <sup>29</sup>Z. Michalopoulou, "Robust multi-tonal matched-field inversion: A coherent approach," *J. Acoust. Soc. Am.* **104**, 163–170 (1998).
- <sup>30</sup>J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, 3rd ed. (Brooks Cole, Pacific Grove, CA, 1991).

# Computed narrow-band time-reversing array retrofocusing in a dynamic shallow ocean

Michael R. Dungan and David R. Dowling<sup>a)</sup>

*Department of Mechanical Engineering and Applied Mechanics, University of Michigan, Ann Arbor, Michigan 48109-2121*

(Received 16 April 1999; accepted for publication 20 March 2000)

A time-reversing array (TRA) can retrofocus acoustic energy, in both time and space, to the original sound-source location without any environmental information. This unique capability may be degraded in time-dependent or noisy acoustic environments, or when propagation losses are prevalent. In this paper, monochromatic propagation simulations (based on the parabolic equation code, RAM) are used to predict TRA retrofocusing performance in shallow-water sound channels having characteristics similar to those measured during the recent SWARM (shallow-water acoustics in a random medium) experiment. Results for the influence of source–array range, source depth, acoustic frequency, bottom absorption, internal wave strength, and round-trip time delay are presented. For a fixed channel geometry, higher frequencies, deeper sources, and lower bottom absorption improve TRA performance and allow retrofocusing at longer ranges. In a dynamic shallow-water channel containing a random superposition of linear internal waves, the size of the retrofocus is slightly decreased and sidelobes are suppressed compared to the static channel results. These improvements last for approximately 1 to 2 min for source-array ranges near 10 km at a frequency of 500 Hz. For longer time delays, the internal waves cause significant TRA retrofocus amplitude decay, and the decay rate increases with increasing internal wave activity and acoustic frequency. © 2000 Acoustical Society of America. [S0001-4966(00)06106-3]

PACS numbers: 43.30.Vh, 43.30.Yj, 43.30.Ft [DLB]

## INTRODUCTION

Acoustic time reversal is the process of recording an acoustic signal and then replaying it backwards so that the last part of the recording is played first and the first part of the recording is played last. When this process is accomplished by an array of transducers (a time-reversing array, or TRA), the replayed sounds will focus at the place of their origin, provided that the acoustic environment is reciprocal and absorption losses are small. Because this retrofocusing occurs without environmental information, TRAs have some unique properties that are of interest in ultrasonics (Fink, 1997; Tanter, 1998) and underwater sound (Jackson and Dowling, 1991; Dowling and Jackson, 1992; Dowling, 1993, 1994; Parvulescu, 1995; Kuperman *et al.*, 1998; Song *et al.*, 1998; Khosla and Dowling, 1998). In fact, acoustic time reversal is now well on its way to becoming an established technique for focusing acoustic waves in unknown environments, and for acoustically characterizing unknown environments by quantitatively interrogating the retrofocus size and field amplitude. In particular, active TRA systems may facilitate moderate range (tens of kilometers) underwater communication through unknown acoustic environments. Novel uses of time-reversal concepts are also occurring in other areas of acoustics (Rose *et al.*, 1999; Roux *et al.*, 1999; Yönak and Dowling, 1999).

The dimensions, field amplitude, and temporal decay of the retrofocus produced by a narrow-band TRA are deter-

mined by the signal frequency, the source–array geometry, and the acoustic environment. In general, higher frequencies lead to better TRA retrofocusing (Jackson and Dowling, 1991). Similarly, larger array apertures and shorter source-array ranges may also improve TRA retrofocusing (Dowling and Jackson, 1992; Hodgkiss *et al.*, 1997). The acoustic environment may aid or limit TRA retrofocusing. For example, inhomogeneous media may produce TRA superfocusing (Saiichev, 1982; Dowling and Jackson, 1992; Derode *et al.*, 1995; Khosla and Dowling, 1998), while environmental dynamics typically cause retrofocus-amplitude decay (Dowling and Jackson, 1992; Dowling, 1994; Khosla and Dowling, 1998).

In a dynamic environment, the refractive-index structure associated with the source-to-array propagation step may change before the array-to-source propagation step because of the requisite time delay for propagation, recording, and signal processing. Therefore, when the time-reversed sound is reintroduced into the medium, it travels back through an altered version of the original environment, and retrofocusing is degraded because of this mismatch. Longer time delays produce larger mismatches that lead to greater TRA retrofocus degradation. The studies mentioned above show that this degradation may consist of lower focal amplitude, larger focal size, or stronger and more prevalent sidelobes (spurious retrofoci).

The environment may also limit TRA retrofocusing performance through absorption losses (Tanter *et al.*, 1998) and the restrictions associated with modal propagation (Kuperman, 1997). Hence, one of the most challenging applications of TRAs lies in shallow ocean sound channels where internal

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: drd@engin.umich.edu



waves produce dynamic inhomogeneities, modal propagation is assured, and bottom absorption, bubbles, and rough-surface scattering provide important loss mechanisms. For the 100-Hz to 1-kHz simulation results presented here, only bottom absorption losses are included because they are likely to be the dominant loss mechanism at the source–array ranges of interest (1 to 50 km). The effects of bubbles are usually more prevalent at higher frequencies. Assessments of rough surface scattering and TRA performance are provided in Kuperman *et al.* (1998) and Khosla and Dowling (1998).

This paper presents the results of monochromatic numerical simulations of TRA retrofocusing performance in static and dynamic shallow ocean sound channels. The main independent parameter is the source–array range,  $R$ . The influence of variations in source depth, acoustic frequency, bottom absorption, internal wave strength, and round-trip time delay on TRA retrofocusing performance is presented. Properties of the sound channel are drawn from measurements made during the SWARM experiment (Apel *et al.*, 1997; Finette, 1998) and reference sources (Jensen *et al.*, 1994). For simplicity and clarity, only results from range-independent environments are presented, and only one array geometry has been considered: a linear vertical array having 24 elements that spans the nominal water column depth, and only sends and receives narrow-band signals. The effects of subsampling a water-column-spanning array are presented in Kuperman *et al.* (1998).

The remainder of this paper is organized into four sections. The next section describes the acoustic environment used in this study, the internal wave model, and the computational technique. Sections II and III present the simulation results for static and dynamic sound channels, respectively. The final section summarizes the findings and states the conclusions drawn from this study.

## I. COMPUTATIONAL APPROACH

This study utilizes parabolic-equation (PE) computations to predict how TRA retrofocusing is influenced by the environmental parameters, source–array geometry, and acoustic frequency. The sound propagation computations were performed using the wide-angle PE-code RAM (Collins, 1993, 1994, 1998). The following three subsections describe the environments investigated, the internal wave model, and the implementation of RAM.

### A. Environment

The basic or generic environment for this study is a range-independent sound channel (see Fig. 1) having a depth  $D=75$  m and a two-layer bottom. The  $r$  coordinate runs in the horizontal direction and the  $z$  coordinate has its origin at the water surface and increases downward. The sound-speed profile in the channel, shown in Fig. 2, was obtained by smoothing and combining oceanic measurements of water temperature and salinity as a function of depth. These measurements were provided by Finette (1998), and were made at the SWARM shallow site where the actual ocean depth is 75 m (Apel *et al.*, 1997). The unusual strength of the downward refraction in the profile of Fig. 2 is common in mid to late summer at the SWARM site. In addition, such a strongly

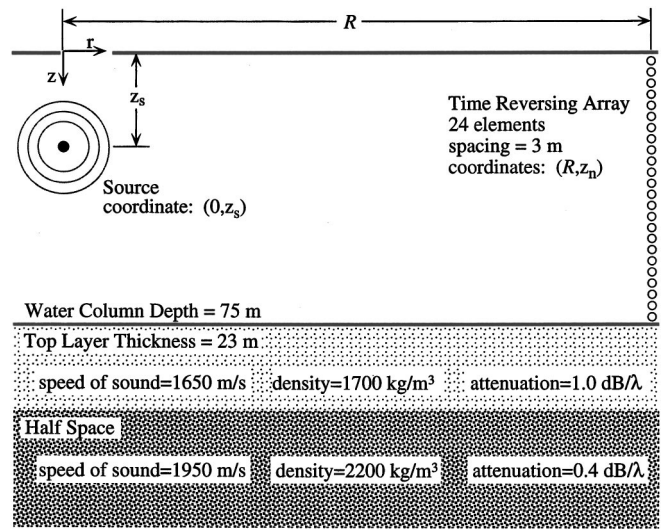


FIG. 1. Schematic of the sound channel and time-reversing array (TRA) geometry. The horizontal coordinate is  $r$ . The vertical coordinate,  $z$ , increases downward. The omnidirectional source at  $(0, z_s)$  broadcasts at time  $t_1$ . The TRA receives the signal at  $(R, z_n)$  and retransmits to form the retrofocus at time  $t_2$ . In a dynamic acoustic environment, TRA retrofocus properties may be degraded when  $t_2 - t_1 > 0$ .

downward-refracting profile indicates favorable conditions for an active internal wave field. Such conditions were sought for the SWARM experiment (Apel *et al.*, 1997). For TRA operations, the main effect of the downward-refracting profile is to enhance the importance of array elements in the lower half of the water column while leaving the upper array elements somewhat underutilized.

The density, speed of sound, and thickness of the two bottom layers are based on core and chirp measurements made at the SWARM site (Apel *et al.*, 1997). The top layer has a thickness of 23 m, sound speed of 1650 m/s, and density of  $1700 \text{ kg/m}^3$ . The second layer is much thicker and has a sound speed of 1950 m/s and a density of  $2200 \text{ kg/m}^3$ . Attenuation values of 1.0 decibel per wavelength in the upper layer, and 0.40 decibel per wavelength in the lower layer were extrapolated from tabulations in Jensen *et al.* (1994) as conservative matches to the available geological measurements. The top of the sound channel was treated as a flat pressure release surface.

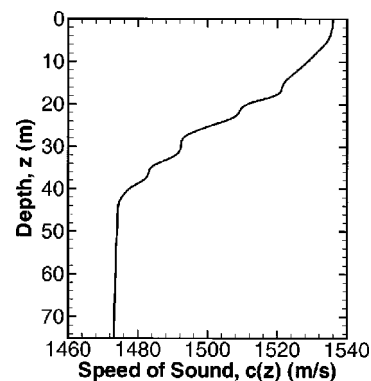


FIG. 2. Speed of sound versus depth for the standard sound channel. This profile is based on measurements made during the 1995 SWARM experiment and was provided to the authors by Dr. Steven Finette of the Naval Research Laboratory.

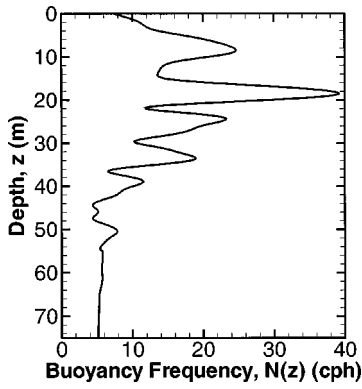


FIG. 3. Buoyancy frequency versus depth for the standard sound channel. This profile is based on measurements made during the 1995 SWARM experiment and was provided to the authors by Dr. Steven Finette of the Naval Research Laboratory.

The environmental parameters given above will be the same throughout the rest of this paper unless stated otherwise. For example, results from differing levels of absorption in this environment are presented in Sec. II.

### B. Linear internal waves

The dynamic effects of internal waves on TRA retrofocusing were predicted by inserting a random superposition of linear internal waves into the sound channel described above (similar to the formulation in Dashen *et al.*, 1979 and Tielbörger *et al.*, 1997). The interaction of tides and the local continental shelf can lead to nonlinear soliton-type internal waves, while random shallow-water linear internal waves are merely the near-coast extension of deep-water linear internal waves (Garrett and Munk, 1979, provides a review of deep-water internal wave properties). Here, the internal wave model is intended to represent the shallow SWARM site in the absence of soliton activity. Figure 3 shows the Brunt-Väisälä or buoyancy frequency profile,  $N(z)$ , used to determine and animate the internal waves. This profile was obtained by smoothing water-column density measurements from the SWARM shallow site (Finette, 1998).

Internal waves cause heaving motions within the water column that alter the local sound speed by lifting and lowering the otherwise static contours of constant sound speed. The vertical deflections produced by these internal waves are governed by a second-order differential equation (Dashen *et al.*, 1979)

$$\frac{d^2}{dz^2} W(z) + k_h^2 \left[ \frac{N^2(z) - \omega^2}{\omega^2 - f_c^2} \right] W(z) = 0, \quad (1)$$

which, along with the waveform boundary conditions of  $W(0) = W(D) = 0$ , creates an eigenvalue problem. Here,  $W(z)$  is the vertical displacement produced by an internal wave having radian frequency  $\omega$ ,  $k_h$  is the horizontal wave number of the internal wave, and  $f_c = 2\Omega \sin(\text{°latitude})$  is the inertial frequency with  $\Omega$  equal to the rotation rate of the earth. In this study,  $f_c$  was evaluated at a latitude of  $39.25^\circ$ , corresponding to the center of the SWARM site. When  $\omega$  and  $k_h$  are independent parameters, (1) has eigenmode solutions  $W(k_h, j, z)$  at a discrete set of frequency eigenvalues

$\omega_j(k_h)$ . The internal wave modes and frequencies particular to this study were calculated from (1) using standard shooting techniques (see Press *et al.*, 1986). Hence, a dynamic three-dimensional simulation of the internal-wave-induced vertical displacement  $\eta$  can be obtained by a triple sum over internal-wave modes ( $j$ ), horizontal wave numbers ( $k_h$ ), and azimuthal angles ( $\theta$ ),

$$\eta(r, \phi, z, t) = \text{Re} \left\{ \sum_j \sum_{k_h} \sum_{\theta} F(k_h, j) W(k_h, j, z) \times \exp[ik_h r \cos \theta \cos \phi + ik_h r \sin \theta \sin \phi - i\omega(k_h, j)t] \right\}, \quad (2a)$$

where  $\phi$  is the azimuthal angle associated with the  $r$ - $z$  coordinate system. The  $F$  term is a Gaussian random complex weighting having a uniform distribution in  $\theta$  and a Garrett-Munk spectrum in mode number and horizontal wave number

$$\langle \|F(k_h, j)\|^2 \rangle = \frac{4}{\pi} E_0 M [j^2 + j_*^2]^{-p/2} k_j k_h^2 (k_h^2 + k_j^2)^{-2}, \quad (2b)$$

where

$$k_j = \frac{\pi j f_c}{\int_0^D N(z) dz}, \quad (2c)$$

$$M = \left\{ \sum_{j=1}^{\infty} (j^2 + j_*^2)^{-p/2} \right\}^{-1}, \quad (2d)$$

$E_0$  is the internal-wave energy level,  $j_*$  is the mode weighting parameter, and  $p$  is the mode weighting exponent. The sum given in (2a) was calculated using a variable azimuthal resolution so that random numbers were assigned to approximately equal tile areas of the  $k_h$ - $\theta$  plane for each value of  $j$ . Hence, the number of  $\theta$  evaluations in (2a) increased linearly with increasing  $k_h$ . The parametric ranges and increments used were:  $1 \leq j \leq 15$ ,  $5 \times 10^{-4} \text{ m}^{-1} \leq k_h \leq 1.5 \times 10^{-1} \text{ m}^{-1}$ , with steps of  $\Delta k_h = 5 \times 10^{-4} \text{ m}^{-1}$ , and  $0 \leq \theta \leq 360^\circ$  with  $\Delta \theta = 2^\circ$  when  $k_h = 1.5 \times 10^{-1} \text{ m}^{-1}$ . The internal wave mode shapes  $W(k_h, j, z)$  evolve gradually with increasing  $k_h$ , so mode shapes were interpolated between the minimum and maximum values of  $k_h$  with intervening calculations every  $10^{-2} \text{ m}^{-1}$ . It was found that additional refinements influenced the statistics of the simulated internal wave fields less than a change in the random number seed. In addition, the plane  $\phi = 0$  was chosen throughout this study, so the second term in the exponential function in (2a) was consistently zero.

The Garrett-Munk spectrum parameters ( $E_0, p, j_*$ ) were chosen to match a temporal power spectrum of vertical displacements measured at the SWARM shallow site (75-m depth) without soliton activity (Finette, 1998). This was accomplished by using (2a) at a fixed depth of 25 m to generate a synthetic vertical displacement time series, computing the power spectrum from this synthetic time series, and then comparing the resulting synthetic spectrum with the SWARM measurements. Adjustments were made to the

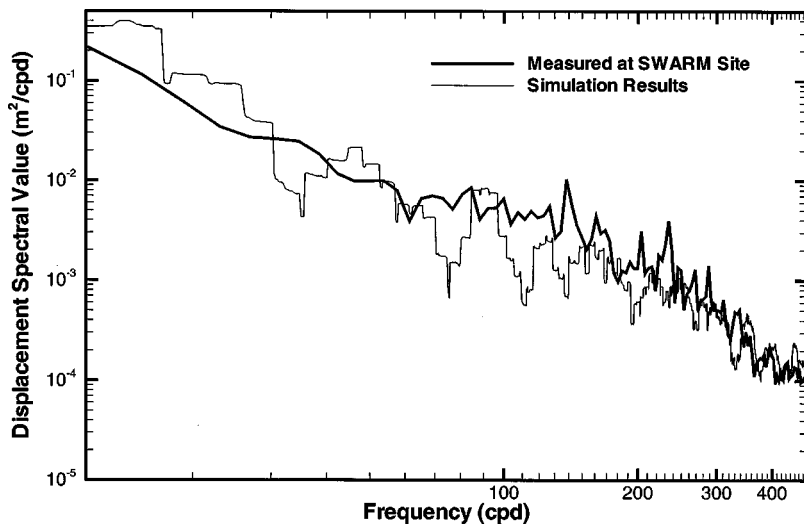


FIG. 4. Comparison of the synthetic internal-wave vertical-displacement temporal power spectrum and a spectrum measured during the 1995 SWARM experiment in the absence of soliton internal waves. The measured spectrum was provided to the authors by Dr. Steven Finette of the Naval Research Laboratory.

spectrum parameters until a visually acceptable match of the two spectra was achieved. Figure 4 shows a sample result for the final parameter choices:  $j_* = 1$ ,  $p = 4$ , and  $E_o = 24.6 \text{ J/m}^2$ . Other independent research suggests similar values for these parameters (Tielbürger, 1997; Yoo and Yang, 1998; Yang and Yoo, 1998). Spectral matches obtained with shallower (15 m) and deeper (40 m) synthetic data were not quite as good as that shown in Fig. 4, but deviations from the measured spectrum at these locations were comparable to those produced by changes in the random number seed.

Sample internal-wave vertical-displacement fields,  $\eta(r, 0, z, t)$ , are shown in Fig. 5 for  $t = 0$  (a), and  $t = 20$  min (b) over a 50-km range. Comparison of Fig. 5(a) and (b) shows that the short wavelength features of the internal wave field decorrelate in 20 min, while the long wavelength features persist, as expected. The disruption by internal waves of the initial layered sound-speed profile (Fig. 2) is accomplished by merely deflecting the static sound-speed profile by

$\eta(r, 0, z, t)$  at each depth and range location in the input file to RAM. Corrections to the internal-wave-induced sound-speed changes due to the adiabatic gradients in the water column variables were on the order of 1% and were therefore neglected.

### C. RAM implementation

The main value of a computational study like this one is the ease with which parametric studies may be completed. Here, the main computational tool is the one-way wide-angle PE code RAM (Collins, 1993, 1994, 1998) modified to output the real and imaginary field components at each grid point. RAM is based on a split-step Padé solution method. In the forward propagation step (at time  $t_1$ ), a point source with radian frequency  $\omega$  located at  $(0, z_s)$  ensonifies the sound channel described above. Elements of the TRA located at  $(R, z_n)$ , record this field, complex conjugate it [complex conjugation is the equivalent of time reversal for narrow-band

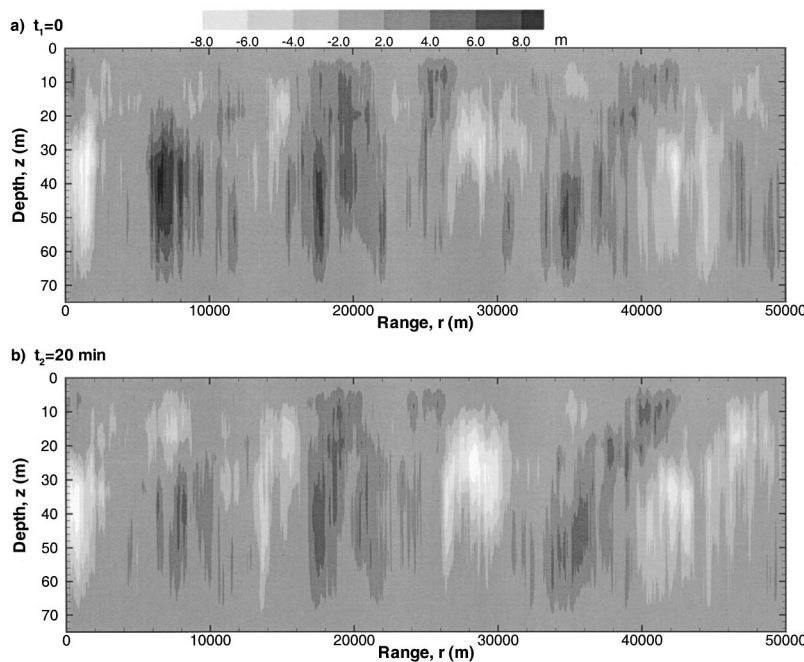


FIG. 5. Sample synthetic internal-wave vertical displacement fields over a 50-km range at an initial time (a), and 20 min later (b). In 20 min most of the small-scale fluctuations have changed, but the larger scale deflections persist. The gray scale ranges from white (less than  $-8\text{-m}$  downward deflection) to black (more than  $+8\text{-m}$  upward deflection).

signals (Jackson and Dowling, 1991)], amplify it, and re-broadcast it (at time  $t_2$ ). Thus, the response of the array is entirely described by the complex-conjugate Green's function,  $G_{c\omega}$  [see Dowling and Jackson (1992), Eq. (10), or Kuperman *et al.* (1998), Eq. (7)]

$$G_{c\omega}(r, z, t_2 - t_1) = \sum_{n=1}^N G_{2\omega}(r, z; R, z_n) G_{1\omega}^*(R, z_n; 0, z_s) \quad (3)$$

where  $r$  and  $z$  are the range and the depth coordinates, respectively,  $N$  is the number of array elements,  $G_{m\omega}(r, z; r', z')$  is the Green's function at radian frequency  $\omega$  for acoustic propagation from  $(r', z')$  to  $(r, z)$  at time  $t_m$ , and the asterisk denotes complex conjugation. The time delay,  $t_2 - t_1$ , becomes important in dynamic environments. For all retrofocusing computations presented here, the array had 24 elements with 3-m vertical spacing. Here, the use of RAM to compute the fields  $G_{1\omega}$  and  $G_{2\omega}$  limits the acoustic propagation to the  $r$ - $z$  plane. Out-of-plane scattering from the fully three-dimensional synthetic internal wave fields is neglected. Neglect of azimuthal scattering is acceptable as long as the average water column properties are range independent, and the internal-wave-induced horizontal gradients in the speed of sound are much smaller than the vertical sound-speed gradients. Both conditions are satisfied here.

To characterize the strength of the retrofocused field at different ranges and depths, the TRA *transmission loss* is defined

$$TL_{c\omega} = -20 \log \left( \frac{G_{c\omega}(r, z, t_2 - t_1)}{G_{c\omega}(1 \text{ m}, z_s, 0)} \right). \quad (4)$$

Here, the reference amplitude is that obtained at the retrofocus of the same array operated in free space when the source–array range is 1 m.

Studies of one-way propagation and computational reciprocity, as well as comparisons to modal sum calculations, were conducted to determine the input parameters best suited for the retrofocus problem at hand. Such parameters include the number of Padé terms, depth and range grid spacing, and computational bottom depth and attenuation. Convergence studies confirmed the choices of eight Padé terms, a depth grid spacing of 1/60th of a wavelength or less, and a range grid spacing of no more than four-thirds of a wavelength. A short-range step is required to obtain the details in the retrofocus peak region. The computational bottom depth was taken to be at least three times the water column depth. Within a few wavelengths of the computational bottom, the attenuation ramps up to a final value of 10.0 dB per wavelength to prevent artificial reflections.

Figure 6 shows a sample TRA retrofocus calculation in the standard sound channel without internal waves for a source depth of  $z_s = 25$  m at a frequency of 500 Hz and a source–array range of  $R = 10$  km. The horizontal (range) axis is zoomed in to the region near the source. The array lies off the figure to the right. As expected, the retrofocus peak appears in the correct location and is clearly the highest amplitude feature on the vertical path through  $r = 0$ . In fact, close inspection of Fig. 6 shows that all sound amplitudes above and below the retrofocus are well below the up- and down-

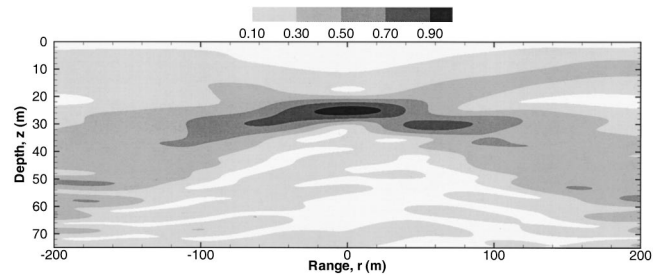


FIG. 6. Sample retrofocus field without internal waves for a source depth of  $z_s = 25$  m, an acoustic frequency of  $f = 500$  Hz, and a source–array range of  $R = 10$  km. The peak amplitude occurs at the correct location and the nearest sidelobe is roughly 60 m up-range from the original source location.

range sidelobes. This is a clear indication that nearly all of the propagating sound energy passes near the source location, and that the retrofocus is not spurious.

## II. STATIC SOUND CHANNEL RESULTS

The results of this investigation into the retrofocusing capability of a time-reversing array are presented in two parts. This section contains results for static sound channels lacking internal wave motions. The parameters explored here include source depth, source–array range, acoustic frequency, and bottom attenuation. The next section covers a dynamic sound channel containing internal waves.

Figure 7(a) shows how the retrofocus field at  $r = 0$  depends on the source depth  $z_s$  in the standard sound channel at a frequency of 500 Hz and  $R = 10$  km. The darkness of the gray scale is proportional to  $G_{c\omega}(0, z, 0) / G_{c\omega}(0, z_s = 63 \text{ m}, 0)$  where the loudest retrofocus was obtained at a source depth  $z_s = 63$  m. This figure shows that retrofocusing is loudest in the lower third of the sound channel, as would be expected with the pressure release boundary condition at  $z = 0$  and the strongly downward-refracting sound-speed profile (Fig. 2). Figure 7(b) shows the same retrofocus information normalized by the peak amplitude on each vertical slice. In an ideal environment, the central dark diagonal curve in Fig. 7(b) would be straight and most of the figure would be white. Here, the vertical sidelobe structure is weakest above the source and when the source resides in the lower half of the sound channel below the thermocline.

Figure 8 shows how the retrofocus field at  $r = 0$  changes with increasing source–array range  $R$  at four acoustic frequencies,  $f = 100$  Hz, 250 Hz, 500 Hz, and 1 kHz, for a source depth of  $z_s = 25$  m. For both this and the next figure, the gray-scale darkness is proportional to  $G_{c\omega}(0, z, 0)$  divided by its maximum value. This scaling eliminates the effect of focal amplitude decay caused by increasing  $R$ .

Several interesting trends are shown in Fig. 8. First, higher frequencies produce smaller retrofoci whose size is essentially proportional to the acoustic wavelength. Here, the vertical retrofocus size is approximately three wavelengths (or less) out to ranges of 10 km at all four frequencies. A second trend shown in Fig. 8 is that higher frequencies lead to effective retrofocusing at longer ranges. Better retrofocusing at longer ranges requires more weakly absorbed modes, and higher acoustic frequencies allow a greater number of (all types of) modes for fixed sound channel properties. A

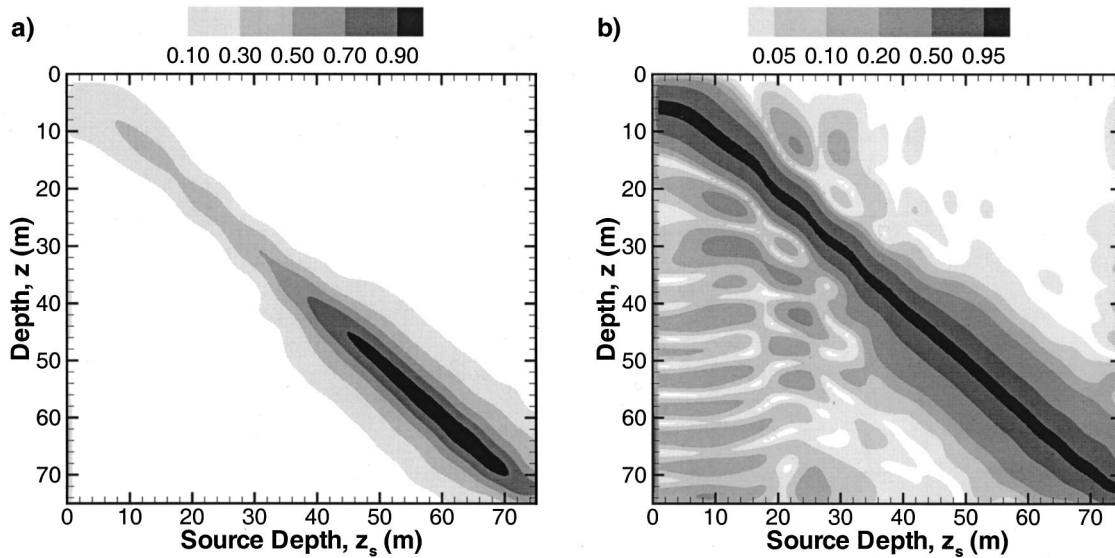


FIG. 7. Retrofocus field amplitude through the channel depth on a vertical line at  $r=0$  versus source depth without internal waves at an acoustic frequency of  $f=500$  Hz and a source–array range of  $R=10$  km. In (a), the gray scale is normalized to the overall peak amplitude, which is found at a source depth of  $z_s=63$  m. Because of the modal structure of the sound channel, source locations in the lower third of the channel lead to stronger (louder) retrofoci. In (b), the gray scale is normalized by the peak amplitude on each vertical slice so that the sidelobe structure of each retrofocus is more clearly displayed. Minimal sidelobes occur when the retrofocus is in the lower half of the sound channel.

third trend seen in Fig. 8 is the downward drift and vertical growth of the focal spot at all frequencies with increasing range. This third trend has also been found in TRA experiments performed in the Mediterranean Sea (Hodgkiss *et al.*, 1997). The retrofocus drift is most apparent at  $f=100$  Hz [Fig. 8(a)], where there is essentially only one propagating mode for  $R>20$  km.

In Fig. 9, results similar to those shown in Fig. 8 are displayed for a deeper source placed at  $z_s=50$  m. Here, the

same three trends are found except that vertical drift of the retrofocus location is less pronounced because this lower source lies near the antinode of the most persistent propagating mode.

To complement the results in Figs. 8 and 9, Table I provides the computed-field transmission loss,  $TL_{c\omega}$ , for the retrofocus-field peak amplitude through the depth at source–array ranges of  $1\text{ km}\leq R\leq 50\text{ km}$  for four acoustic frequencies,  $f=100$  Hz, 250 Hz, 500 Hz, and 1 kHz, and source

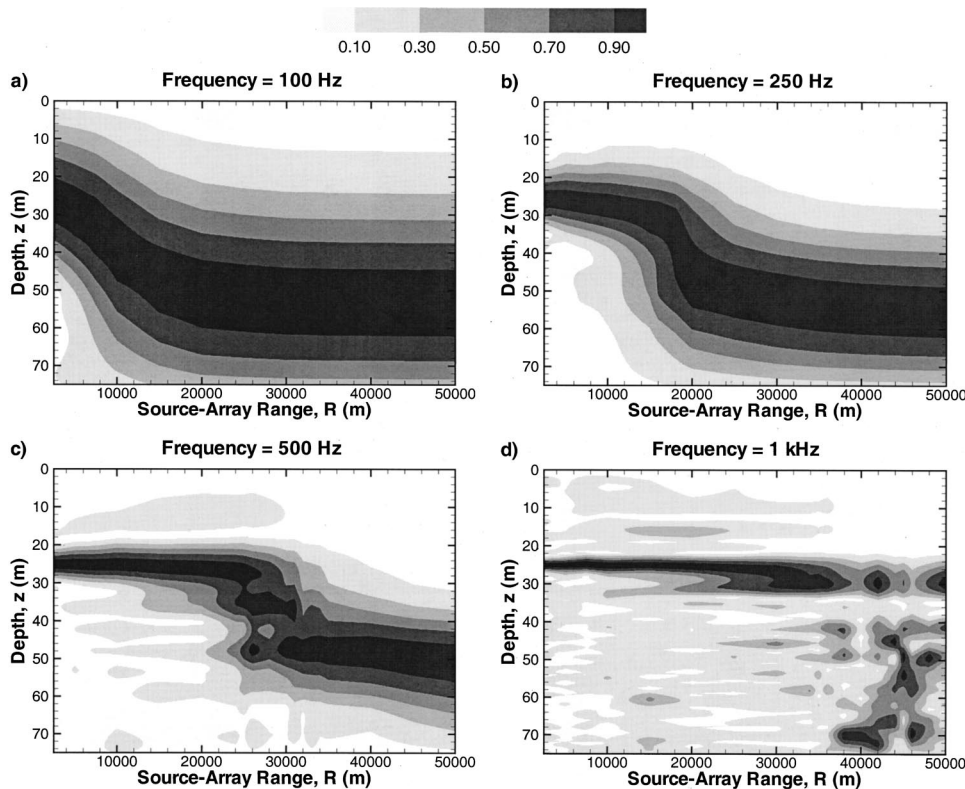


FIG. 8. Retrofocus field amplitude through the channel depth on a vertical line at  $r=0$  versus source–array range  $R$  without internal waves for a source depth of  $z_s=25$  m at acoustic frequencies of  $f=100$  Hz (a), 250 Hz (b), 500 Hz (c), and 1 kHz (d). Here, each vertical slice is normalized by its maximum value so that the vertical growth and migration of the retrofocus with increasing  $R$  is emphasized. Dark regions lie near the local field maximum at any  $R$ . As expected, higher frequencies produce tighter retrofoci.

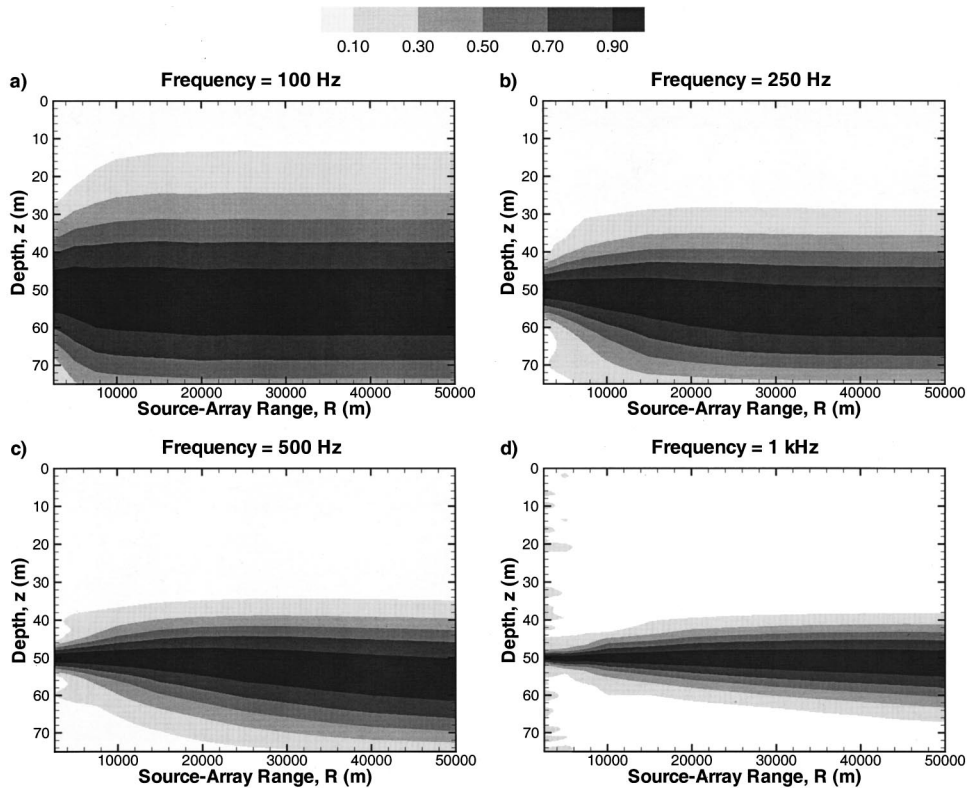


FIG. 9. Retrofocus field amplitude through the channel depth on a vertical line at  $r=0$  versus source-array range  $R$  without internal waves for a source depth of  $z_s=50$  m at acoustic frequencies of  $f=100$  Hz (a), 250 Hz (b), 500 Hz (c), and 1 kHz (d). Here, each vertical slice is normalized by its maximum value so that the vertical growth and migration of the retrofocus with increasing  $R$  is emphasized. Dark regions lie near the local field maximum at any  $R$ . As expected, higher frequencies produce tighter retrofoci.

depths of  $z_s=25$  m (Table Ia) and  $z_s=50$  m (Table Ib). These results show that  $TL_{c\omega}$  is not always monotonically increasing with increasing  $R$ . Table I also lists the depth of the retrofocus peak,  $z_{rp}$  (which indicates the vertical retrofocus drift), and the amplitude of the largest vertical sidelobe in dB relative to the retrofocus peak amplitude. For the shallower source, sidelobe amplitudes generally increase with

increasing source-array range until absorption losses in the modal propagation cause the retrofocus location to shift. At this juncture, sidelobe amplitudes sometimes exceed that of the intended retrofocus (for example:  $f=1$  kHz,  $z_s=25$  m, and  $R=40$  km). As  $R$  is increased further, the retrofocus eventually settles onto the lowest propagating mode and sidelobes no longer exist [see Fig. 8(a)–(c)]. Sidelobes are

TABLE I. TRA retrofocus parameters for various source-array ranges ( $R$ ) and frequencies ( $f$ ) for two source depths:  $z_s=25$  m (a) and  $z_s=50$  m (b). The TRA transmission loss,  $TL_{c\omega}$ , is computed from (4) for the peak value of the retrofocus field through the depth. The vertical location of this retrofocus peak is  $z_{rp}$ . The sidelobe ratio is based on the strongest sidelobe, and a negative sidelobe ratio implies that the sidelobe amplitude is less than the retrofocus peak. If sidelobes are absent or the sidelobe ratio is less than  $-70$  dB, table entries are left blank.

Range (km)	Frequency											
	100 Hz			250 Hz			500 Hz			1000 Hz		
	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)
(a) Source depth=25 m												
1.0	37.3	23.5	-20.1	37.0	25.5	-15.5	37.0	25.0	-15.0	37.6	25.0	-7.1
2.0	41.7	25.5	-19.2	41.6	25.5	-17.1	41.8	25.0	-16.1	41.9	25.0	-9.1
5.0	49.4	29.0	-17.1	50.0	26.0	-24.7	50.1	25.0	-16.5	50.7	25.0	-11.3
10.0	56.1	39.5		58.7	27.5	-15.1	60.0	25.0	-12.5	60.2	25.0	-9.1
20.0	61.9	51.0		68.0	46.5		74.1	26.0	-11.6	73.7	25.5	-4.2
30.0	67.9	53.0		66.1	52.0		82.3	36.5	+0.7	86.3	26.5	-0.6
40.0	74.1	53.0		66.5	54.5		77.2	48.0	-38.2	94.4	29.0	+3.7
50.0	80.2	53.5		68.7	55.5		73.1	51.0	-57.0	103.6	29.5	-2.0
(b) Source depth=50 m												
1.0	35.9	50.0	-18.5	36.1	50.0	-19.9	36.1	50.0	-17.4	35.3	50.0	-13.3
2.0	40.3	51.0	-22.5	40.2	50.0	-21.9	40.2	50.0	-20.1	40.4	50.0	-12.6
5.0	46.8	50.5	-32.3	46.6	50.0	-29.4	46.6	50.0	-15.8	46.7	50.5	-17.0
10.0	52.7	52.5		52.0	51.0	-37.5	52.0	50.0	-37.8	52.5	50.0	-22.7
20.0	60.8	53.5		57.8	53.5		57.6	51.0	-54.3	57.8	50.0	-36.6
30.0	67.7	53.5		61.7	55.5		61.0	52.5		61.2	50.5	-35.2
40.0	74.1	53.5		65.1	56.0		63.4	54.5		63.6	51.0	-40.9
50.0	80.2	53.5		68.2	56.0		65.4	56.0		65.7	51.5	

TABLE II. TRA retrofocus parameters for various source–array ranges ( $R$ ) and bottom absorption levels for two source depths:  $z_s = 25$  m (a) and  $z_s = 50$  m (b) at a frequency of 500 Hz. Other parameters are the same as for Table I.

Range (km)	Quarter attenuation			Half attenuation			Nominal attenuation			Double attenuation		
	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)	$TL_{c\omega}(z_{rp})$ (dB)	$z_{rp}$ (m)	Sidelobe ratio (dB)
(a) Source depth=25 m												
1.0	35.6	25.0	-12.7	36.1	25.0	-13.7	37.0	25.0	-15.0	38.5	25.0	-15.4
2.0	39.4	25.0	-13.9	40.3	25.0	-14.8	41.8	25.0	-16.1	44.5	25.0	-17.1
5.0	44.6	25.0	-18.9	46.7	25.0	-19.6	50.1	25.0	-16.5	56.0	25.5	-13.3
10.0	50.1	25.0	-13.8	53.7	25.0	-15.5	60.0	25.0	-12.5	70.8	26.0	-5.4
20.0	56.3	25.0	-16.0	62.7	25.5	-14.0	74.1	26.0	-11.6	75.5	48.5	-34.3
30.0	61.3	25.5	-14.4	70.3	25.5	-7.9	82.3	36.5	+0.7	68.2	52.5	
40.0	66.1	25.5	-13.8	77.6	26.5	-8.8	77.2	48.0	-38.2	68.5	56.0	
50.0	69.6	25.5	-10.5	83.1	27.5	0.0	73.1	51.0	-57.0	70.5	57.5	
(b) Source depth=50 m												
1.0	35.2	50.0	-13.3	35.5	50.0	-14.6	36.1	50.0	-17.4	37.1	50.0	-22.4
2.0	38.5	50.0	-14.1	39.1	50.0	-17.1	40.2	50.0	-20.1	41.7	50.0	-18.2
5.0	43.5	50.0	-17.4	44.8	50.0	-18.3	46.6	50.0	-15.8	48.8	50.0	-25.2
10.0	47.8	50.0	-21.9	49.6	50.0	-17.9	52.0	50.0	-37.8	54.6	51.0	-51.9
20.0	52.6	50.0	-17.5	55.0	50.0	-35.8	57.6	51.0	-54.3	60.2	54.0	
30.0	55.7	50.0	-34.3	58.3	50.5	-61.2	61.0	52.5		63.8	56.5	
40.0	58.1	50.0	-38.6	60.7	51.0	-57.9	63.4	54.5		66.9	58.0	
50.0	59.9	50.0	-44.4	62.6	51.5	-59.9	65.4	56.0		69.7	58.5	

uniformly small for the deeper source. In addition, examination of Figs. 8 and 9 and Table I shows that when the retrofocus occurs within a wavelength or two of the intended depth ( $z_s$ ) the shallower source generally produces a quieter but smaller retrofocus.

Taken together, these results clearly show that the shape and number of propagating modes in the sound channel are critical to forming a robust retrofocus. The shallower retrofocus requires acoustic penetration of the strongly downward-refracting portion of the sound-speed profile, and this is only possible with the highest-order propagating modes. A higher acoustic frequency (or deeper sound channel) allows more of the highest-order modes to propagate further. Retrofocusing at the deeper source does not require acoustic modes that penetrate into the strongly downward-refracting portion of the sound-speed profile, and thus is less sensitive to the presence or lack of the highest-order modes. This point is supported by the retrofocus size results in Figs. 8 and 9, which show that deeper retrofoci tend to have larger vertical extent.

To confirm these interpretations of the relative amplitudes and sizes of the TRA retrofocus, a parameter study on the bottom attenuation level was undertaken at an acoustic frequency of  $f = 500$  Hz with different attenuation values in the two bottom layers. In practice, of course, once an ocean site has been selected there is no control over these variables, but understanding the trends may lead to more effective use of TRA systems. The results of this attenuation study are presented in Table II for one quarter (0.25 and 0.10 dB/ $\lambda$ ), one half (0.50 and 0.20 dB/ $\lambda$ ), one (1.00 and 0.40 dB/ $\lambda$ ), and two (2.00 and 0.80 dB/ $\lambda$ ) times the nominal attenuation values for source depths of  $z_s = 25$  m (Table IIa) and  $z_s = 50$  m (Table IIb). When the retrofocus occurs near its intended location, greater bottom loss always leads to lower retrofocus amplitudes. However, comparisons within Table II show that

the deeper source is much less influenced by the level of bottom losses, which confirms the description of sound-channel retrofocusing given above.

From these studies of retrofocusing in static sound channels, a common theme emerges: deeper sources lying below the strongly downward-refracting portion of the sound-speed profile generally lead to a louder, larger TRA retrofocus independent of the acoustic frequency, source–array range, and attenuation level. Results given in the next section also show that the TRA retrofoci produced by shallow and deep sources have differing properties in dynamic shallow-water environments.

### III. DYNAMIC SOUND CHANNEL RESULTS

This section presents TRA retrofocusing results for a dynamic sound channel that contains a linear superposition of synthetic internal waves whose temporal spectrum was matched to measurements taken during the SWARM experiment. Throughout this section, the static characteristics of the sound channel are the same as those described in Sec. I and shown in Fig. 1. The main effect of dynamic evolution of the sound channel is to degrade the retrofocus amplitude during the time delay ( $t_2 - t_1$ ) between the source broadcast (time= $t_1$ ) and the formation of the retrofocus (time= $t_2$ ). The independent parameters explored in this section are source depth, source–array range, acoustic frequency, internal wave strength, and time delay.

Figure 10 shows a sample of how the TRA retrofocus field degrades for  $z_s = 25$  m,  $f = 500$  Hz, and  $R = 10$  km. The first frame [Fig. 10(a)] corresponds to the ideal case of retrofocusing in a *frozen* ocean containing SWARM-strength internal waves (i.e.,  $t_2 - t_1 = 0$ ). The second, third, and fourth frames [Fig. 10(b)–(d)] show the retrofocus field at time delays of  $t_2 - t_1 = 1, 2,$  and  $5$  min, respectively. The gray scale

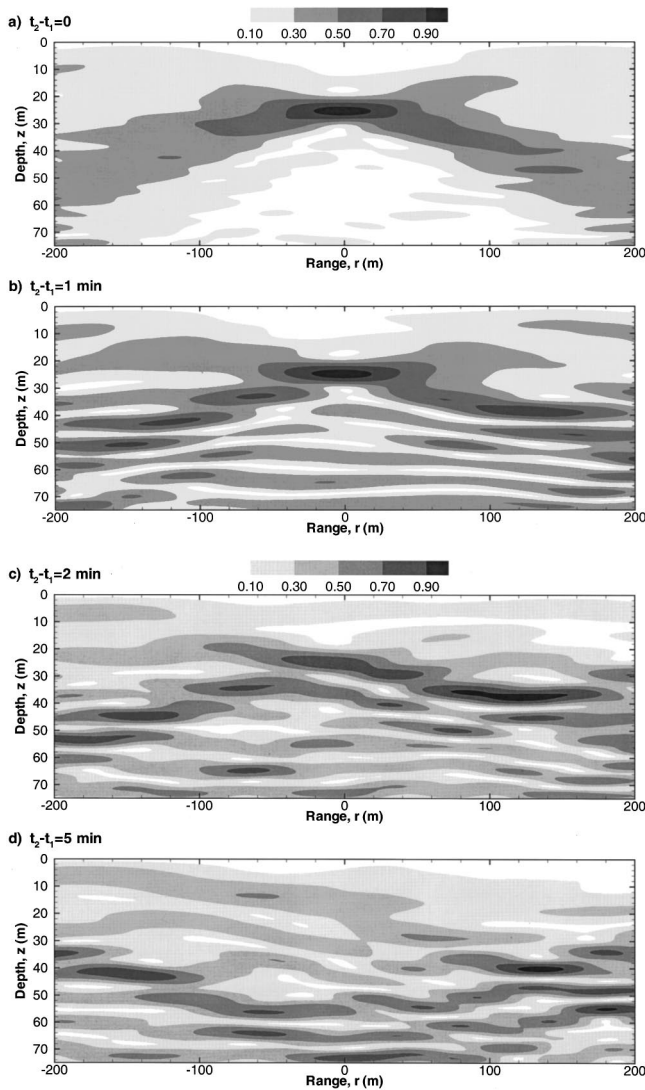


FIG. 10. Sample retrofocus fields with SWARM-strength synthetic internal waves for a source depth of  $z_s=25$  m, an acoustic frequency of  $f=500$  Hz, and a source–array range of  $R=10$  km. The time delay is  $t_2-t_1=0, 1, 2,$  and  $5$  min in frames (a), (b), (c), and (d), respectively. This series of frames shows how the retrofocus is degraded in a dynamic sound channel. Comparison of frame (a) with Fig. 6 shows that the internal waves lead to a slight decrease in the retrofocus size and a reduction of the nearby sidelobes.

is the same in all three frames. Comparison of these three frames shows that the retrofocus is degraded by amplitude loss, sidelobe growth, and position migration, but not by an obvious change in retrofocus size. In fact, a comparison of Fig. 10(a) and Fig. 6 shows that the internal waves lead to only a slight reduction in the size of the retrofocus and a somewhat stronger suppression of the nearest sidelobes. Hence, the retrofocus amplitude is chosen to be the main performance metric as  $t_2-t_1$  increases. It would certainly be critical to any underwater communication scheme based on amplitude modulation.

Figure 11 provides a sample of how the amplitude at the intended retrofocus location changes with increasing  $t_2-t_1$  when  $R=10$  km at four internal-wave energy levels for  $f=100$  Hz [Fig. 11(a)], 250 Hz [Fig. 11(b)], 500 Hz [Fig. 11(c)], and 1 kHz [Fig. 11(d)], and a source depth of  $z_s$

$=25$  m. Here, the four internal-wave energy levels have been obtained by multiplying the value of  $E_0$  obtained by matching the synthetic internal-wave spectrum to the measured SWARM spectrum by: 4, 1, 0.1, and 0.01, adjusting  $\eta$ , and then recomputing the forward and backward propagating sound fields at the same time delays. All the results in Fig. 11 are for one internal wave realization. The black dots in Fig. 11(c) denote the locations of the retrofocus fields shown in Fig. 10(b)–(d). As expected, the more active internal-wave fields and the higher frequencies lead to the quickest loss of amplitude. However, the field decay is not always monotonic and there are cases where the internal waves slightly enhance the retrofocus or cause it to reemerge after a hiatus of many minutes [see Fig. 11(b)]. Similar time-dependent features of shallow-water TRA retrofocusing have been observed during the experiments in the Mediterranean Sea (Song, 1998).

Results similar to those shown in Fig. 11 are obtained at  $z_s=50$  m, except that the retrofocus amplitude decay rates are different. When results like these are averaged over multiple realizations for fixed internal-wave energy level at a variety of source–array ranges, a performance envelope for a water-column-spanning TRA can be plotted. Figure 12 presents such performance envelopes for SWARM-level internal waves,  $0.5 \text{ km} \leq R \leq 50 \text{ km}$ ,  $0.2 \leq t_2-t_1 \leq 300$ ,  $f=100$  Hz, 250 Hz, 500 Hz, and 1 kHz, and source depths of  $z_s=25$  m [Fig. 12(a)] and  $z_s=50$  m [Fig. 12(b)]. Here, the plotted curves are the loci of the points where the amplitude at the intended retrofocus location has fallen by a factor of 2 (6 dB) from its value when  $t_2-t_1=0$ . These results are based on 20 internal-wave realizations with the extreme highest and lowest retrofocus ranges ignored. Acceptable focusing should occur below and to the left of these lines, while retrofocusing is marginal or not assured above and to the right of them. The horizontal error bars denote the 50%-confidence interval for each curve. As the widths of these intervals suggest, considerable variability was found, especially for the deeper source [Fig. 12(b)]. These half-life curves are truncated when the retrofocus is lost because of absorption. The black dots in Fig. 12(a) denote the locations of the retrofocus fields shown in Fig. 10(b)–(d). The upward-sloping diagonal line in both frames is the causality boundary. Realizable TRA systems must operate above and to the left of this line.

Comparison of the two frames of Fig. 12 shows that, for sufficiently long time delays (greater than 10–20 min at 1 kHz, for example), the shallower source leads to a more robust retrofocus. However, the deeper source generally allows longer-range retrofocusing for time delays less than 10 min. The explanation of this changeover may lie in the fact that all important propagating modes must arrive in phase at the source location for proper retrofocusing. Shallow-source retrofocusing relies primarily on the higher-order modes, while deeper-source retrofocusing utilizes all propagating modes (this is why deeper retrofoci are also louder). Hence, internal-wave-produced scattering and phase perturbations of the low-order modes do not influence the shallow retrofocus as much as the deeper one because the shallow retrofocus is not produced by as many modes. Thus, the deeper retrofocus



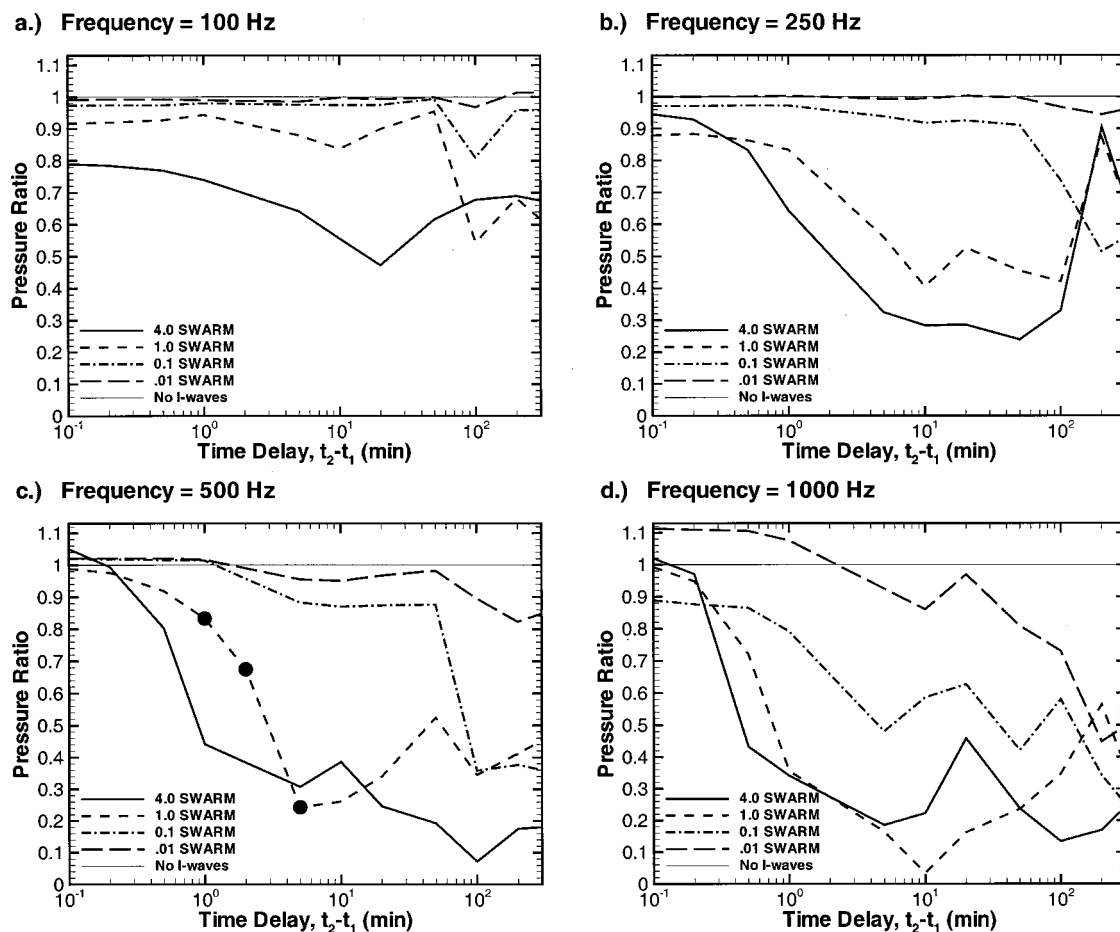


FIG. 11. Sample retrofocus field amplitude versus time delay for four internal-wave energy levels when  $R = 10$  km. The vertical axis is the retrofocus field amplitude at  $z = z_s = 25$  m divided by that found without internal waves. The various curves are for 4.0 (—), 1.0 (---), 0.1 (-·-·-), and 0.01 (----) times the measured SWARM internal-wave energy level. The four frames are for acoustic frequencies of 100 Hz (a), 250 Hz (b), 500 Hz (c), and 1 kHz (d). The retrofocus persists for longer periods of time at lower frequencies and lower internal-wave activity levels. The black dots in (c) correspond to the fields shown in Fig. 10(b)–(d).

decays more quickly because it is sensitive to internal-wave-induced perturbations of all propagating modes.

#### IV. SUMMARY AND CONCLUSIONS

This paper reports the findings of a broad computational study of time-reversing array retrofocusing in the shallow ocean with a water-column-spanning array. The influences of source depth, source–array range, acoustic frequency, bottom absorption, internal-wave strength, and time delay have all been presented. Although the many specific parametric and geometric choices have been guided by measurements made during the SWARM experiment, the parametric studies presented here should encompass a large fraction of all shallow ocean environments, and should be readily extendable to mildly range-dependent environments.

There are three main conclusions that can be drawn from this study. First, modal propagation sets limits on TRA retrofocusing in a static sound channel. When the acoustic frequency is higher and more modes are available, TRA retrofocusing is better and extends to longer ranges. Similarly, when the initial source lies in the lower portion of the sound channel where the speed of sound gradient is weak and many modes have both nodes and antinodes, time-reversing array

retrofocusing is enhanced compared to that computed for shallow sources that primarily excite more rapidly attenuated high-order acoustic modes. In fact, the ultimate source–array range for TRA retrofocusing in quiescent shallow-water sound channels will be set by mode stripping caused by bottom and bubble absorption, and scattering losses from the bubbles and the ocean’s free surface.

The second main conclusion is that internal-wave-induced fluctuations actually aid TRA retrofocusing for short time delays (a minute or less at source–array ranges of 10 km) mainly by suppressing sidelobes, but do eventually degrade retrofocusing at longer times (tens of minutes at source–array ranges of 10 km). The beneficial effects come about because internal waves induce mode coupling that allows low-order modes to properly scatter into higher orders as the backward-propagating signal nears the original source location. Such scattering increases the acoustic energy in the higher-order modes compared to the same signal in a sound channel without internal waves, and leads to a slightly smaller retrofocus with lower sidelobes. However, the beneficial effects of this internal-wave-induced mode coupling are lost when the time delay is large enough for the internal-wave dynamics to significantly change the refractive index structure in the sound channel. As expected, significant

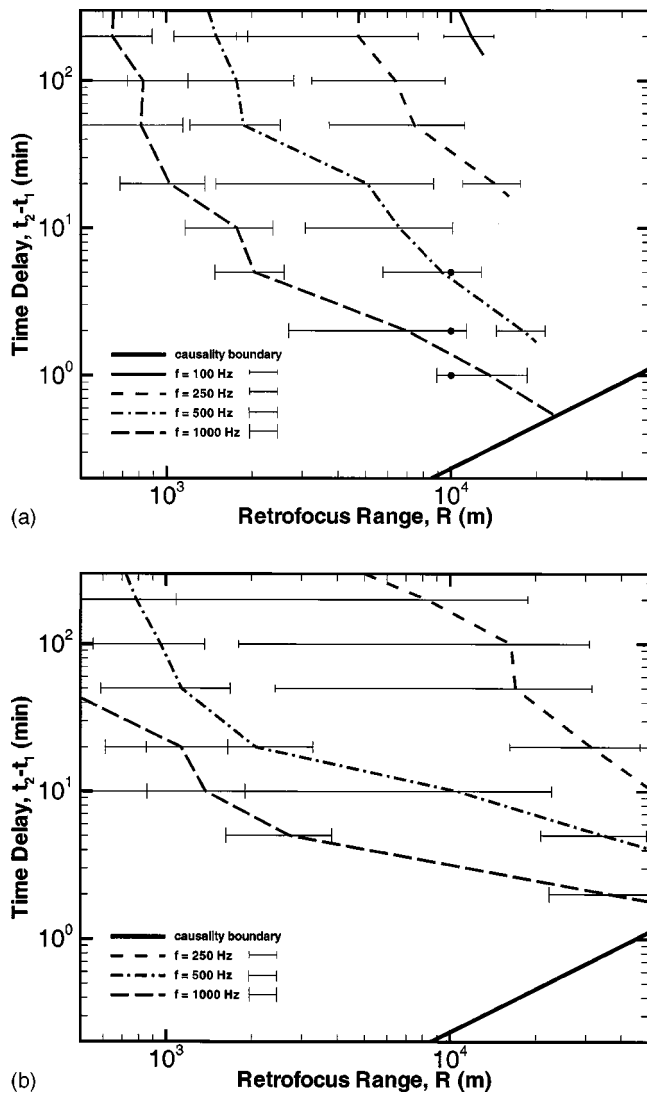


FIG. 12. Time-reversing array performance envelopes for variable source–array range ( $R$ ) and time delay ( $t_2 - t_1$ ) with SWARM-strength internal waves. The contours are the loci of the retrofocus half life at each frequency  $f = 100$  Hz (—), 250 Hz (---), 500 Hz (-·-·-), and 1 kHz (·-·-·). The plotted curves correspond to the average of 20 trials. Above and to the right of these contours, the TRA retrofocus amplitude is predicted to have fallen to half its value in a frozen shallow ocean ( $t_2 - t_1 = 0$ ). The upward-sloping diagonal line is the causality boundary. Realizable TRA systems must operate above and to the left of this line. Frame (a) is for a source depth  $z_s = 25$  m, and frame (b) is for a source depth  $z_s = 50$  m. The deeper source leads to retrofocusing at longer ranges, but the shallow source produces a longer-lasting retrofocus. The 100-Hz curve lies off the figure in (b). The error bars denote the 50%-confidence interval for each half-life curve. The black dots in (a) correspond to the fields shown in Fig. 10(b)–(d).

changes occur more rapidly at longer source–array ranges and higher acoustic frequencies. Decreasing the frequency increases the delay time before internal-wave degradation of the retrofocus but also increases the retrofocus size. However, some minimum number of propagating modes (ten or so) is required to retrofocus properly (depending on source depth), and this sets a lower limit on the acoustic frequency that could be used for successful TRA operations.

Finally, for an active internal-wave environment such as that found during the SWARM experiment, the source–array range limitation imposed by the internal waves is likely to be more severe than that imposed by bottom losses. In less ac-

tive environments, absorption losses are likely to set the ultimate source–array range.

## ACKNOWLEDGMENTS

This work was sponsored by the Office of Naval Research Ocean Acoustics Program. We also want to thank Professor William Kuperman for suggesting this study, Dr. Michael Collins for providing RAM and his insights for properly using it, Dr. Darrell Jackson for assistance with the internal wave model development, and Dr. Steven Finette for generously providing the SWARM measurements.

- Apel, J. R., Badiey, M., Chiu, C.-S., Finette, S., Headrick, R., Kemp, J., Lynch, J. F., Newhall, A., Orr, M. H., Pasewark, B. H., Tielburger, D., Turgut, A., von der Heydt, K., and Wolf, S. (1997). "An overview of the 1995 SWARM shallow-water internal wave acoustic scattering experiment," *IEEE J. Ocean Eng.* **22**, 465–500.
- Collins, M. D. (1993). "A split-step Pade solution for parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742.
- Collins, M. D. (1994). "Generalization of the split-step Pade solution," *J. Acoust. Soc. Am.* **96**, 382–385.
- Collins, M. D. (1998). "New and improved parabolic equation models," *J. Acoust. Soc. Am.* **104**, Pt. 2, 1808.
- Dashen, R., Munk, W. H., Watson, K. M., and Zachariassen, F. (1979). In *Sound Transmission Through a Fluctuating Ocean*, edited by S. M. Flatte (Cambridge University Press, Cambridge).
- Derode, A., Roux, P., and Fink, M. (1995). "Robust acoustic time reversal with high-order multiple scattering," *Phys. Rev. Lett.* **75**, 4206–4209.
- Dowling, D. R., and Jackson, D. R. (1992). "Narrow-band performance of acoustic phase-conjugate arrays in dynamic random media," *J. Acoust. Soc. Am.* **91**, 3257–3277.
- Dowling, D. R. (1993). "Phase-conjugate array focusing in a moving medium," *J. Acoust. Soc. Am.* **94**, 1716–1718.
- Dowling, D. R. (1994). "Acoustic pulse compression using passive phase-conjugate processing," *J. Acoust. Soc. Am.* **95**, 1450–1458.
- Fink, M. (1997). "Time reversed acoustics," *Phys. Today* **50**, No. 3, 34–40.
- Finette, S. (1998). Private communication.
- Garrett, C., and Munk, W. (1979). "Internal waves in the ocean," *Annu. Rev. Fluid Mech.* **11**, 339–369.
- Hodgkiss, W. S., Kuperman, W. A., and Song, H. C. (1997). Private communication.
- Jackson, D. R., and Dowling, D. R. (1991). "Phase-conjugation in underwater acoustics," *J. Acoust. Soc. Am.* **89**, 171–181.
- Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H. (1994). *Computational Ocean Acoustics* (AIP, New York), p. 41.
- Khosla, S. R., and Dowling, D. R. (1998). "Time-reversing array retrofocusing in simple dynamic underwater environments," *J. Acoust. Soc. Am.* **104**, 3339–3350.
- Kuperman, W. A. (1997). Private communication.
- Kuperman, W. A., Hodgkiss, W. S., Song, H. C., Akal, T., Ferla, C., and Jackson, D. R. (1998). "Phase-conjugation in the ocean: Experimental demonstration of an acoustic time reversal mirror," *J. Acoust. Soc. Am.* **103**, 25–40.
- Parvulescu, A. (1995). "Matched-signal (MESS) processing by the ocean," *J. Acoust. Soc. Am.* **98**, 943–960.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1986). *Numerical Recipes* (Cambridge University Press, Cambridge), Chap. 16.
- Rose, J. H., Bilgen, M., Roux, P., and Fink, M. (1999). "Time-reversal mirrors and rough surfaces: Theory," *J. Acoust. Soc. Am.* **106**, 716–723.
- Roux, P., De Rosny, J., and Fink, M. (1999). "Time reversal mirrors and rough surfaces: Experiment," *J. Acoust. Soc. Am.* **106**, 724–732.
- Saichev, A. I. (1982). "Effect of compensating distortions due to scattering in an inhomogeneous medium by the use of a reflector turning the front," *Radio Eng. Electron. Phys.* **27**, 23–30.
- Song, H. C. (1998). Private communication.
- Song, H. C., Kuperman, W. A., and Hodgkiss, W. S. (1998). "A time-reversal mirror with variable range focusing," *J. Acoust. Soc. Am.* **103**, 3234–3240.

- Tanter, M., Thomas, J.-L., and Fink, M. (1998). "Focusing and steering through absorbing and aberrating layers: Application to ultrasonic propagation through the skull," *J. Acoust. Soc. Am.* **103**, 2403–2410.
- Tielbürger, D., Finette, S., and Wolf, S. (1997). "Acoustic propagation through an internal wave field in a shallow water waveguide," *J. Acoust. Soc. Am.* **101**, 789–808.
- Yang, T. C., and Yoo, K. (1998). "Frequency spectrum of linear internal waves in shallow water and a modified Garrett–Munk model," *J. Acoust. Soc. Am.* **104**, 1765(A).
- Yönak, S. H., and Dowling, D. R. (1999). "Photoacoustic detection and localization of small gas leaks," *J. Acoust. Soc. Am.* **105**, 2685–2694.
- Yoo, K., and Yang, T. C. (1998). "Measurements of modal spectrum of linear internal waves in shallow water and comparison with the Garrett–Munk model," *J. Acoust. Soc. Am.* **104**, 1765(A).

# Cell model calculations of dynamic drag parameters in packings of spheres

Olga Umnova

*Department of Environmental and Mechanical Engineering, Faculty of Technology, The Open University, Milton Keynes MK7 6AA, United Kingdom*

Keith Attenborough<sup>a)</sup>

*School of Engineering, The University of Hull, Hull HU6 7RX, United Kingdom*

Kai Ming Li

*Department of Mechanical Engineering, Hong Kong Polytechnic University, Kowloon, Hong Kong*

(Received 26 October 1998; accepted for publication 16 February 2000)

An external flow approach is used to predict the viscous drag due to oscillating flow in an air-filled stack of fixed identical rigid spheres. Analytical expressions for dynamic and direct current (dc) permeability, high-frequency limit of tortuosity, and the characteristic viscous dimension are derived using a cell model with an adjustable cell radius which allows for hydrodynamic interactions between the spherical particles. The resulting theory requires knowledge of two fixed parameters: the volume porosity and the particle radius. The theory also requires a value for the cell radius. Use of the cell radius corresponding to that of the sphere circumscribing a unit cell of a cubic lattice arrangement is proposed. This is found to enable good agreement between predictions of the new theory and both published data and numerical results for simple cubic and random spherical packings. © 2000 Acoustical Society of America. [S0001-4966(00)04405-2]

PACS numbers: 43.35.Bf [HEB]

## INTRODUCTION

Models for the acoustical properties of porous granular materials may be classified as “internal” and “external” flow models. Internal flow models (or pore-based models) derive expressions for the viscous force that assumes that the fluid phase exists in tubular pores in the rigid frame. Two recent pore-based models for rigid-framed absorbers<sup>1,2</sup> have been shown to give improved agreement with acoustical data compared with simpler models that assume uniform identical pores. One of them<sup>1</sup> requires information about porosity, tortuosity, flow resistivity, and the pore geometry. The pore geometry is described by two characteristic dimensions. These may be deduced from laboratory measurements, including the Brunauer–Emmett–Teller (BET) method, or approximated from other measured parameters. The other model<sup>2</sup> requires information about the “pore” size distribution. This may be obtained through use of laboratory techniques such as water suction or mercury injection for granular materials. These measurements are tedious and time consuming if the “pores” are small. The question arises of whether it is possible to derive alternative models for the acoustic behavior of porous granular materials based, for example, on knowledge of the constituent grain sizes and packing characteristics.

External-flow models consider the fluid flow around constituent particles rather than in hypothetical pores. They require knowledge of porosity (or particle volume fraction), particle shape, particle size, and the properties of the packing. The drag force between phases in dilute suspensions is calculated as the sum of the forces between the moving fluid

and each particle. In concentrated suspensions and granular materials, there will be interactions between the flows around particles. Cell models such as those by Kuwabara<sup>3</sup> and Happel,<sup>4</sup> modified by Strout,<sup>5</sup> for the oscillatory motion of the fluid have been developed to account for such interactions. In these models, each cell consists of a central particle core and a concentric shell of fluid. The hydrodynamic interactions between particles are taken into account through the boundary conditions on the outer surface of the cell. In the Kuwabara model, the vorticity of fluid velocity is supposed to be zero. In the Happel model the shear stress is assumed zero at the cell boundary. Cell models have been found to give good predictions for the acoustical properties of concentrated suspensions of spherical particles.<sup>5</sup> A cell model, with additional averaging over the distribution of the Voronoi cells, has been applied to calculations of the flow resistivity of fibrous materials.<sup>6</sup> In highly concentrated systems the applicability of the cell model approach becomes questionable. Neighboring solid particles will be in contact and hence will penetrate the hypothetical cell around each individual particle. Nevertheless, in this paper, we will show that adjustment of the cell radius enables reasonable values of DC permeability, tortuosity, and the characteristic viscous dimension of a material consisting of rigidly fixed stacked spheres, even for some types of close packing. In fact, we propose use of a cell radius that has a simple geometrical interpretation.

Section I reinterprets the high- and low-frequency limits of the complex tortuosity function derived by Johnson *et al.*<sup>7</sup> in terms of dynamic drag. In Sec. II these results are compared with the corresponding limits of a modified form of the Kuwabara/Strout cell model. This enables derivation of explicit analytical expressions for the acoustical constants and

<sup>a)</sup>Author to whom correspondence should be addressed.

the geometrical factor as functions of porosity, particle radius, and cell radius. In Sec. III we find an appropriate expression for cell radius and compare the predictions of the resulting theory with published data and with numerical results obtained independently for different types of packing. Section IV offers concluding remarks.

## I. COMPLEX DENSITY, TORTUOSITY, AND PERMEABILITY IN THE EXTERNAL FLOW MODEL

Consider a fluid-filled granular material of porosity  $\phi$  and consisting of rigid identical spherical particles of radius  $R$ . Suppose that the solid phase density is much greater than that of the fluid (continuous phase). In this case, the momentum conservation equation for the fluid is

$$\phi \rho_0 \frac{\partial v}{\partial t} = -\phi \frac{\partial p}{\partial x} - D(\omega)v, \quad (1)$$

where  $\rho_0$  is the fluid density,  $p$  is the pressure variation in the sound wave,  $v$  is the fluid velocity averaged over unit volume of the fluid, and  $D(\omega)$  is the frequency-dependent drag coefficient.

The waves are supposed to be harmonic and, consequently,  $\partial/\partial t \rightarrow -i\omega$ .

The complex density of the fluid is defined by

$$\rho(\omega) = -\frac{\partial p}{\partial x} \frac{1}{-i\omega v} = \rho_0 \left( 1 + \frac{D(\omega)}{-i\omega \rho_0 \phi} \right). \quad (2a)$$

The complex density in a rigid-framed porous material is related to the complex tortuosity  $\alpha(\omega)$  and complex permeability  $k(\omega)$  defined by Johnson,<sup>7</sup>

$$\begin{aligned} \alpha(\omega) &= \frac{\rho(\omega)}{\rho_0}, \\ k(\omega) &= \frac{i\eta\phi}{\omega\rho(\omega)} = \frac{i\eta\phi}{\omega\rho_0(1 + (-D(\omega)/i\omega\rho_0\phi))}. \end{aligned} \quad (2b)$$

Johnson<sup>7</sup> has considered the analytical properties of the linear response function  $\alpha(\omega)$  as well as its low- and high-frequency limits viz.,

$$\begin{aligned} \lim_{\omega \rightarrow 0} \alpha(\omega) &= \frac{\phi}{k_0} \frac{i\eta}{\omega\rho_0}, \\ \lim_{\omega \rightarrow \infty} \alpha(\omega) &= \alpha_\infty + \left( \frac{i\eta}{\omega\rho_0} \right)^{1/2} \frac{2\alpha_\infty}{\Lambda}, \end{aligned} \quad (3)$$

where  $\alpha_\infty$  is the high-frequency limit of tortuosity (known simply as tortuosity in many publications; for example, see Ref. 2),  $k_0$  is the DC permeability, and  $\Lambda$  is the characteristic viscous dimension of the material.

From (3) and (2), the following limits for the drag term can be obtained:

$$\lim_{\omega \rightarrow 0} D(\omega) = \frac{\phi^2 \eta}{k_0}, \quad (4a)$$

$$\lim_{\omega \rightarrow \infty} D(\omega) = -i\omega\rho_0\phi \left[ (\alpha_\infty - 1) + \left( \frac{i\eta}{\omega\rho_0} \right)^{1/2} \frac{2\alpha_\infty}{\Lambda} \right]. \quad (4b)$$

## II. THE DRAG FORCE IN THE PRESENCE OF HYDRODYNAMIC INTERACTIONS BETWEEN RIGID PARTICLES

Consider the oscillatory flow of incompressible viscous fluid with macroscopic velocity  $\mathbf{v}e^{-i\omega t}$  around a fixed spherical particle. The fluid velocity field around a sphere can be represented by<sup>8</sup>

$$\mathbf{u} = [\mathbf{v} + \text{curl} \text{curl}(f(r)\mathbf{v})]e^{-i\omega t},$$

where  $f(r)$  is the potential function.

The components of the fluid velocity in a polar coordinate system (with the polar axis  $z$  parallel to  $\mathbf{v}$ ) are

$$\begin{aligned} u_\theta &= -v \sin \theta \left( 1 - \frac{df}{rdr} - \frac{d^2f}{dr^2} \right), \\ u_r &= v \cos \theta \left( 1 - 2\frac{df}{rdr} \right). \end{aligned}$$

The potential function obeys the following equation:

$$\frac{d}{dr} \left( \Delta \Delta f - \frac{\omega\rho_0}{i\eta} \Delta f \right) = 0, \quad (5)$$

where  $\Delta = d^2/dr^2 + (2/r)d/dr$  is the radial part of the spherical Laplacian operator.

The nonslip boundary conditions ( $u_\theta|_R = u_r|_R = 0$ , where  $R$  is the particle radius) are applied on the particle surface. In terms of the potential function they can be rewritten as

$$\left. \frac{df}{rdr} \right|_R = \frac{1}{2}, \quad (6)$$

$$\Delta f|_R = \frac{3}{2}. \quad (7)$$

In the Kuwabara/Strout model<sup>3,5</sup> the vorticity of the fluid velocity is assumed to be zero at the outer cell boundary ( $\text{curl } \mathbf{u}|_b = 0$ , where  $b$  is the cell radius). In terms of the potential function that means that

$$\left. \frac{d\Delta f}{dr} \right|_b = 0. \quad (8)$$

In the original version of the Strout/Kuwabara model, the radial component of fluid velocity is assumed to vanish at the external cell boundary. Instead, we choose the value that results from matching the magnitude of the macroscopic fluid velocity  $v$  with the  $z$  component of the fluid particle velocity averaged over the cell volume. Hence

$$\begin{aligned} \langle u_z \rangle_v &= \frac{2\pi}{\frac{4}{3}\pi(b^3 - R^3)} \int_0^\pi \int_R^b (u_r(r, \theta) \cos \theta \\ &\quad - u_\theta(r, \theta) \sin \theta) r^2 \sin \theta d\theta dr \\ &= v - \frac{2v}{b^3 - R^3} \left( b^2 \left. \frac{df}{dr} \right|_b - \frac{R^3}{2} \right) = v. \end{aligned}$$

It should be noted that the nonslip boundary condition (6) has been used here. From this equation, the boundary condi-

tion for the radial component of the fluid velocity on the outer cell boundary can be obtained. In terms of potential function it has the form

$$\left. \frac{df}{rdr} \right|_b = \frac{\Theta}{2}, \quad (9)$$

where

$$\Theta = \left( \frac{R}{b} \right)^3. \quad (10)$$

The magnitude of the drag force,  $F$ , parallel to the direction of the velocity,  $v$ , is calculated from<sup>8</sup>

$$F = \int (-P \cos \theta + \sigma_{rr} \cos \theta - \sigma_{r\theta} \sin \theta) df,$$

where the integration is performed over the spherical particle surface.

The components of the stress tensor are

$$\sigma_{rr} = 2\eta \frac{\partial u_r}{\partial r}, \quad \sigma_{r\theta} = \eta \left( \frac{\partial u_r}{r \partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} \right).$$

The pressure  $P$  on the particle surface is calculated from the equations of fluid motion in the polar coordinate system and is connected with the potential function in the following way:

$$P = P_0 + \eta v R \cos \theta \times \left( \left. \frac{d\Delta f}{rdr} \right|_R - \Delta \Delta f|_R - \left( \frac{\omega \rho_0}{i \eta} \right) \left( \left. \frac{df}{rdr} \right|_R - \Delta f|_R \right) \right).$$

Using boundary conditions (6), (7) the expression for the drag force becomes

$$F = \frac{4}{3} \pi R^3 v \eta \left( -3 \left. \frac{d\Delta f}{rdr} \right|_R + \Delta \Delta f|_R - \frac{\omega \rho_0}{i \eta} \right).$$

The drag term is calculated as the sum of the drag forces on individual particles,

$$D(\omega) = \frac{nF}{v}, \quad (11)$$

where  $n = (1 - \phi)/(4/3) \pi R^3$  is the number of particles per unit volume.

As a result of solving Eq. (5) with boundary conditions (6)–(9), we get

$$D(\omega) = +i\omega\rho_0(1-\phi) \left( 1 + \frac{3}{2}(\Theta^{-1}-1) \times \frac{\exp(2s(\Theta^{-1/3}-1))A_1+A_2}{\exp(2s(\Theta^{-1/3}-1))B_1-B_2} \right), \quad (12)$$

where

$$\begin{aligned} A_1 &= (s\Theta^{-1/3}-1)(s^2+3s+3), \\ A_2 &= (s\Theta^{-1/3}+1)(s^2-3s+3), \\ B_1 &= (s\Theta^{-1/3}-1)(-s^2(\Theta^{-1}-1)+3s+3), \\ B_2 &= (s\Theta^{-1/3}+1)(s^2(\Theta^{-1}-1)+3s-3), \end{aligned}$$

and  $s = \sqrt{(\omega\rho_0/i\eta)R^2}$ .

The high-frequency limit of the drag term is

$$D(\omega \rightarrow \infty) \rightarrow -i\omega\rho_0(1-\phi) \left[ \frac{1}{2} + \frac{9}{2(1-\Theta)R} \left( \frac{i\eta}{\omega\rho_0} \right)^{1/2} \right].$$

Comparison of this expression with (4b) yields explicit expressions for the high-frequency limit of tortuosity and the characteristic viscous dimension in terms of porosity, grain radius, and cell radius,

$$\Lambda = \frac{4(1-\Theta)\phi\alpha_\infty}{9(1-\phi)}R, \quad (13)$$

$$\alpha_\infty = 1 + \frac{1-\phi}{2\phi}. \quad (14)$$

The latter expression coincides with the well-known result for tortuosity obtained by Berryman<sup>9</sup> and is independent of the cell radius.

Comparison of the low-frequency limit of  $D(\omega)$ , i.e.,

$$D(\omega \rightarrow 0) \rightarrow \frac{9\eta(1-\Theta)(1-\phi)\Omega_k}{2R^2},$$

with (4a) gives an explicit expression for dc permeability,

$$k_0 = \frac{2}{9} \frac{\phi^2}{(1-\phi)(1-\Theta)\Omega_k} R^2, \quad (15)$$

where

$$\Omega_k = \frac{5}{5-9\Theta^{1/3}+5\Theta-\Theta^2} \quad (16)$$

is Kuwabara's low-frequency correction factor.

These results for  $\Lambda$ ,  $\alpha_\infty$ , and  $k_0$  are exact to first non-trivial order in solid concentration  $1-\phi$ .

Our modification of the boundary condition for the radial component of fluid velocity does not change the expression for  $k_0$  and  $\Lambda$  that would have resulted from use of the original version of Strout/Kuwabara theory. However, it changes the predicted value of tortuosity significantly. The assumption of zero  $r$ -component of velocity at the outer cell boundary gives

$$\alpha_\infty = 1 + \frac{(1-\phi)}{2\phi}(1+2\Theta). \quad (14')$$

This expression leads to values of tortuosity that exceed the results of numerical calculations and data by a significant margin.

The geometrical factor is defined by

$$M = \frac{8\alpha_\infty k_0}{\Lambda^2 \phi}, \quad (17)$$

and corresponds to  $2n^2/s$  in the notation of Attenborough,<sup>10</sup> where  $n$  is the dynamic shape factor and  $s$  is the static shape factor. For materials with cylindrical pores

$$\alpha_\infty = \frac{1}{\cos^2 \theta}, \quad k_0 = \frac{1}{8} \phi a^2 \cos^2 \theta, \quad \Lambda = a,$$

where  $a$  is the pore radius and  $\theta$  is the angle between the pores and the sound propagation direction. Hence the geo-

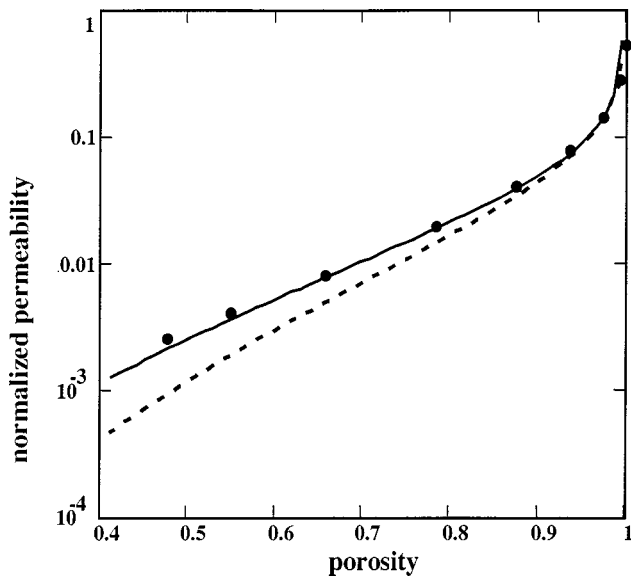


FIG. 1. Dimensionless dc permeability  $k_0/d^2$  [ $d=\sqrt[3]{(4\pi)/(3(1-\phi))}R$  is the base vector of the lattice] for SC packing as a function of porosity, circles—numerical results (Ref. 12), solid line—cell model (19), dashed lines—cell model (18).

metrical factor is 1 for a material with cylindrical pores. Allard<sup>1</sup> suggests that  $M$  could take values up to 10 for materials with noncylindrical pores. On the other hand, Johnson<sup>7</sup> has found that  $M=1$  gives the best agreement between his predictions and data for glass beads. Numerical simulations with an intersecting tubes model<sup>11</sup> indicate that  $M$  is approximately constant (i.e., changes within a factor of two) over the range of porosity that causes  $k_0$  to vary over eight orders of magnitude.

### III. COMPARISONS WITH NUMERICAL RESULTS AND DATA

So far we have not defined the parameter  $\Theta$ —the ratio of the particle volume to the cell volume. In the original version of the Kuwabara/Strout model it was taken to be equal to the volume fraction of the solid phase in the material,

$$\Theta = 1 - \phi. \quad (18)$$

Instead, we propose use of larger spherical cells based on spheres that circumscribe the unit cells of the packing. In this case for simple cubic (SC) packing,

$$\Theta = \frac{3}{\sqrt{2}\pi}(1-\phi) \cong 0.675(1-\phi), \quad (19)$$

Chapman and Higdon have calculated the drag parameters and the dynamic permeability of different packings of identical spherical particles numerically.<sup>12</sup>

In Fig. 1, their numerical predictions of DC permeability for SC packing for the range of porosities up to the close packing limit  $\phi_c=0.4764$  are compared with cell model results. If the model with modified cell radius (19) is used, there is tolerable agreement (within 18%) even at a porosity corresponding to close packing. With the unmodified cell radius (18), the agreement with numerical results is good

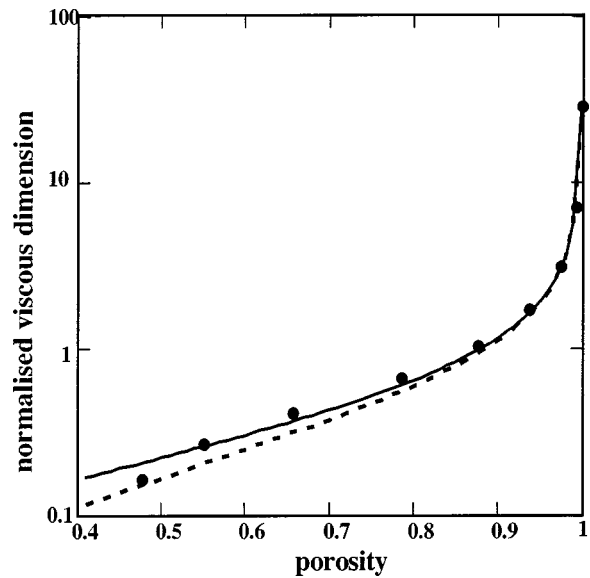


FIG. 2. Dimensionless characteristic viscous dimension  $\Lambda/d$  for SC packing as a function of porosity, circles—numerical results (Ref. 12) for SC packing, solid line—cell model (19) results, dashed lines—cell model (18) results.

only for very dilute systems. In Fig. 2 the corresponding comparisons for characteristic viscous dimension are presented. The results of cell model with modified radius are closer to numerical values than those with the unmodified radius, except for close packing. Overall, the error is not more than 20% for the modified cell model, whereas with the unmodified theory and some values of porosity, it exceeds 40%.

In fact, our choice (19) of the connection between porosity and the parameter  $\Theta$  in SC packing corresponds to the minimum of the mean value of the absolute difference between numerical results<sup>12</sup> and cell model estimates. With the cell radius given by  $\Theta/(1-\phi)=0.675$ , then, for DC permeability,

$$\varepsilon_k = \frac{1}{N} \sum_{i=1}^N \frac{|k_0^{\text{num}} - k_0^{\text{cell}}|}{k_0^{\text{num}}} \times 100\% = 5.8\%,$$

where  $N$  is the number of points with different porosity where comparisons are made. Similarly with  $\Theta/(1-\phi)=0.675$ , then, for characteristic viscous dimension,

$$\varepsilon_\Lambda = \frac{1}{N} \sum_{i=1}^N \frac{|\Lambda^{\text{num}} - \Lambda^{\text{cell}}|}{\Lambda^{\text{num}}} \times 100\% = 7.3\%.$$

If  $\Theta/(1-\phi)=0.6$ , then  $\varepsilon_k \cong 8.2\%$  and  $\varepsilon_\Lambda = 10.7\%$ , if  $\Theta/(1-\phi)=0.8$  then  $\varepsilon_k = 11.5\%$  and  $\varepsilon_\Lambda = 10.7\%$ .

In Fig. 3, the cell model predictions of formation factor  $F = \alpha_\infty / \phi$  are compared with numerical results. Since the value of formation factor does not depend on parameter  $\Theta$  in the cell model, both versions give the same results. The deviation between the numerical results and the cell model estimates does not exceed 12% even for close packing. Figure 4 shows that modified cell model predictions for magnitude and phase of dynamic permeability  $k(\omega)$  are also in reasonable agreement with numerical results<sup>12</sup> for SC packing.

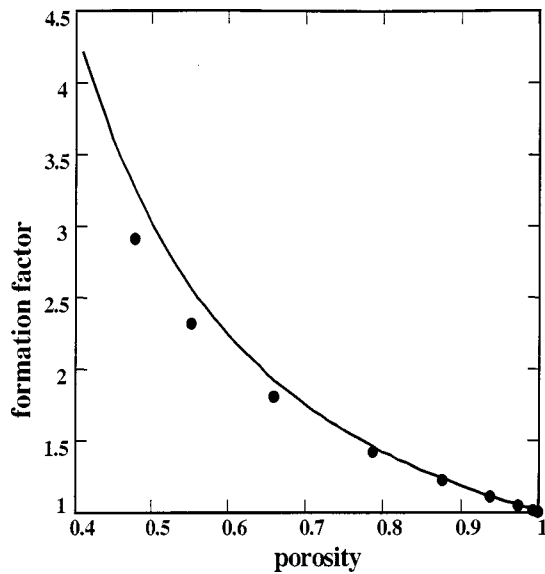


FIG. 3. Formation factor  $F = \alpha_{\infty} / \phi$  as a function of porosity, circles—numerical results (Ref. 12) for SC packing, solid line—cell model results.

Nevertheless, the results indicate that the cell model is not sufficiently accurate to reproduce the differences between the various forms of packing obtained numerically. Moreover, the modified cell model results differ significantly from the numerical estimates for close packings. This indicates that the cell model is not able to describe such dense systems.

Wong *et al.*<sup>16</sup> have presented data for the formation factor and dc permeability of fused glass beads with porosities between 0.023 and 0.399. For the highest porosity  $\phi = 0.399$ , the measured value of formation factor is close to 4. For this value of porosity, the modified cell model predicts  $F = 4.375$ . The measured value for dc permeability is approximately  $3 \times 10^{-12} \text{ m}^2$  for a mixture of glass beads with diameters between  $44 \times 10^{-6}$  and  $53 \times 10^{-6} \text{ m}$ . The modified cell model predicts  $1.95 \times 10^{-12} \text{ m}^2 \leq k_0 \leq 2.83 \times 10^{-12} \text{ m}^2$ , where the first figure corresponds to smaller particles.

For a mixture of beads with diameters between  $88 \times 10^{-6}$  and  $106 \times 10^{-6} \text{ m}$ , the measured value of DC permeability is approximately  $10^{-11} \text{ m}^2$ , and the modified cell model predicts  $7.79 \times 10^{-12} \text{ m}^2 \leq k_0 \leq 11.31 \times 10^{-12} \text{ m}^2$ .

Predictions of the cell model [with parameter  $\Theta$  calculated using either (18) or (19)] are compared with data<sup>13-15</sup> for permeability, tortuosity, and characteristic viscous dimension of random packings of bronze spheres and glass beads in Tables I and II. The number of nearest neighbors in close random packing is similar to that in (6) in simple cubic packing, so we use expression (19) to connect cell radius with porosity in random packing also. The predictions of the theory with modified cell radius agree well with two series of data.<sup>13,14</sup> However, there are discrepancies between predictions and the data for high-frequency constants of two data sets [Ref. 13, series 3 and Ref. 15]. Neither the original nor modified cell models give predictions that match these data in most cases. On the other hand, it is noticeable that the geometrical factors for these data have unusually large values. Moreover there is a significant difference between these data and those for tortuosity and characteristic viscous di-

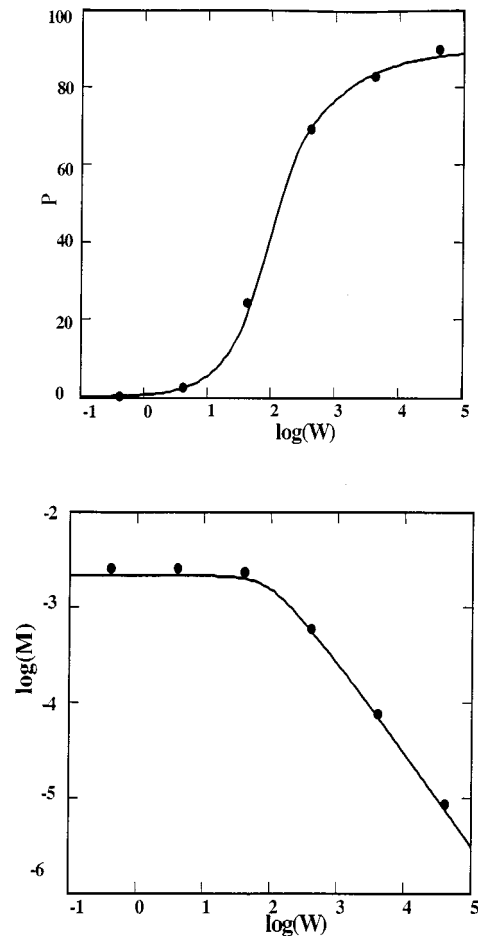


FIG. 4. Phase  $\tan^{-1}(\text{Im}(k(\omega))/\text{Re}(k(\omega)))$ , in degrees, and magnitude  $M(\omega) = |k(\omega)/d^2|$  of dynamic permeability as a function of dimensionless frequency  $W = (\omega \rho_0 d^2 / \eta)$  for SC lattice. Circles—numerical results (Ref. 12), solid line—cell model (19) results.

TABLE I. Comparison of cell model predictions and data for dc permeability, tortuosity, characteristic viscous dimension, and geometrical parameter in packings of bronze spheres. (Brackets indicate parameter values deduced from those of the others.)

Parameters	$k_0$ (mm <sup>2</sup> )	$\alpha_{\infty}$	$\Lambda$ (mm)	$M$	
$\phi = 0.43$ $R = 0.0375 \text{ mm}$	$7 * 10^{-6}$	(1.61)	$1.24 * 10^{-2}$	1.34	Bronze spheres data series 1 <sup>a</sup>
	$7.57 * 10^{-6}$	1.66	$1.29 * 10^{-2}$	1.42	Cell model (19) prediction
	$2.97 * 10^{-6}$	1.66	$0.90 * 10^{-2}$	1.14	Cell model (18) prediction
$\phi = 0.41$ $R = 0.055 \text{ mm}$	$12.00 * 10^{-6}$	(1.97)	$1.79 * 10^{-2}$	1.44	Bronze spheres data series 2 <sup>a</sup>
	$13.40 * 10^{-6}$	1.72	$1.76 * 10^{-2}$	1.74	Cell model (19) prediction
	$5.00 * 10^{-6}$	1.72	$1.2 * 10^{-2}$	1.4	Cell model (18) prediction
$\phi = 0.41$ $R = 0.25 \text{ mm}$	$1.90 * 10^{-4}$	(2.70)	$5.32 * 10^{-2}$	3.54	Bronze spheres data series 3 <sup>a</sup>
	$2.77 * 10^{-4}$	1.72	$7.99 * 10^{-2}$	1.74	Cell model (19) prediction
	$1.03 * 10^{-4}$	1.72	$5.44 * 10^{-2}$	1.4	Cell model (18) prediction

<sup>a</sup>See Ref. 13.



TABLE II. Comparison of cell model predictions and data for dc permeability, tortuosity, characteristic viscous dimension, and geometrical parameter in packings of glass beads.

Parameters	$k_0$ (mm <sup>2</sup> )	$\alpha_\infty$	$\Lambda$ (mm)	$M$	
$\phi=0.5$ $R=0.85$ mm	$8.0 \times 10^{-3}$	1.42	$4.40 \times 10^{-1}$	1.0	Fused glass bead data series 1 <sup>a</sup>
	$7.45 \times 10^{-3}$	1.50	$3.75 \times 10^{-1}$	1.26	Cell model (19) prediction
	$3.43 \times 10^{-3}$	1.50	$2.83 \times 10^{-1}$	1.02	Cell model (18) prediction
$\phi=0.5$ $R=0.475$ mm	$1.83 \times 10^{-3}$	1.56	$2.18 \times 10^{-1}$	1.02	Fused glass bead data series 2 <sup>a</sup>
	$2.33 \times 10^{-3}$	1.50	$2.10 \times 10^{-1}$	1.26	Cell model (19) prediction
	$1.07 \times 10^{-3}$	1.50	$1.58 \times 10^{-1}$	1.02	Cell model (18) prediction
$\phi=0.5$ $R=0.25$ mm	$0.44 \times 10^{-3}$	1.49	$1.10 \times 10^{-1}$	1.02	Fused glass bead data series 3 <sup>a</sup>
	$0.64 \times 10^{-3}$	1.50	$1.10 \times 10^{-1}$	1.26	Cell model (19) prediction
	$0.30 \times 10^{-3}$	1.50	$0.83 \times 10^{-1}$	1.02	Cell model (18) prediction
$\phi=0.4$ $R=0.73$ mm	$1.51 \times 10^{-3}$	1.37	$0.90 \times 10^{-1}$	5.1	Stacked glass beads data <sup>b</sup>
	$2.15 \times 10^{-3}$	1.75	$2.25 \times 10^{-1}$	1.48	Cell model (19) prediction
	$0.76 \times 10^{-3}$	1.75	$1.51 \times 10^{-1}$	0.8	Cell model (18) prediction

<sup>a</sup>See Ref. 14.

<sup>b</sup>See Ref. 15.

mension obtained with materials having similar values for the radius and volume fraction of particles.

The frequency dependencies of predicted and measured<sup>14</sup> values of the real and the imaginary parts of the scaled dynamic permeability  $k(\omega)/k_0$  in a random packing are presented in Fig. 5.

However, cell model estimates of DC permeability are significantly lower than numerical results.<sup>17,18</sup> For example, for close SC packing, the numerical result<sup>17</sup> is  $k_0/R^2\phi = 2.065 \times 10^{-2}$ , whereas the modified cell model predicts  $k_0/R^2\phi = 1.14 \times 10^{-2}$ .

Calculation of formation factor for close SC packing using (14') and (19) gives  $F=4.068$ . This exceeds both the numerical result<sup>12</sup>  $F=2.906$  and the modified cell model prediction  $F=3.253$ . For data series 2,<sup>14</sup> (14') and (19) give  $\alpha_\infty=1.838$ . Again this value exceeds both the measured value  $\alpha_\infty=1.56$  and the modified cell model result  $\alpha_\infty=1.5$ .

#### IV. CONCLUDING REMARKS

The Kuwabara/Strout cell model has been modified to match the macroscopic and the cell-volume averaged velocities and by increasing the cell volume. Analytical expressions for acoustical/drag parameters have been obtained in terms of three parameters: the volume fraction of the solid phase, the particle radius, and the cell radius. If the cell radius is defined by that of the sphere that circumscribes a unit cell in a cubic lattice, then the values of drag parameters predicted by the new cell model are in good agreement with numerical results<sup>12</sup> for simple cubic packing and measured

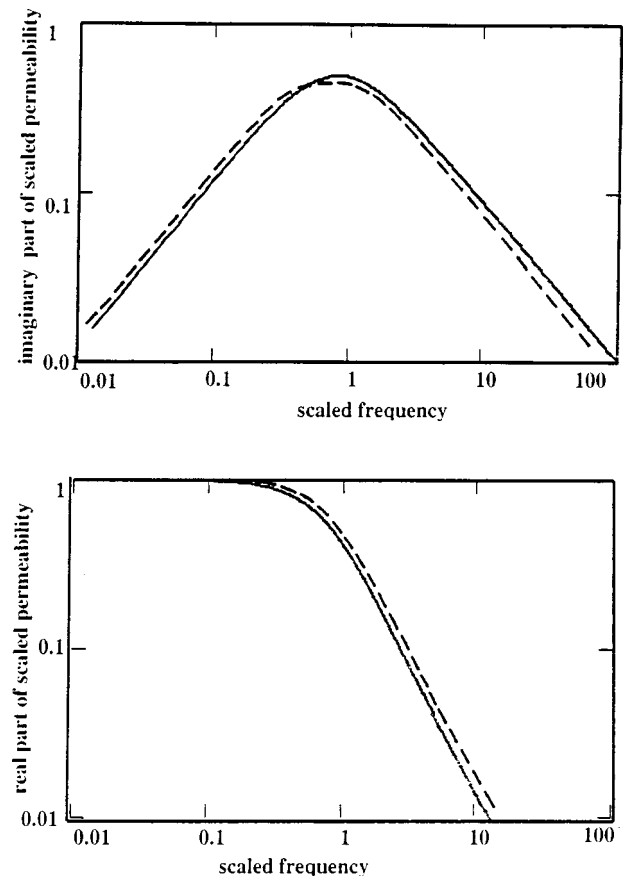


FIG. 5. Real and imaginary parts of the scaled dynamic permeability  $k(\omega)/k_0$  as a function of the scaled frequency  $\omega/\omega_c$  [ $\omega_c = \eta(1-\phi)/\rho_0\alpha_\infty k_0$ ]. The solid line represents cell model predictions and the broken line corresponds to fused glass bead data (Ref. 14) series 1–3.

data for sintered bronze spheres and fused glass beads.<sup>13,14,16</sup> This is encouraging for our ultimate aim of using the cell model approach for deducing the acoustical properties of granular materials. Further publications will be concerned with new similarity relationships between viscous and thermal characteristics and with the calculation of bulk acoustical parameters.

#### ACKNOWLEDGMENT

The work was supported by EPSRC (U.K.) through Grant No. GR/L61804.

<sup>1</sup>J-F. Allard, *Propagation of Sound in Porous Media* (Elsevier Science, Essex, U.K., 1993).

<sup>2</sup>K. Attenborough, "Models for acoustical properties of air-saturated granular media," *Acta Acust.* **1**, 213–226 (1993).

<sup>3</sup>S. Kuwabara, "The forces experienced by randomly distributed parallel circular cylinders or spheres in a viscous flow at small Reynolds numbers," *J. Phys. Soc. Jpn.* **14**, 527–532 (1959).

<sup>4</sup>J. Happel, "Viscosity of suspensions of uniform spheres," *J. Appl. Phys.* **28**, 1288–1292 (1957).

<sup>5</sup>T. A. Strout, "Attenuation of sound in high concentration suspensions: Development and application of an oscillatory cell model," Ph.D. Thesis, University of Maine, 1991.

<sup>6</sup>V. Tarnow, "Calculation of the dynamic air flow resistivity of fiber materials," *J. Acoust. Soc. Am.* **102**, 1680–1688 (1997).

<sup>7</sup>D. L. Johnson, J. Koplik, and R. Dashen, "Theory of dynamic permeabil-

- ity and tortuosity in fluid-saturated porous media," J. Fluid Mech. **176**, 379–402 (1987).
- <sup>8</sup>L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Butterworth–Heinemann, Oxford, 1987).
- <sup>9</sup>J. G. Berryman, "Confirmation of Biot's theory," Appl. Phys. Lett. **37**, 382–384 (1980).
- <sup>10</sup>K. Attenborough, "Acoustical characteristics of rigid porous absorbers and granular materials," J. Acoust. Soc. Am. **73**, 785–799 (1983).
- <sup>11</sup>D. L. Johnson, J. Koplik, and L. M. Schwartz, "New pore-size parameter characterizing transport in porous media," Phys. Rev. Lett. **57**, 2564–2567 (1986).
- <sup>12</sup>A. M. Chapman and J. J. L. Higdon, "Oscillatory Stokes flow in periodic porous media," Phys. Fluids A **4**, 2099–2116 (1992).
- <sup>13</sup>S. R. Baker and I. Rudnick, "Measurement of the Biot structural factor-sigma for sintered bronze spheres," IEEE Trans. Antennas Propag. **33**, 118 (1986); *Proceedings of the 1985 IEEE Ultrasonic Symposium*, San Francisco, 1985 (IEEE, New York, 1986).
- <sup>14</sup>E. Charlaix, A. P. Kushnick, and J. P. Stokes, "Experimental study of dynamic permeability in porous media," Phys. Rev. Lett. **61**, 1595–1598 (1988).
- <sup>15</sup>J. F. Allard, M. Henry, J. Tizianel, L. Kelders, and W. Lauriks, "Sound propagation in air saturated random packings of beads: Surface waves over bead layers," J. Acoust. Soc. Am. **104**, 2004–2007 (1998).
- <sup>16</sup>P.-Z. Wong, J. Koplik, and J. P. Tomanic, "Conductivity and permeability of rocks," Phys. Rev. B **30**, 6606–6614 (1984).
- <sup>17</sup>M.-Y. Zhou and P. Sheng, "First-principle calculations of dynamic permeability in porous media," Phys. Rev. B **39**, 12027–12039 (1989).
- <sup>18</sup>A. S. Sangani and A. Acrivos, "Slow flow through a periodic array of spheres," Int. J. Multiphase Flow **8**, 343–360 (1982).

# Transient Lamb waves: Comparison between theory and experiment

M. Deschamps and O. Poncelet

*Laboratoire de Mécanique Physique, Université Bordeaux I, UMR CNRS 5469, 351, Cours de la Libération, 33405 Talence Cedex, France*

(Received 27 February 1998; accepted for publication 16 February 2000)

This paper presents an experimental investigation of the generation of transient Lamb waves inside an immersed plate by a time limited signal. The transient Lamb waves are the solutions of the characteristic equation of the plate for complex frequency and real slowness, in contrast to the leaky Lamb waves, for which the slowness is complex and the frequency is real. By using a large transducer in agreement with the real slowness assumption, the excitation of these transient Lamb waves is observed experimentally both in reflection and transmission for a tone burst excitation. Particular attention is paid to the  $A_0$  mode for which the amplitude increases exponentially in time. As a result, the total reflected waveform contains in this case a transient contribution, coming from the generation of the  $A_0$  mode, that precedes the forced regime. It is shown that there is a very good agreement between the theoretical and the experimental time exponential amplitude variations of the transient Lamb waves. For the  $A_0$  mode, the transient signal is qualitatively explained by a group velocity higher than the phase velocity. © 2000 Acoustical Society of America.

[S0001-4966(00)04505-7]

PACS numbers: 43.35.Cg [HEB]

## INTRODUCTION

The study of Lamb waves that propagate in an immersed plate is of considerable interest in view of the various applications in nondestructive testing and in quantitative nondestructive evaluation.<sup>1-3</sup>

Most investigations on Lamb waves in immersed plates have been carried out considering real frequency and complex wave number, in order to account for attenuating wave modes during their propagation.<sup>4,5</sup> This decision was often taken in accordance with specific ultrasonic applications in which a space leakage of energy was observed.<sup>4,6</sup> In this context, the free mode properties of fluid-loaded plates have been studied theoretically in terms of plane waves<sup>7,8</sup> as well as of spatial bounded beams.<sup>9-11</sup>

Another way to investigate this problem consists of studying the acoustic plate response to a tone burst. In such a case, the transient part that appears after the steady state is directly related to the excitation of Lamb waves.<sup>12</sup> Based on this observation, an experimental method to isolate the resonances of the plate has been developed and correlated to theoretical works.<sup>13</sup>

A more general study has been performed with respect to the transmission and the reflection of space and time bounded beams on a viscoelastic anisotropic solid layer.<sup>14</sup> The results of this study clearly show that the amplitude of the transmitted or reflected signal depends upon the receiver position, and consequently, the spectrum, from which the transfer function may be derived, must also be regarded as a function of the receiver location. In such a situation, it is difficult to extract a specific Lamb wave signature from the time response. This is mainly due to the fact that the excitation is limited in time as well as in space.

In this regard, a recent paper<sup>15</sup> has placed emphasis on the difference between transient Lamb waves and leaky

Lamb waves by comparing their dispersion curves. More precisely, the different nature of these two extreme possibilities for generating Lamb waves in an immersed plate has been discussed from a physical point of view, leading to the conclusion that no intrinsic relation between both solutions exists. The reason is that these two kinds of Lamb waves are, in essence, different and they model two distinct experimental conditions.

Plate waves with real slowness and complex frequency vary with time, not in space when following a phase front. In this way, we have decided to call them transient Lamb waves in contrast with the transient leaky Lamb waves,<sup>16,17</sup> which are characterized by a complex frequency and a real wave number (in this case the slowness must be complex too). Transient Lamb waves can be generated by time limited excitation using a source which is infinite in the spatial domain. On the other hand, leaky Lamb waves vary in space but are steady in time. They can be generated by a harmonic source which is localized in space. In addition, it has been shown that, although the wave vectors associated with each solution are both complex, the corresponding physical meaning is fundamentally different. Apart from the difference in physical nature, the biggest contrast between the solutions is exhibited by the  $A_0$  mode for which the amplitude of a transient Lamb wave increases exponentially with time.

The purpose of this study is to experimentally analyze the reflected and the transmitted waveforms from an immersed plate by transient Lamb waves. The imaginary part of the frequency, measured directly from the waveform, will be compared with the predicted value obtained from the dispersion curves. Special attention is focused on the  $A_0$  mode.

This paper begins by describing the theory, followed by the experimental setup and the excitation/measurement conditions. Section III deals with the comparison between the

theoretical calculations and the experimental results. Finally, the relative values of the group velocity and the phase velocity are studied in order to explain the peculiar behavior of the  $A_0$  mode.

## I. THEORY

For a plate immersed in a fluid, the dispersion equation of Lamb waves is given by the following transcendental equation in implicit form:<sup>8</sup>

$$(S(\omega, s_x) + i\tau)(A(\omega, s_x) - i\tau) = 0, \quad (1)$$

with

$$A(\omega, s_x) = 4s_x^2 \varphi \psi \tan(\omega \varphi e/2) + (s_T^2 - 2s_x^2)^2 \tan(\omega \psi e/2),$$

$$S(\omega, s_x) = 4s_x^2 \varphi \psi \cotan(\omega \varphi e/2) + (s_T^2 - 2s_x^2)^2 \cotan(\omega \psi e/2).$$

$$\psi = \sqrt{c_L^{-2} - s_x^2}, \quad \varphi = \sqrt{c_T^{-2} - s_x^2},$$

$$\tau = \frac{\rho_f \psi}{\rho_s \sqrt{c_F^{-2} - s_x^2} c_T^4},$$

where  $c_L$  and  $c_T$  are, respectively, the longitudinal and the shear bulk wave velocities in the solid; the quantity  $e$  being the thickness of the plate and  $c_F$  the sound velocity in the fluid. The variables  $\rho_f$  and  $\rho_s$  represent the densities of the fluid and the solid. The quantity  $s_x$  stands for the  $x$ -component of the complex slowness vector  $\mathbf{S}$  of each wave, in agreement with the generalized Snell's law. Throughout this paper, the quantities related to acoustic fields can be complex. Solving the above characteristic equation is complicated and generally requires numerical analysis. In addition to numerical techniques which provide the exact solution of the dispersion curves, other methods give approximate values, such as the Resonance Scattering Theory and the derivative of the reflection coefficient phase.<sup>18</sup> The solutions yield the antisymmetric and symmetric modes for  $A(\omega, s_x) - i\tau = 0$  and  $S(\omega, s_x) + i\tau = 0$ , respectively. These solutions are complex due to the fluid coupling. As a consequence, this equation can be satisfied if either  $s_x$  or  $\omega$  is complex or if, of course, both are simultaneously complex.

The physical meaning associated with these solutions is that the displacement field of Lamb modes is composed either of harmonic inhomogeneous waves (only  $s_x$  is complex), or of transient homogeneous waves (only  $\omega$  is complex), or of transient inhomogeneous waves (both  $s_x$  and  $\omega$  are complex).

One can describe such a general plane wave solution at any point  $\mathbf{M}$  in space and time  $t$  by taking the real part of the following acoustic displacement:<sup>19,20</sup>

$$\mathbf{d} = a \mathbf{P} \exp i(\omega t - \mathbf{K} \cdot \mathbf{M}), \quad (2)$$

where  $\mathbf{K} = \mathbf{K}' - i\mathbf{K}''$ .

In the above expression,  $a$  is the complex amplitude,  $\mathbf{K} = \omega \mathbf{S}$  denotes the complex wave vector that is the so-called wave bivector. This vector is given by

$$\mathbf{K} = (\omega' \mathbf{S}' + \omega'' \mathbf{S}'') - i(\omega' \mathbf{S}'' - \omega'' \mathbf{S}'), \quad (3)$$

where  $\mathbf{S} = \mathbf{S}' - i\mathbf{S}''$ .

The vector  $\mathbf{P}$  represents the normalized polarization vector and  $\omega = \omega'(1 + i\Omega)$  is the complex angular frequency. The dimensionless parameter  $\Omega = \omega''/\omega'$  expresses the imaginary to real frequencies ratio. The propagation vector is denoted  $\mathbf{K}'$  and the attenuation (or damping) vector is represented by  $\mathbf{K}''$ . The real positive scalar  $\omega'$  stands for the angular frequency. The parameter  $\omega''$  is the extinction coefficient (with  $\omega'' > 0$ ) or the "source inflation" coefficient (with  $\omega'' < 0$ ). This coefficient describes the exponential time decay/rise of the wave.

The bivector  $\mathbf{S}$  represents the complex slowness vector. The introduction of this vector instead of the complex wave vector  $\mathbf{K}$  is of primary importance, since the frequency and the slowness vector are two independent quantities when the characteristics of the fluid and the plate are fixed, contrary to  $\omega$  and  $\mathbf{K}$ . Consequently, the solutions for  $s_x$  (when  $\omega$  is a parameter) cannot *a priori* provide the solutions for  $\omega$  (when  $s_x$  is a parameter), and *vice versa*. Thus, there exists an infinite number of dispersion curves for each of the  $A_n$  (or  $S_n$ ) modes. It is immediately apparent that the problem can be solved by two different basic approaches. On the one hand, by setting the frequency, one can search the slowness and, on the other hand, by setting the slowness the solution can be obtained in terms of the frequency.

For convenience, a complex slowness vector  $\mathbf{S}_{\text{Lamb}} = s_x \mathbf{x}$  (where  $\mathbf{x}$  is the unit vector oriented along the plate) can be associated with the Lamb waves. Both real components of this slowness bivector,  $\mathbf{S}'_{\text{Lamb}}$  and  $\mathbf{S}''_{\text{Lamb}}$  are oriented following  $\mathbf{x}$  and their norm is given by, respectively,  $s'_x$  and  $s''_x$ , which can be directly obtained from the solutions of Eq. (1). The associated wave bivector is such that  $\mathbf{K}_{\text{Lamb}} = \omega \mathbf{S}_{\text{Lamb}}$ . It is of interest to inspect the wave bivectors of both the transient and the leaky Lamb waves separately. For the leaky Lamb waves ( $\omega$  is real), the vectors  $\mathbf{S}'_{\text{Lamb}}$  and  $\mathbf{S}''_{\text{Lamb}}$  are oriented following the same direction. The imaginary part of  $\mathbf{K}_{\text{Lamb}}$ ,  $\mathbf{K}''_{\text{Lamb}} = \omega' \mathbf{S}''_{\text{Lamb}}$ , which represents the amplitude decrease direction of the acoustic field at a fixed time, is therefore oriented following the propagation vector  $\mathbf{K}'_{\text{Lamb}} = \omega' \mathbf{S}'_{\text{Lamb}}$ . For transient Lamb waves ( $s''_x = 0$ ), the imaginary part of  $\mathbf{K}_{\text{Lamb}}$ , such that  $\mathbf{K}''_{\text{Lamb}} = -\omega'' \mathbf{S}'_{\text{Lamb}}$ , is not necessarily oriented following the propagation vector  $\mathbf{K}'_{\text{Lamb}}$ . Clearly depending on the sign of the imaginary part of the frequency  $\omega''$  (or  $\Omega$ ) the two vectors can be either of opposite direction or identically oriented.

In practice, it is necessary to know exactly the type of the initial excitation. For this purpose, a virtual incident plane wave is associated with the vector  $\mathbf{S}_{\text{Lamb}}$  and its slowness vector  $\mathbf{S}_i$  is such that  $\mathbf{S}_i \cdot \mathbf{x} = s_x$ . Figure 1 illustrates the two most extreme and restrictive cases for which the frequency and the slowness can not be simultaneously complex. For both cases, the displayed incident plane wave is the unique plane wave in the fluid that can generate the Lamb mode. For this wave, the moduli of the reflection and transmission coefficients tend to infinity.<sup>21</sup> In Figs. 1 and 15 concerning the structure of Lamb waves, this plane wave will be drawn (dashed lines) in order to immediately visualize the

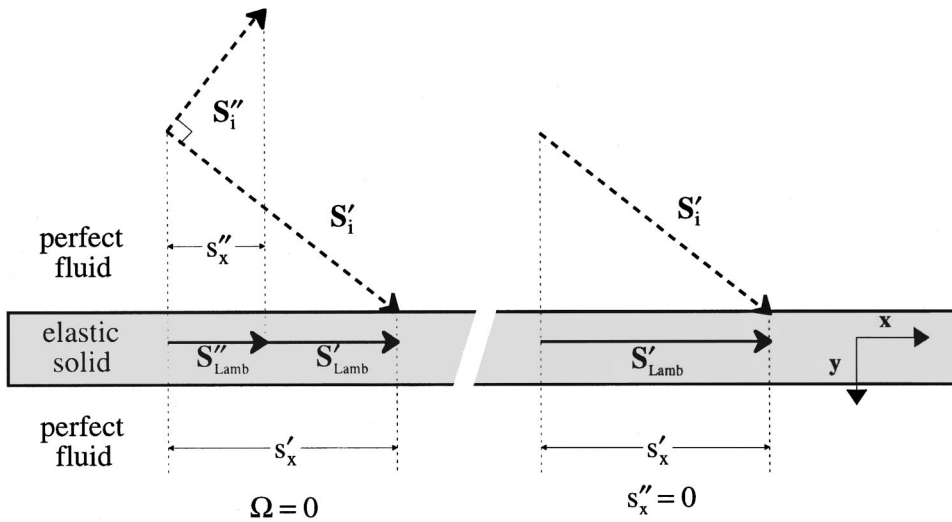


FIG. 1. Schematic representation of Lamb waves via the slowness vector  $S_{Lamb}$ .

nature (temporal or spatial limited beams) of the Lamb waves.

To fix the idea, Fig. 2 shows a comparison between the Lamb waves' dispersion curves for complex slowness (when  $\omega$  is real) and for complex frequency (when  $s_x$  is real). The phase velocities  $v = s_x'^{-1}$  of the full set of Lamb modes are plotted versus the product real frequency thickness, using the properties  $c_F = 1.5 \text{ mm}/\mu\text{s}$ ,  $c_L = 6.37 \text{ mm}/\mu\text{s}$ ,  $c_T = 3.1 \text{ mm}/\mu\text{s}$ ,  $\rho_f = 1$  and  $\rho_s = 2.8$ . As pointed out in Refs. 15 and 16, it is apparent that the Lamb waves' dispersion curves are different for these two calculations. In the solution for complex frequencies, the major feature is the existence of a cutoff frequency beyond which the  $S_0$  and  $A_0$  modes disappear simultaneously. The imaginary parts  $\Omega$  of the characteristic equation solutions [Eq. (1)], when searching complex frequency solutions for real values of  $s_x$ , are plotted in Figs. 3 and 4. Except for the  $A_0$  mode, all the Lamb modes have a positive imaginary part of the complex frequency which corresponds to an exponential decrease of the signal with time. For the case of the  $A_0$  mode, the value of  $\Omega$  is always negative and means that the temporal shape of this mode is an increasing exponential function. This particular behavior will be shown experimentally later.

In many guided ultrasonic waves applications, the real couples  $(v, \omega)$ , such that the reflection coefficients of the

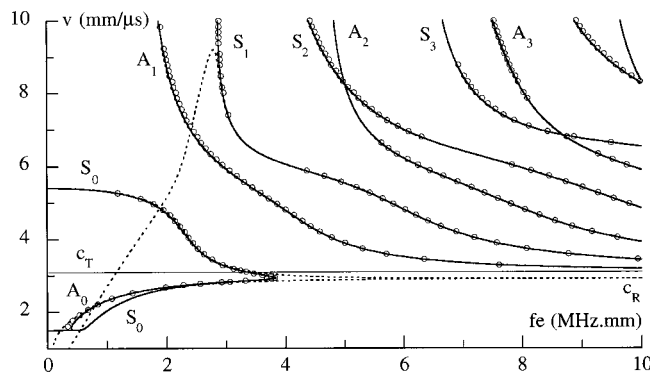


FIG. 2. Comparison between the zeros of the reflection coefficient (O) and the dispersion curves of Lamb waves in a fluid-loaded plate for complex slowness solutions'  $s_x$  (dashed lines) and for complex frequency solutions  $\omega$  (solid lines).

plate (for a harmonic homogeneous incident plane wave) are null, may be identified by the velocity dispersion of plate waves (Cremer's coincidence principle). It can be seen in Refs. 1 and 22 that, assuming real frequencies, the loci of the real reflection zeros and the loci of the real part of the reflection complex poles (related to the leaky Lamb waves) have a similar behavior, except for some regions like: high and low phase velocities and near the Rayleigh wave velocity. However, if we now compare the real zeros with the dispersion curves calculated for complex frequency and real slowness (see Fig. 2), we can see that the reflection coefficient zeros globally better match the complex frequency poles than the complex slowness poles, especially in the three regions mentioned above.

Additionally, depending on the properties of the plate, the differences between the zeros of the reflection coefficient and the dispersion curves can be much more pronounced. However, these zeros agree again very well with the dispersion curves of transient Lamb waves. For instance, Fig. 5 shows the superposition of the zeros with both kinds of Lamb wave dispersion curves, for the isotropic plane of an unidirectional graphite/epoxy anisotropic plate. Since the investigated plane is normal to the fiber direction, the dispersion equation (1) is still valid even though the plate is assumed to be anisotropic. The numerical values are identical to those given in Ref. 23. In view of Fig. 5, it is unquestion-

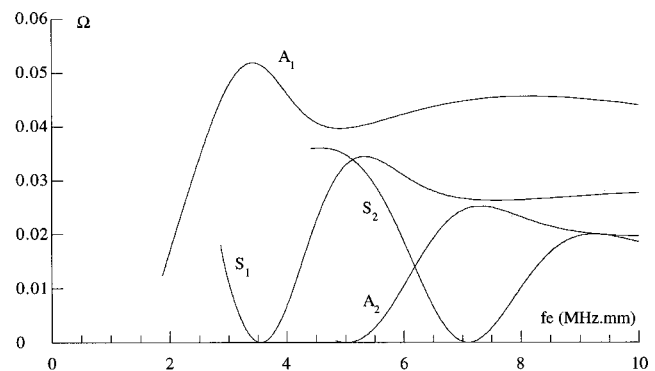


FIG. 3. Imaginary part ( $\Omega$ ) of the transient Lamb waves versus the product  $fe$  ( $s_x'' = 0$ ).

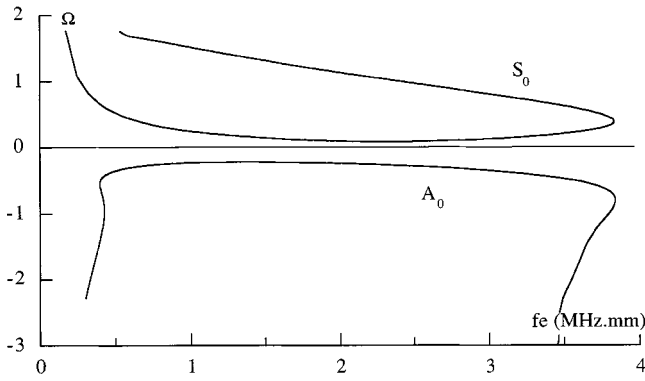


FIG. 4. Imaginary part ( $\Omega$ ) of the transient Lamb waves  $A_0$  and  $S_0$  versus the product  $fe$  ( $s'_x=0$ ).

able that the dispersion curves of the transient Lamb waves best match the real zeros of the reflection coefficient. A complementary inspection of all principal planes may be found in Ref. 24. Similar and much more spectacular results may be observed for an embedded layer.<sup>25</sup> It is not the purpose of this paper to comment on the different matches between the two approaches, but only to show them. As a matter of fact, some detailed discussions about this subject are presented in Refs. 1, 25, and 22.

In conclusion, we may state that the use of complex frequencies appears to be considerably more attractive for some specific applications. This feature will be used in the experiments to determine the excitation conditions of transient Lamb waves. The experimental setup, which will be presented in the next section, is in agreement with the assumption of transient and homogeneous plane waves.

## II. EXPERIMENTAL SETUP

The experimental configuration to study the reflected and the transmitted waveforms is presented in Fig. 6. A goniometer with axis of rotation  $O$  supports an aluminum sample (with the same characteristics as those given in Sec. I) which is immersed in water. The sample rotation  $\theta_i$  gives the incidence angle. A broad band rectangular plane wave

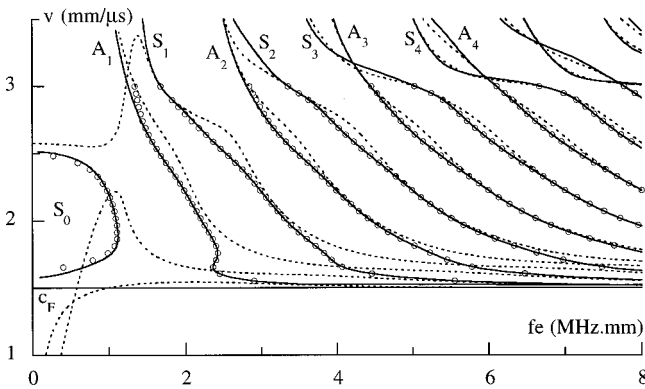


FIG. 5. Comparison between the zeros of the reflection coefficient ( $\circ$ ) and the dispersion curves of the permanent leaky Lamb waves (complex slowness solutions  $s_x$ ; dotted line), as well as the dispersion curves of the transient Lamb waves (complex frequency solutions  $\omega$ ; solid line) for a unidirectional composite (principal plane perpendicular to the fiber). In this plane, the acoustic properties are  $c_L=2.970$  mm/ $\mu$ s,  $c_T=1.489$  mm/ $\mu$ s and  $\rho_s=1.578$ .

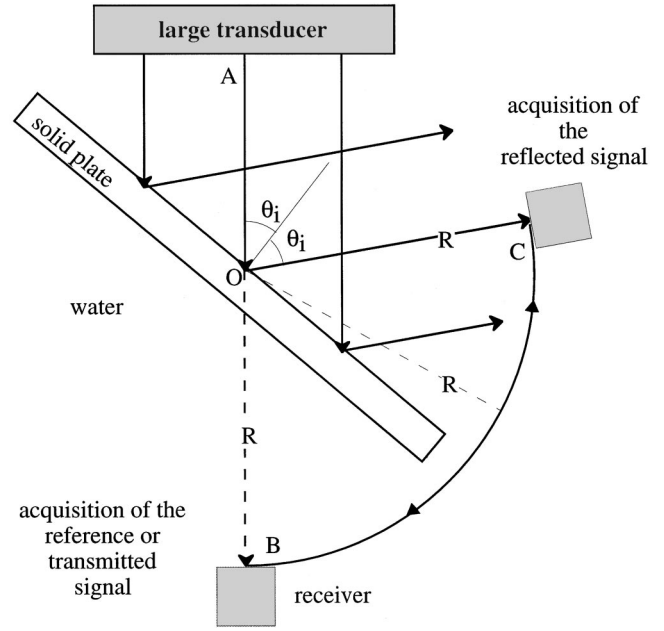


FIG. 6. Reflection and transmission of a plane wave by an immersed plate: experimental setup.

emitter is in a fixed position. Its dimensions are 80 by 40 mm, and its center frequency is about 3 MHz. Measurements of the acoustic field at about 30 cm from the emitter show that the wave front is a plane with a constant amplitude in an area of about 50 mm in width. At 2 MHz in water, this corresponds to about 70 wavelengths. Hence, at the center of the beam the assumption of the plane wave is well justified and the spatial bounded beam effects are considerably limited. The receiver is a typical 3/4" broadband transducer of 2.25 MHz center frequency. It is supported by a second goniometer which is concentric with the first. When investigating beam reflection, the second goniometer is rotated over an angle that is exactly the double of the sample rotation, i.e.,  $2\theta_i$ , in agreement with the specular direction. Obviously, when working in transmission, the rotation of the second goniometer is equal to  $180^\circ$ . In both cases the emitter-to-receiver path length remains constant. The receiver and emitter distances, with respect to the point  $O$ , are respectively  $OB=OC=R=11$  cm and  $AO=28$  cm. It should also be noted that the rotation point  $O$  is located on the sample's first interface. The reason will be explained later.

The emitter is excited by a tone burst signal with a number of periods which depend on the experiment (according to the length of the transient Lamb waves signal). Note that due to interferences within the plate, the waveform is very sensitive to small variations of the excitation frequency. The excitation frequency has to be very accurate to guarantee good measurements. The temporal responses of the receiver are captured with 2048 points at a sampling frequency of 100 MHz. In addition, to increase the signal to noise ratio, 10 signals are averaged.

To compare the theoretical results with those obtained from the experiments, four different situations, denoted  $E_1$ ,  $E_2$ ,  $E_3$  and  $E_4$ , are inspected. The angle of incidence, the thickness of the plate, the number of cycles and the emission frequency are listed in Table I.

TABLE I. The angle of incidence, the thickness of the plate, the burst number and the emission frequency for the four experimental situations  $E_1$ ,  $E_2$ ,  $E_3$ , and  $E_4$ .

Experiment	Thickness	Frequency	Angle of incidence	Number of cycles
$E_1$	1 mm	2.8 MHz	14°	25
$E_2$	1 mm	1.32 MHz	37°	10
$E_3$	1 mm	2 MHz	37°	10
$E_4$	5 mm	1.32 MHz	37°	10

The corresponding frequency-thickness/angle of incidence points are plotted in Fig. 7 in addition to the dispersion curves of the transient Lamb waves. The angles and frequencies for the experiments  $E_1$  and  $E_2$  permit optimal excitation of the transient  $A_1$  and  $A_0$  modes, respectively. As illustrated above in Figs. 2 and 5, the dispersion curves for complex frequencies closely follow the curves of the reflection coefficient zeros. Using this fact, to approach the best conditions of transient Lamb wave generation, the excitation frequency was tuned from their nominal values until a null reflection of the harmonic part of the incident signal is obtained. For  $E_1$ , the angle of incidence is less than the shear critical angle in the plate, but large enough so that the receiver does not disturb the incident beam (if the angle of incidence is too small then the receiver is too close to the emitter and it disturbs the incident field). The excitation frequency is close to the central frequency of the emitter. For  $E_2$ , the  $A_0$  frequency is below the center frequencies of both transducers but they are sufficiently broad band that the experiments could be done accurately. The other two experiments,  $E_3$  and  $E_4$ , were performed at the fixed angle of incidence  $\theta_i=37^\circ$  and for two different frequency-thickness values. The  $E_3$  experiment was chosen to inspect the waveforms when the experimental situations are close to (but not matching) the optimal transient Lamb wave generation.  $E_4$  was chosen to examine the case of no transient Lamb wave generation.

The predicted poles of the reflection coefficient [solution  $\omega$  of Eq. (1)] and the experimentally applied frequencies corresponding to the null reflection, associated with the experiments  $E_1$  and  $E_2$ , are given in Table II.

The dimensionless factor  $\Omega=f''/f'$ , which is equivalent

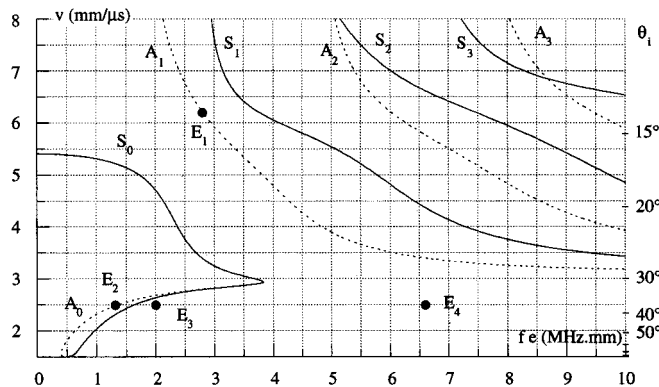


FIG. 7. Dispersion curves of transient Lamb waves for an aluminum plate immersed in water: antisymmetric modes (dashed lines); symmetric modes (solid lines); ● experimental point (frequency-thickness, angle of incidence). Vertical left-hand axis: phase velocity of Lamb wave. Vertical right-hand axis: angle of incidence.

TABLE II. Real ( $f'$ ) and imaginary ( $\Omega=f''/f'$ ) parts of the complex frequency solutions of Eq. (1); measured frequency of null reflection ( $f$ ) for the two experiments  $E_1$  and  $E_2$ .

Experiments	$f'$	$\Omega$	$f$	$\theta_i$
$E_1$	2.805 MHz	+0.043	2.812 MHz	14°
$E_2$	1.345 MHz	-0.22	1.317 MHz	37°

to the imaginary part of the solutions, indicates the wave attenuation (or amplification) as time increases since it is directly connected with this attenuation over a period of time. A comparison between different values of this parameter is then possible. For instance, the amplitude of the transient Lamb wave excited at 37° has a larger exponential variation than the Lamb wave generated at 14°.

### III. EXPERIMENTAL RESULTS

This section is concerned with the reflected and transmitted waveforms from an immersed layer for the conditions presented in Table I. Different reflected and transmitted waveforms of such situations are presented in Figs. 8–14. Note that the time origin does not correspond to the real time of wave excitation. All the waveforms within each figure have been time shifted by an identical delay time. No comparison of absolute times can be made between the different figures.

The classical time response of a tone burst in the case of Lamb wave generation, in experiment  $E_1$ , is shown in Figs. 8 and 9 for reflection and transmission, respectively. It has been recently shown that the envelope of the incident waveform decreases exponentially with time in accordance with  $\Omega>0$ , except for the  $A_0$  mode for which  $\Omega<0$ <sup>15</sup> (see Figs. 3 and 4). For a harmonic incident signal in a rectangular window, the complex harmonic Lamb waves are generated at the abrupt variations of the amplitude located at the beginning and at the end of the excitation. This is clearly visible in Fig. 8, where, in addition to the reflected waveform of the generated  $A_1$  mode (solid line), the reference signal is plotted (dashed line). The latter signal is obtained when the sample is withdrawn and when the transducers are face to face. As already mentioned, the distances  $AB$  and  $AO+OC$  are identical. Consequently, the time of flight corresponding to the

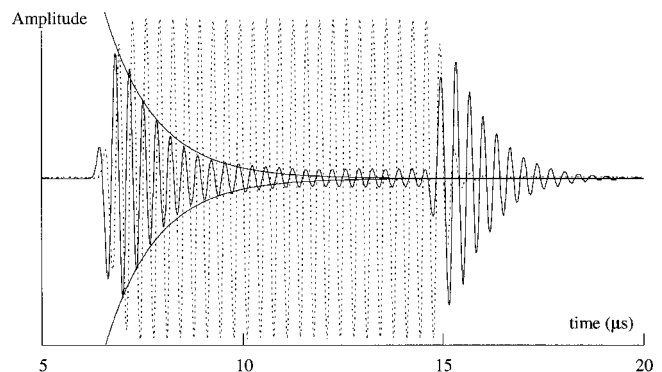


FIG. 8. Experiment  $E_1$  comparison between the reference waveform (dashed line) at the frequency  $f=2.812$  MHz and the experimental waveform reflected from a 1 mm thick aluminum plate at  $\theta_i=14^\circ$  (solid line). The decay envelope is that predicted for the transient  $A_1$  mode.

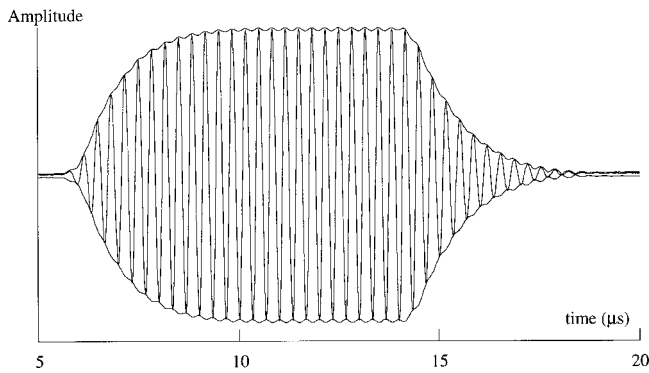


FIG. 9. Experiment  $E_1$  experimental waveform transmitted by a 1 mm thick aluminum plate at  $\theta_i = 14^\circ$  at the frequency  $f = 2.812$  MHz and its envelope.

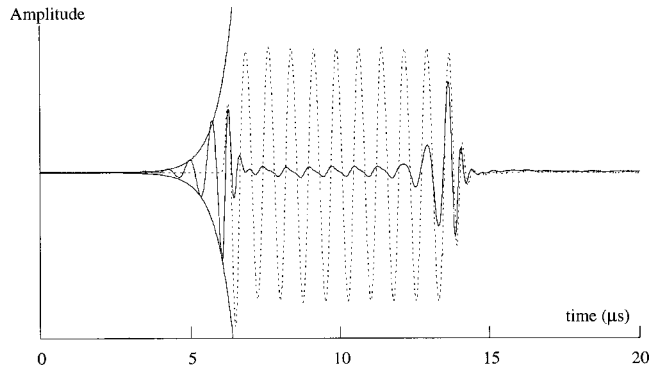


FIG. 10. Experiment  $E_2$  comparison between the reference waveform (dashed line) at the frequency  $f = 1.317$  MHz and the experimental waveform reflected from a 1 mm thick aluminum plate at  $\theta_i = 37^\circ$  (solid line). The exponential envelope is that predicted for the transient  $A_0$  mode.

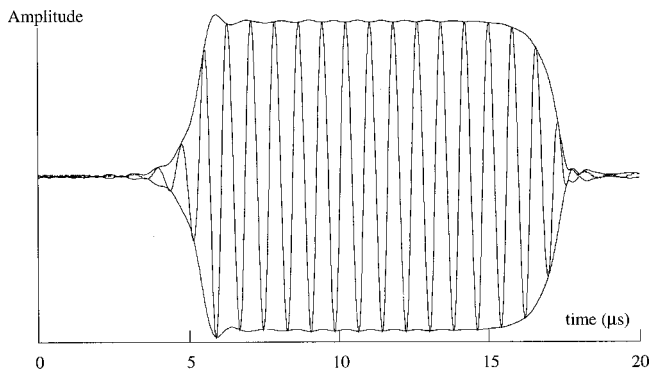


FIG. 11. Experiment  $E_2$  experimental waveform transmitted by a 1 mm thick aluminum plate at  $\theta_i = 37^\circ$  at the frequency  $f = 1.317$  MHz and its envelope.

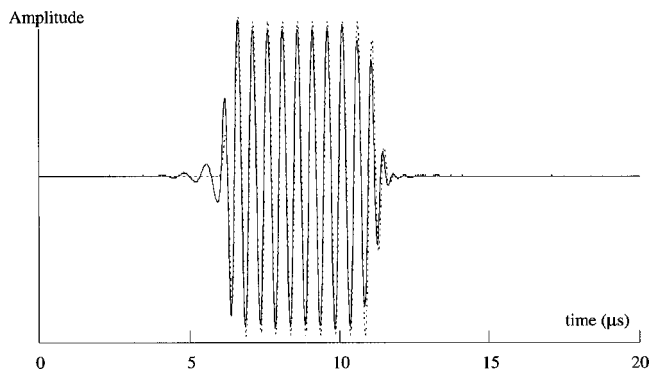


FIG. 12. Experiment  $E_3$  comparison between the reference waveform (dashed line) at the frequency  $f = 2$  MHz and the experimental waveform reflected from a 1 mm thick aluminum plate at  $\theta_i = 37^\circ$  (solid line).

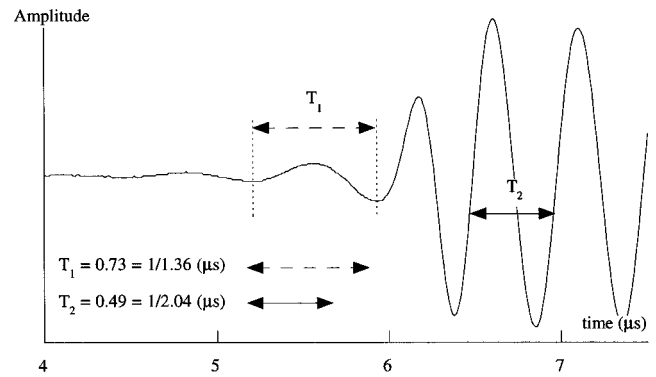


FIG. 13. Zoom of the beginning of the reflected signal presented in Fig. 12.

propagation in the fluid of both the reference and the reflected signals must be identical as well. Comparing the relative arrival time of both wave fronts, it is clear that the excitation of the transient  $A_1$  mode occurs exactly at the two transient parts of the tone burst. It should be noted that the steady state response tends to zero and maximum, respectively, in reflection and in transmission. This agrees with the fact that the steady state reflection coefficient is equal to zero. Finally, a very good agreement is observed between the theoretical time exponential decay and the experiment. The envelope of the amplitude of the Lamb wave directly depends on the imaginary part of the complex frequency solution of Eq. (1). These predicted positive and negative exponential functions (shown by a solid line in Fig. 8, for example) are of the form  $\pm \exp[-\omega''(t-t_0)]$ , for  $t > t_0$ , where  $t_0$  is estimated by matching the temporal position of the exponential functions to the experimental waveform (the rate of the exponential decay is, of course, not changed by this matching, only its position in time). In transmission, the end of the signal exhibits the same exponential behavior. However, the beginning is totally different. In order to make the interpretation easier, the experimental envelope is shown in Fig. 9. In fact, the envelope of the transient part at the onset of the tone-burst excitation is of the form:  $\pm(1 - \exp[-\omega''(t-t_0)])$ , for  $t > t_0$ , where  $t_0$  would be the beginning of the signal. The steady state is then reached when the complex pole associated with the Lamb wave generation no longer contributes to the waveform.

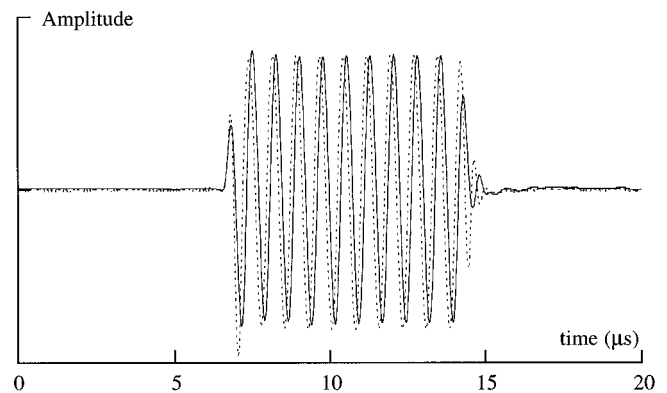


FIG. 14. Experiment  $E_4$  comparison between the reference waveform (dashed line) at the frequency  $f = 1.32$  MHz and the experimental waveform reflected from a 5 mm thick aluminum plate at  $\theta_i = 37^\circ$  (solid line).



Although the difference between leaky Lamb waves (real frequency and complex slowness) and transient Lamb waves (complex frequency and real slowness) has never been clearly pointed out, the observation of such temporal behavior of most transient Lamb waves was widely commented on.<sup>26,27</sup> However, some important and new observations can be drawn from the inspection of the results from experiment  $E_2$ . The excitation of the transient  $A_0$  mode in reflection and transmission is illustrated in Figs. 10 and 11. For this mode the predicted  $\Omega$  is negative and the experimental time response is totally different. The exponential envelopes during the times of the transient parts are now exponential amplifications. Furthermore they finish (they do not start contrary to all the other modes) at the beginning and at the end of the excitation. The example presented in Fig. 10, for  $\theta_i = 37^\circ$  at the frequency  $f = 1.317$  MHz, clearly illustrates this feature. The reflected signal temporally precedes the reference signal. At first sight, it seems that the plate response anticipates the excitation, but this behavior does not result in violation of either energy conservation or causality since the acoustic reflected field is analyzed only in the specular direction. This phenomenon can be explained by the fact that, for the  $A_0$  mode, the group velocity is always greater than the phase velocity. A more detailed discussion will be given in the next section. The good agreement between the predicted exponential increase, by means of the imaginary part of the frequency ( $\Omega = -0,22$ ), and the measurements clearly indicates that this peculiar reflection is the correct solution of this reflection/refraction problem.

Regarding the transmitted wave front, in Fig. 11, it should be noted that, contrary to the  $A_n$  modes ( $n \neq 0$ ), the part of the waveform for which the envelope is of the form  $\pm(1 - \exp[-\omega''(t - t_0)])$  with  $\omega'' < 0$  is now located at the end of the signal (in which  $t_0$  represents the end of the signal). The change in the waveform pattern is consistent with the negative value of  $\omega''$ . Another consequence is that the steady state is reached immediately at the time corresponding to the beginning of the plate excitation, i.e., just after the first contribution of the pole at the beginning of the waveform. Note that, for a correct analysis in this particular case, it is of prime importance to simultaneously visualize the signal and its envelope. Without the envelope pattern, the end of the waveform may be mistakenly interpreted as a decreasing exponential function. It is also interesting to observe that the global shape of the  $A_0$  mode can be obtained by simply reversing, with respect to time, the global shape of the  $A_n$  modes. Apart from the amplitude value of the steady state, which is directly in relation to the reflection or the transmission coefficients, the shape of the waveform is mainly governed by the imaginary part of the complex pole.

Now let us examine the results of experiment  $E_3$ , in which the experimental parameters are close, but not equal, to those of the optimal excitation parameters of the  $A_0$  mode. To this end, only the frequency is shifted from 1.32 to 2 MHz in comparison with experiment  $E_2$ . The reflection response is shown in Fig. 12. Obviously, in this case the reflection coefficient is nonzero. Its modulus is, in fact, theoretically equal to 0.95. Nevertheless, observation of Fig. 12 reveals that, in spite of this near-total reflection, the pole

corresponding to the  $A_0$  mode contributes to the first transient part of the signal, as can be seen at the beginning of the waveform. A close comparison between the frequency of the steady state and the frequency of the transient part emphasizes this point. Figure 13 shows a temporal zoom of the beginning of the waveform presented in Fig. 12. From 5 to 6  $\mu\text{s}$  the frequency of the signal is about 1.32 MHz and equal to the value of the real part of the complex pole for  $\theta_i = 37^\circ$ . After 6  $\mu\text{s}$ , the frequency of the signal corresponding to the experiment  $E_3$  becomes 2 MHz. This means that the  $A_0$  transient mode can be generated for a frequency associated with a near-total reflection. The pole that influences the reflection/refraction is not *a priori* the closest point on the dispersion curve but it is the point at the same angle of incidence. This follows from the fact that the emitter is quasi-planar, as mentioned in Sec. II. From an experimental point of view, the angle of incidence can therefore be assumed to be unique and equal to  $37^\circ$ .

Of course, when the experimental point (frequency-thickness, angle of incidence) is really far from a dispersion curve, no pole contributes to the waveforms. A final example of particular interest in this regard is experiment  $E_4$ , for which no Lamb wave is generated, as shown in Fig. 14. Such circumstances are easily obtained starting from the conditions of the experiment  $E_2$  simply by changing the thickness of the plate. There is no influence of any Lamb wave in the reflected field. The plate can be considered as a pure reflector, in the sense that the envelope of the reflected wave is obtained from the shape of the incident wave by a simple constant multiplication. Moreover, no time shift is observed, since the path lengths in water are the same for both the reflected and the reference fields.

#### IV. RELATION BETWEEN INCIDENT WAVE FRONT AND THE WAVE FRONT OF LAMB WAVES

This section outlines the relationship between the sign of the imaginary part of the frequency, the vector  $\mathbf{K}_{\text{Lamb}}''$  and the group velocity of transient Lamb waves.

Let us first consider the vector  $\mathbf{K}_{\text{Lamb}}''$ . As discussed in Sec. I, depending on the sign of  $\Omega$ , this vector can either be oriented following  $\mathbf{K}_{\text{Lamb}}'$  or in the opposite direction (Fig. 15). From a physical point of view, this corresponds neither to a leakage nor to an amplification during the propagation along the plate. Let us follow the Lamb phase wave front during its propagation to justify this statement. Without loss of generality, we examine, for example, the plane of null phase. As time  $t$  increases, this particular plane moves along the plate on the position  $x$ , such that

$$x = \frac{\omega'}{K_{\text{Lamb}}'} t. \quad (4)$$

The amplitude term, at any time and position, is expressed by

$$\exp[-(\omega''t + K_{\text{Lamb}}''x)]. \quad (5)$$

The amplitude of the phase wave front during its propagation is then given by

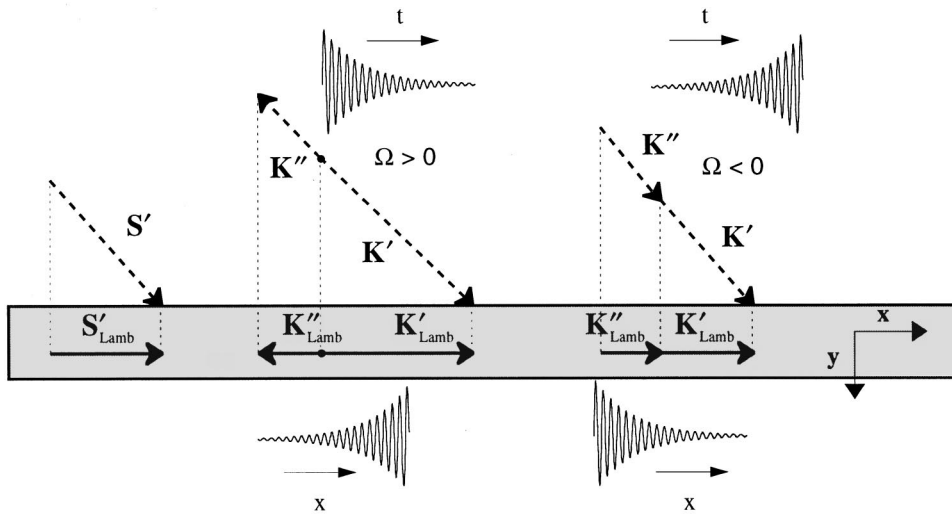


FIG. 15. Structures of transient Lamb waves. Left: slowness vector. Center: wave vector associated with a positive value of the imaginary part of the frequency. Right: wave vector associated with a negative value of the imaginary part of the frequency. In addition to each wave vector, the temporal signal as well as the spatial variations of the displacement field are shown. The corresponding virtual incident plane wave is drawn in a dashed line.

$$\exp\left[-\left(\omega'' + \frac{K''_{\text{Lamb}}}{K'_{\text{Lamb}}}\omega'\right)t\right]. \quad (6)$$

This term is, of course, by essence independent of the position. Using Eq. (6) and taking into account the expressions of the wave vector  $\mathbf{K}'_{\text{Lamb}} = \omega' \mathbf{S}'_{\text{Lamb}}$  and of the vector  $\mathbf{K}''_{\text{Lamb}} = -\omega'' \mathbf{S}'_{\text{Lamb}}$  for a transient plane wave, it can be readily seen that

$$\omega'' + \frac{K''_{\text{Lamb}}}{K'_{\text{Lamb}}}\omega' = 0. \quad (7)$$

This relation holds independent of the sign of the imaginary part of the frequency. The physical meaning is that the wave amplitude of a transient plane wave does not vary during its propagation along the plate. In other words, the vector  $\mathbf{K}''_{\text{Lamb}}$  cannot be attributed to a loss of energy of the wave. In contrast, for the leaky Lamb waves, i.e.,  $\omega'' = 0$ , the amplitude term is given by

$$\exp\left(-\frac{S''_{\text{Lamb}}}{S'_{\text{Lamb}}}\omega' t\right). \quad (8)$$

Clearly, in this case, since  $\mathbf{S}''_{\text{Lamb}}$  and  $\mathbf{S}'_{\text{Lamb}}$  have the same direction, this non-null amplitude variation corresponds to a loss of energy, which denotes the wave leakage due to the fluid loading. This is the basic difference between these two types of waves. The transient Lamb waves propagate in the  $x$  direction with a constant amplitude, although  $\mathbf{K}''_{\text{Lamb}} \neq \mathbf{0}$ , while the leaky Lamb waves lose energy during the propagation ( $\mathbf{S}''_{\text{Lamb}} \neq \mathbf{0}$ ). Strangely, even in some references<sup>16,28</sup> that deal with eigenmodes of multilayered structures in terms of complex frequencies, this physical aspect of such waves is not treated, mainly because the waves with complex frequency are equally assumed to be leaky waves.

Let us focus our attention now on the temporal exponential amplification of the  $A_0$  mode which contrasts the usual exponential decay of all the other modes. For angles of incidence lower than the shear critical angle in the plate, the transient exponential decay can be physically interpreted as a superposition of multiple reflections within the plate.<sup>29</sup> Nevertheless, such an argument does not apply when explaining

the temporal amplification of the  $A_0$  mode and cannot help in understanding why the reflected wave front precedes the reference wave front. For a better explanation of these phenomena let us compare the relative values of group and phase velocities of Lamb waves.

From a numerical point of view, we restrict the comparison to the case of a plate in vacuum, for which the calculation can be easily done because all quantities are real. For the experiments presented, the phase and the group velocities are not much affected by the presence of the fluid. The group velocity  $c_g$  is given by

$$c_g = \frac{\partial \omega}{\partial k} = \left(\frac{1}{c_\varphi} - \frac{\omega}{c_\varphi^2} \frac{\partial c_\varphi}{\partial \omega}\right)^{-1}, \quad (9)$$

where  $k = \omega/c_\varphi$ ,  $c_\varphi$  being the phase velocity of the Lamb wave.

The  $A_0$  mode stands out not only by the sign of its imaginary part of the frequency. As shown in Fig. 16, this mode is also the unique mode for which the group velocity is always larger than the phase velocity. Using this observation, an interpretation of the peculiar transient phenomena of the  $A_0$  mode can be found based on the inspection of geometrical ray trajectories.

The reflection of an incident wave front is schematically presented in Fig. 17. At two different times  $t = t_0$  and  $t = t_1$ , such that  $t_1 > t_0$ , the incident wave fronts (solid line) as

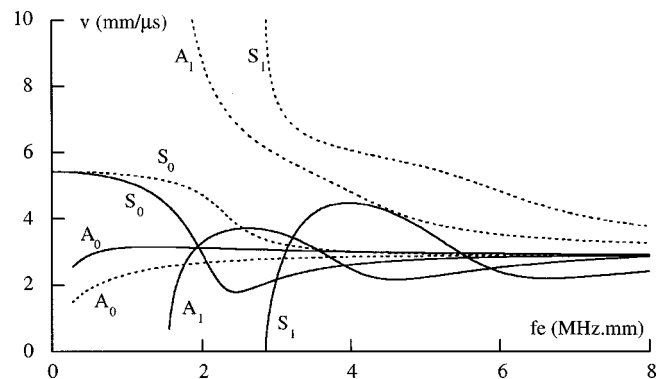


FIG. 16. Comparison between the group velocity (solid line) and phase velocity (dashed line) of Lamb waves for an aluminum plate in vacuum.

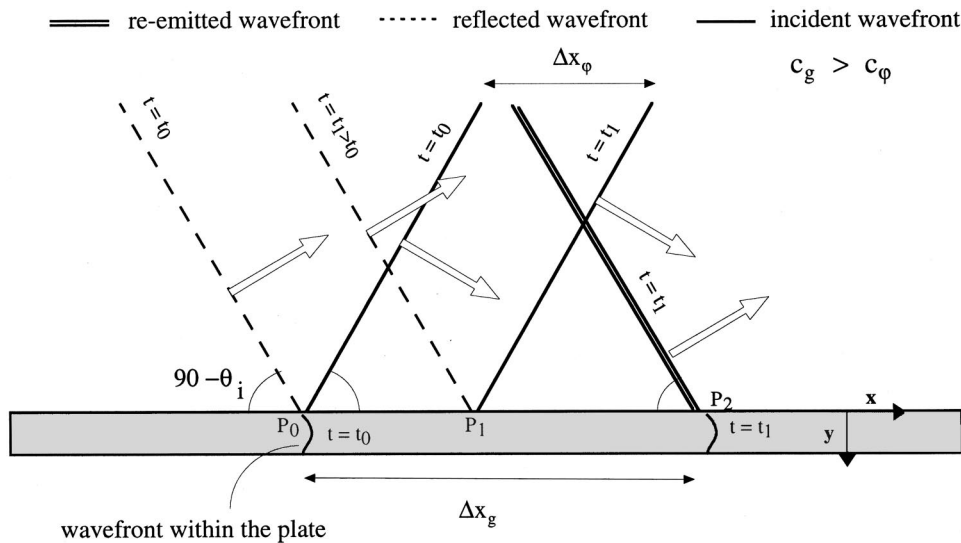


FIG. 17. The incident, reflected and reemitted wave fronts in the fluid at times  $t=t_0$  and  $t=t_1$ .

well as the two fronts associated with the specular reflection (dashed line) are drawn. The group velocity of Lamb waves is also the front speed within the plate, while the corresponding phase velocity represents the displacement speed of the incident (or reflected) front on the interface. As a result, during the time interval  $\Delta t = t_1 - t_0$ , the incident front moved a distance  $\Delta x_\phi = c_\phi \Delta t$  in the  $x$  direction. In the meantime, the Lamb wave front generated at  $t = t_0$  moved a distance  $\Delta x_g = c_g \Delta t$  in the same direction.

As for the plate in vacuum, assuming that the group velocity is greater than the phase velocity, it is evident that the Lamb wave front is located at a position  $P_2$  at the time  $t = t_1$ , which is different from the position  $P_1$  of the incident wave front. During its propagation, the Lamb wave creates a wave in the fluid. Its associated wave front is identified, at time  $t = t_1$ , by the double lines in Fig. 17. It is important to note that this spatial reemission must be considered as the consequence of a transient phenomenon. In fact, within the framework of the plane wave assumption, all wave fronts are spread out from the interface to infinity, and when the steady state is reached the acoustic response is identical at any location of the receiver.

To understand this more clearly, let us assume that the beginning of the excitation coincides with the time  $t = t_0$  at the point  $P_0$ , and that the receiver is somewhere on the wave front reemitted at the point  $P_2$ . It is not surprising to observe that the received signal originates at  $t = t_1$ , and that it arrives ahead of the reflected front. Over this time, starting from zero, the amplitude increases until the steady state is established at this measurement position.

Similar comments may be made concerning all the other modes, except that, in this case, due to the fact that the group velocity is less than the phase velocity, the establishment of the steady state occurs only after the passage of the reemitted wave front. Moreover, the acoustic response is nonzero during the extinction of the reflected wave front while this reemitted wave front exists. Note that, as should be expected for this analysis, when the phase and the group velocities are equal, the imaginary part of the frequency is zero.

This study clearly indicates that there exists a close relation between the imaginary part of the solution of Eq. (1)

and the relative values of phase and group velocities. By using the resonance scattering theory, an approximative expression can be found linking the imaginary part of the solution and the difference between the phase and the group velocities.<sup>30</sup> It should be noted that, in the present study, the explanations given are purely schematic. In other words, the considerations depict a plausible scenario which is in qualitative agreement with observation. However, specific theoretical calculations are required to correctly deal with this problem. This is beyond the scope of the present work.

## V. CONCLUSION

The predictions concerning the transient Lamb waves that have been mentioned in previous reports have been successfully compared with experimental observations. Mainly, the peculiar behavior of the  $A_0$  mode, for which the amplitude increases exponentially as time increases, has been observed experimentally. This apparent violation of causality can be explained by a simple consideration of the relative values of group and phase velocities of Lamb waves. This last point requires more attention, and an additional theoretical study is now being conducted to express the relationship between the velocity variation and the imaginary part of the complex frequency pole associated with the Lamb wave.

<sup>1</sup>D. E. Chimenti and S. I. Rokhlin, "Relationship between leaky Lamb modes and reflection coefficient zeros for a fluid-coupled elastic layer," *J. Acoust. Soc. Am.* **88**, 1603–1611 (1990).

<sup>2</sup>J. Kaplunov, J. Nolde, and N. Veksler, "Determination of parameter of elastic layer by measured dispersion curves of zero-order Lamb-type waves," *Proc. Est. Acad. Sci., Phys., Math.* **41**, 39–48 (1992).

<sup>3</sup>M. R. Karim, A. K. Mal, and Y. Bar-Cohen, "Inversion of leaky Lamb wave data by simplex algorithm," *J. Acoust. Soc. Am.* **88**(1), 482–491 (1990).

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

<sup>5</sup>L. G. Merkulov, "Damping of normal modes in a plate immersed in a liquid," *Sov. Phys. Acoust.* **10**(2), 169–173 (1964).

<sup>6</sup>D. E. Chimenti, "Guided waves in plates and their use in materials characterization," *Appl. Mech. Rev.* **50**(5), 247–284 (1997).

<sup>7</sup>M. F. M. Osborne and S. D. Hart, "Transmission, reflection and guiding of an exponential pulse by a steel plate in water. I. Theory," *J. Acoust. Soc. Am.* **17**, 1–18 (1945).

<sup>8</sup>L. E. Pitts, T. J. Plona, and W. G. Mayer, "Theoretical similarities of

- Rayleigh and Lamb modes of vibration," J. Acoust. Soc. Am. **60**, 374–377 (1976).
- <sup>9</sup>T. J. Plona, L. E. Pitts, and W. G. Mayer, "Ultrasonic bounded beam reflection and transmission effects at a liquid/solid-plate/liquid interface," J. Acoust. Soc. Am. **59**, 1324–1328 (1976).
- <sup>10</sup>M. Rousseau and Ph. Gagnon, "Asymptotic analysis of non-specular effects for the reflection and transmission of a Gaussian acoustic beam incident on a solid plate," J. Acoust. Soc. Am. **80**, 325–332 (1986).
- <sup>11</sup>K. Van Den Abeele and O. Leroy, "Complex harmonic wave scattering as the framework for investigation of bounded beam reflection and transmission at plane interfaces and its importance in the study of vibrational modes," J. Acoust. Soc. Am. **93**, 308–323 (1993).
- <sup>12</sup>J.-L. Izbicki, G. Maze, and J. Ripoche, "Étude de la réémission libre d'une plaque résonnante," *Acustica* **55**, 27–29 (1984).
- <sup>13</sup>N. Gespa, *La Diffusion Acoustique* (Cedocar, Paris, 1987).
- <sup>14</sup>M. Deschamps and P. Chevée, "Reflection and transmission of a transient bounded beam by a dispersive and orthotropic plate," *Acta Acust. (China)* **1**, 3–14 (1993).
- <sup>15</sup>O. Poncelet and M. Deschamps, "Lamb waves generated by complex harmonic inhomogeneous plane waves," J. Acoust. Soc. Am. **102**(1), 292–300 (1997).
- <sup>16</sup>L. E. Alsop, "The leaky-mode period equation. A plane-wave approach," *Bull. Seismol. Soc. Am.* **60**, 1989–1998 (1970).
- <sup>17</sup>J. H. Rosenbaum, "The long-time response of a layered elastic medium to explosive sound," *J. Geophys. Res.* **65**, 1577–1613 (1960).
- <sup>18</sup>O. Lenoir, J. Duclos, J. M. Conoir, and J. L. Izbicki, "Study of Lamb waves based upon the frequency and angular derivatives of the phase of the reflection coefficient," J. Acoust. Soc. Am. **94**, 330–343 (1993).
- <sup>19</sup>M. Hayes, "Inhomogeneous plane waves," *Arch. Ration. Mech. Anal.* **85**, 41–79 (1984).
- <sup>20</sup>M. Deschamps, "Reflection and refraction of the inhomogeneous plane wave," in *Acoustic Interaction with Submerged Elastic Structures*, edited by D. Inman and A. Guran (World Scientific, Singapore, 1996), Series B, Vol. 5, Part I, pp. 164–206.
- <sup>21</sup>K. V. D. Abeele and O. Leroy, "On the influence of frequency and width of an ultrasonic bounded beam in the investigation of materials: Study in terms of heterogeneous plane waves," J. Acoust. Soc. Am. **93**(5), 2688–2699 (1993).
- <sup>22</sup>A. H. Nayfeh and D. E. Chimenti, "Propagation of guided waves in fluid-coupled plates of fiber-reinforced composite," J. Acoust. Soc. Am. **83**(5), 1736–1743 (1998).
- <sup>23</sup>M. R. Karim, A. K. Mal, and Y. Bar-Cohen, "Inversion of leaky Lamb wave data by simplex algorithm," J. Acoust. Soc. Am. **88**(1), 482–491 (1990).
- <sup>24</sup>M. Deschamps and O. Poncelet, "Plate waves with complex frequency and total transmission," *Proceeding of the World Congress on Ultrasonics, Yokohama, Japan, August 24–27 1997*, pp. 190–191.
- <sup>25</sup>A. Bernard, M. Lowe, and M. Deschamps, "Calculation of dispersion curves in bonded joints using either complex frequency or complex slowness and comparison of the results with the minima of the plane wave reflection coefficients," *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1998), Vol. 17, pp. 185–192.
- <sup>26</sup>G. Maze, J.-L. Izbicki, and J. Ripoche, "Méthode d'isolement et d'identification des résonances: Application aux cylindres et tubes," in *La Diffusion Acoustique*, edited by N. Gespa (Cedocar, Paris, 1987).
- <sup>27</sup>G. Maze, J.-L. Izbicki, J. Ripoche, A. Nagl, H. Überall, and K. B. Yoo, "Transient acoustic scattering from layers and plates," J. Acoust. Soc. Am. **80**(1), 295–301 (1986).
- <sup>28</sup>F. Gilbert, "Propagation of transient leaking modes in a stratified elastic waveguide," *Rev. Geophys.* **2**, 123–153 (1964).
- <sup>29</sup>S. K. Numrich and H. Überall, "Scattering of sound pulses and the ringing of target resonances," in *Physical Acoustics XXI*, edited by A. D. Pierce and R. N. Thurston (Academic, New York, 1992), pp. 235–318 (1992).
- <sup>30</sup>J.-L. Izbicki and J.-M. Conoir, "Gradient de phase," *Proceedings of the 2<sup>e</sup> École d'hiver sur la diffusion acoustique, Jan. 30–Feb. 5, 1994, Caudebec, France.*

# Nonlinear saturation of the thermoacoustic instability

S. Karpov and A. Prosperetti<sup>a)</sup>

Department of Mechanical Engineering, The Johns Hopkins University, Baltimore, Maryland 21218

(Received 15 February 1999; accepted for publication 10 February 2000)

A weakly nonlinear theory of the thermoacoustic instability in gas-filled tubes is developed in the time domain by exploiting the difference between the instability time scale and the period of standing waves. By carrying the expansion to fourth order in the perturbation parameter, explicit results for the initial growth, nonlinear evolution, and final saturation are obtained. The dependence of the saturation amplitude upon the temperature difference in the stack, the tube geometry, stack plate spacing, Prandtl number, and other parameters is illustrated. © 2000 Acoustical Society of America. [S0001-4966(00)03605-5]

PACS numbers: 43.35.Ud [HEB]

## INTRODUCTION

The linear theory of thermoacoustic effects, as developed in the well-known series of papers by Rott (1969, 1976, 1980) is by now well established (good reviews are provided by Rott, 1983; Wheatley, 1986; and Swift, 1988). The extension of this work to the nonlinear regime is, however, a difficult task. Cao *et al.* (1996) have presented numerical solutions of the complete set of conservation equations for a simplified thermoacoustic couple at steady state. In order to refine these calculations, Worlikar *et al.* (Worlikar and Knio, 1996, 1998; Worlikar *et al.*, 1998) had to devise a new formulation leading to a manageable numerical task in spite of the scale complexity inherent to the problem.

In an earlier paper (Yuan *et al.*, 1997), we have presented numerical results based on an approximate quasi-one-dimensional extension of the linear theory. These results are interesting, but the calculations are complex and time consuming and it is difficult to carry out extensive parametric studies on their basis. For this purpose, analytic or semianalytic approximations would be more useful but, in this field, very little has been done. The only analytical work dealing with the nonlinear problem is a recent one by Gopinath *et al.* (1998), who calculated the steady temperature distribution in a resonant channel starting from a boundary layer approximation to the complete two-dimensional formulation. The problem was solved only to the extent needed to find the second-order time-independent part of the temperature field.

The purpose of the present paper is to obtain a time-domain description of the evolution of the thermoacoustic instability and to determine its saturation level and its dependence on the parameters of the problem. Thus, this work is quite different from that of Gopinath *et al.*, where the oscillations were externally forced. The complex calculations of these authors suggest that the pursuit of our objective on the basis of the exact equations might be, at least for the time being, overly ambitious. For this reason, we base our weakly nonlinear analysis on the simplified quasi-one-dimensional model developed in earlier papers (Watanabe *et al.*, 1997; Yuan *et al.*, 1997), using a method of multiple time scales similar to the one in our recent study of the linear problem

(Karpov and Prosperetti, 1998). As with all such calculations, the results are valid only in a neighborhood of the instability threshold. We assume that the flow remains laminar, which is justified particularly near threshold (see, e.g., Akhavan *et al.*, 1991; Cooper *et al.*, 1993).

In a typical prime mover arrangement, it is only the fundamental mode that is unstable, while the higher-order modes are stable. In order for the instability to saturate, therefore, it is necessary that the energy injected into the fundamental mode be transferred by nonlinear couplings to the higher modes, where it is dissipated. Since the energy input is proportional to the square of the amplitude of the fundamental, while the energy dissipation is proportional to the square of the amplitude of the higher modes which, due to the nonlinearity, are at least quadratic in the amplitude of the fundamental, at sufficiently high amplitudes a balance is reached that allows the system to reach steady state.

The mathematical model that we use contains a great deal of information on the geometry and other characteristics of the system. Our results are therefore useful to explore the dependence of both the transient and the steady state on these parameters. As a demonstration, we show the dependence of the saturation amplitude on the temperature difference in the stack, the separation of the stack plates, the gas Prandtl number, and other quantities. In particular, it is found that the system is extremely sensitive to the shape of the resonant tube, which determines the mutual relation among the modes and therefore the ease with which energy can be transferred from the fundamental to the higher modes.

## I. MATHEMATICAL MODEL

In Watanabe *et al.* (1997), a nonlinear model for thermoacoustic devices was derived by averaging the exact conservation equations for mass, momentum, and energy over the cross section of the device. The model consists of the equations of continuity

$$\frac{\partial \rho}{\partial t} + \frac{1}{S} \frac{\partial}{\partial x} (S \rho u) = 0, \quad (1.1)$$

momentum

<sup>a)</sup>Electronic mail: prosperetti@jhu.edu

$$\rho \frac{\partial u}{\partial \hat{t}} + \rho u \frac{\partial u}{\partial x} + \frac{\partial p}{\partial x} = -\mathcal{D}(u), \quad (1.2)$$

and energy

$$\begin{aligned} \frac{\partial p}{\partial \hat{t}} + u \frac{\partial p}{\partial x} + \frac{\gamma p}{S} \frac{\partial(Su)}{\partial x} \\ = (\gamma - 1) \left[ \mathcal{H}(T_w - T) - \frac{dT_w}{dx} \mathcal{Q}(u) + u \mathcal{D}(u) \right] \end{aligned} \quad (1.3)$$

supplemented by the perfect-gas equation of state

$$p = \mathcal{R} \rho T. \quad (1.4)$$

In these equations  $\hat{t}$  is the dimensional time,  $\rho$ ,  $p$ ,  $T$ , and  $u$  are the cross-sectional averages of the density, pressure, temperature, and velocity fields of the gas, respectively and  $T_w$  is the prescribed temperature of the solid surfaces in contact with the gas, i.e., the tube walls and the stack plates. Unlike Gopinath *et al.* (1998), who were interested in the temperature distribution caused by a prescribed acoustic field, our focus here is on the evolution of the acoustic field under the destabilizing action of an externally imposed wall temperature gradient. For this reason, and in view of the large thermal capacity of typical stack plates, for simplicity we assume that  $T_w = T_w(x)$  is a prescribed function of  $x$ . With relatively straightforward adjustments, however, the present method can be extended to deal with a  $T_w$  dependent on time over the fast time scale of the acoustic oscillations as well as over the slower time scale of the developing instability. The spatial coordinate  $x$  is directed along the axis of the device,  $S(x)$  is the (in general, nonuniform) cross-sectional area,  $\gamma$  the ratio of specific heats, and  $\mathcal{R}$  the universal gas constant divided by the gas molecular mass.

The key to the success of the model (1.1) to (1.4) lies in the proper specification of the operators  $\mathcal{D}$ ,  $\mathcal{H}$ , and  $\mathcal{Q}$  that account for the exchange of momentum and energy between the gas and the solid surfaces in contact with it. In the model of Watanabe *et al.* (1997), these quantities were specified in such a way that the resulting theory, upon linearization, became identical to the exact linear theory developed by Rott (1969, 1983; see also Swift, 1989). Here, we deviate in two significant aspects from our earlier work. In the first place, since the nature of the perturbation method that we use enables us to distinguish between different Fourier components of the solution, we do not need to commit ourselves to a frequency-independent form for the operators. Second, while in the previous paper we extended the original definition of the operators to the nonlinear domain, here we retain their linear form. The motivation is not only one of simplicity but also, as mentioned at the end of Sec. IV, the fact that both numerical and perturbation results obtained with and without the nonlinear extension are very close. This circumstance is fortunate in view of the uncertainty affecting their proper extension to the nonlinear regime.

Using the forms of  $\mathcal{H}$  and  $\mathcal{Q}$  given in our previous paper (Watanabe *et al.*, 1997), in the linear approximation, we have

$$\begin{aligned} \mathcal{T}(p', u') &\equiv \mathcal{H}(T_w - T) - \frac{dT_w}{dx} \mathcal{Q}(u') \\ &= -i \hat{\omega} f_K p' - \rho_0 c_p \frac{f_V - f_K}{(1 - f_V)(1 - \sigma)} \frac{dT_w}{dx} u', \end{aligned} \quad (1.5)$$

where primes denote linear perturbation values, i.e.,  $p' = p - p_0$ , with  $p_0$  the undisturbed pressure. This expression has been written in the frequency domain and a dependence of  $p'$  and  $u'$  proportional to  $\exp i \hat{\omega} t$  is implied. The parameters  $f_{V,K}$  depend on the ratio of the diffusion lengths  $\delta_{V,K}$  to the plate spacing  $l$  and are given by (Swift, 1988)

$$f = (1 - i) \frac{\delta}{l} \tanh(1 + i) \frac{l}{2\delta}, \quad (1.6)$$

where the index can be  $V$  or  $K$ . The viscous and thermal penetration lengths  $\delta_{V,K}$  are given by

$$\delta_V = \sqrt{\frac{2\nu}{\hat{\omega}}}, \quad \delta_K = \sqrt{\frac{2\alpha}{\hat{\omega}}} = \frac{\delta_V}{\sqrt{\sigma}}, \quad (1.7)$$

with  $\nu$  the kinematic viscosity,  $\sigma$  the Prandtl number, and  $\alpha = \nu/\sigma$  the thermal diffusivity. For a circular cross section with radius  $r_0$ , one has instead (Rott, 1969)

$$f = \frac{2J_1((i-1)(r_0/\delta))}{(i-1)(r_0/\delta)J_0((i-1)(r_0/\delta))}, \quad (1.8)$$

again valid for both  $f_V$  and  $f_K$ . When the diffusion penetration depths are small compared with either  $l$  or  $r_0$ ,  $f$  given by either (1.6) or (1.8) admits the asymptotic approximation

$$f \simeq (1 - i) \frac{\delta}{l}, \quad f \simeq (1 - i) \frac{\delta}{r_0}. \quad (1.9)$$

More generally, for other cross-section shapes, in this limit one may set  $f \simeq 2(1 - i)\delta/d_h$ , where  $d_h = 4S(x)/s(x)$  is the hydraulic diameter defined in terms of the cross-sectional area  $S$  and the ‘‘wetted’’ perimeter  $s$ . The action of the wall drag operator  $\mathcal{D}$  is expressed similarly

$$\mathcal{D}(u') = i \rho_0 \hat{\omega} \frac{f_V}{1 - f_V} u'. \quad (1.10)$$

We reiterate that, upon linearization, in the frequency domain the previous model reduces exactly to Rott’s formulation.

## II. ASYMPTOTIC ANALYSIS

In setting up a perturbation approach to the problem posed in the previous section, we note that experiment (see, e.g., Wheatley, 1986; Atchley, 1990) shows that the initial build-up of the thermoacoustic instability has the character of a modulated standing wave, the frequency of which is essentially dictated by the resonator, while the amplitude is slowly varying in time. Similarly, if the thermoacoustic device is driven externally by a loudspeaker or a piston, the steady-state temperature distribution evolves slowly over the time scale of the acoustic period. These observations suggest the possibility of setting up a perturbation scheme based on

the smallness of the ratio of the characteristic period of oscillation to the characteristic time for the evolution of the thermoacoustic effect.

In the framework of the previous model, the terms in the left-hand side of Eqs. (1.1)–(1.3) describe nonlinear oscillations in a gas column with variable cross-sectional area and temperature stratification and are therefore responsible for the “carrier” frequency of the wave. The heart of the thermoacoustic effect is in the terms in the right-hand sides. The observed slowness of the modulation implies that the effect of these terms is small. As will be seen in the following, this effect arises through integrals over the tube length. Therefore, the effect will be small not only when the terms themselves are small but—as usually happens in practice—when they are large only over a small fraction of the tube length. A formal justification of this statement will be found in Appendix A. In order to set up a perturbation scheme, we introduce a small parameter  $\epsilon$  and set

$$F_K = \frac{f_K}{\epsilon}, \quad F_Q = \frac{1}{\epsilon} \frac{f_V - f_K}{(1 - f_V)(1 - \sigma)}, \quad F_D = \frac{1}{\epsilon} \frac{f_V}{1 - f_V}, \quad (2.1)$$

with  $F_{K,Q}$  formally treated as  $O(1)$  quantities. The considerations of Appendix A show that this does not necessarily imply that the  $f$ 's are of order  $\epsilon$ , but only that

$$\frac{L_S}{L} f_{V,K} \sim \epsilon, \quad (2.2)$$

where  $L_S$  is the length of the stack region, i.e., the region where the  $f$ 's are not small. An explicit definition of  $\epsilon$  is not necessary as the final results do not explicitly depend on this parameter, but one may think of it as the ratio of the standing wave period to the modulation time scale.

We use the method of multiple time scales (see, e.g., Kevorkian and Cole, 1996; Murdock, 1991; Hinch, 1991) and introduce the new time variables

$$t = \hat{t}, \quad \tau = \epsilon \hat{t}, \quad \theta = \epsilon^2 \hat{t}, \quad \eta = \epsilon^3 \hat{t}. \quad (2.3)$$

As a consequence of these definitions, we have

$$\frac{\partial}{\partial \hat{t}} = \frac{\partial}{\partial t} + \epsilon \frac{\partial}{\partial \tau} + \epsilon^2 \frac{\partial}{\partial \theta} + \epsilon^3 \frac{\partial}{\partial \eta} + \dots \quad (2.4)$$

The field variables are also expanded in a power series in  $\epsilon$ ; for example

$$p'(x, \hat{t}) = \epsilon p_1(x, \hat{t}) + \epsilon^2 p_2(x, \hat{t}) + \epsilon^3 p_3(x, \hat{t}) + \epsilon^4 p_4(x, \hat{t}) + \dots, \quad (2.5)$$

with analogous expressions for  $u'$ , etc. These expansions imply that the nonlinearity is taken to be of the same order as the amplitude modulation, which is the interesting case. Indeed, if the modulation were much stronger than the nonlinearity, we would essentially be dealing with the linear problem already studied in our earlier paper (Karpov and Prosperetti, 1998). On the other hand, since the wave exists only due to the thermoacoustic instability, one cannot have a strong nonlinearity coupled with weak thermoacoustic effects.

For clarity, it is important to stress two important aspects in which the work described in this paper differs from most other nonlinear stability studies. In the first place, while the deviation from marginal stability conditions is usually measured by a single control variable, here it is a whole function—the wall temperature distribution  $T_w$ —that determines the stability properties of the linear system. The situation might be reduced to the more usual one by assuming that the temperature distribution in the stack has a certain functional form dependent upon one parameter which would then play the role of control variable. For example, a linear temperature distribution would be characterized by the temperature gradient. Such an assumption is, however, unnecessarily restrictive and it is preferable to keep the framework general.

Second, perturbation expansions are usually carried out in the neighborhood of linear marginal stability conditions that are known exactly. In order to proceed in this way, we would have to solve the linear problem including the exchange terms  $f_{V,K}$ , find the marginal stability conditions, and then allow for a perturbation. Here, such a procedure would not lead to very transparent results given the complexity of the linear problem, compounded by the fact that, with nonzero  $f$ 's, the linear operator is not self-adjoint. Hence, we prefer to take a different approach choosing as our base state the situation in which there is no coupling between the gas and the solid structure even though, in general, this situation is not marginally stable. The marginal stability condition will be determined as part of the perturbation procedure itself, much as in our earlier paper (Karpov and Prosperetti, 1998), of which the first few steps of the present perturbation procedure constitute a refinement.

Here, therefore, as the parameter  $\epsilon$  is increased from zero, two effects appear: the coupling with the solid structure dampens or amplifies the oscillations, and nonlinear effects influence their development. It will be seen that the first effect sets in at a lower order in  $\epsilon$  than the second one. The first few steps of the perturbation procedure, therefore, furnish successively closer approximations to the linear stability threshold, from which the nonlinear effects eventually take off.

In order to deal with this situation we must clarify the role played by the wall temperature distribution  $T_w(x)$ . In the usual perturbation procedures, the small parameter is the difference between the given value of the control variable and the critical value. Here, we deal with the function  $T_w$ , and the critical conditions are not known exactly. Thus, we represent the given function  $T_w$  as

$$T_w(x) = T_{w0}(x) + \epsilon T_{w1}(x) + \epsilon^2 T_{w2}(x). \quad (2.6)$$

The terms  $T_{w0}, T_{w1}$  will be determined so as to satisfy the marginal stability conditions, after which the difference  $T_w(x) - [T_{w0}(x) + \epsilon T_{w1}(x)]$  will be the driving force for the development of the instability to the order considered.<sup>1</sup> In principle, the same procedure can be applied to higher orders in  $\epsilon$ . Since the appearance of the instability depends on integral conditions on the  $T_{wj}$ 's, in principle, there is a degree of nonuniqueness here since the  $T_{wj}$ 's can be taken arbitrarily, provided only the integral conditions are satisfied. In spite of

this seeming level of arbitrariness, one would expect the results to be insensitive to any specific choice, and indeed we have verified this numerically as will be described below.

From (2.6), we also have

$$\frac{dT_w}{dx} = G_0 + \epsilon G_1 + \epsilon^2 G_2, \quad (2.7)$$

where

$$G_j = \frac{dT_{wj}}{dx}, \quad j=0,1,2,\dots \quad (2.8)$$

Furthermore, since the equilibrium density is connected to  $p_0$  and  $T_w$  by the equation of state, when we write

$$\begin{aligned} \rho(x, \hat{t}) = & \rho_0 + \epsilon \rho_1(x, \hat{t}) + \epsilon^2 \rho_2(x, \hat{t}) + \epsilon^3 \rho_3(x, \hat{t}) \\ & + \epsilon^4 \rho_4(x, \hat{t}) + \dots, \end{aligned} \quad (2.9)$$

the corrections  $\rho_1$ , etc. will contain time-independent components related to  $T_{w1}$ , etc. With all these definitions we may write

$$\mathcal{T}(p', u') = \epsilon^2 \mathcal{T}_1 + \epsilon^3 \mathcal{T}_2 + \epsilon^4 \mathcal{T}_3, \quad (2.10)$$

$$\mathcal{D}(u') = \epsilon^2 \mathcal{D}_1 + \epsilon^3 \mathcal{D}_2 + \epsilon^4 \mathcal{D}_3, \quad (2.11)$$

where

$$\mathcal{T}_1 = -i\omega F_{Kp_1} - \rho_0 c_p F_{Q} G_0 u_1, \quad \mathcal{D}_1 = i\rho_0 \omega F_{V} u_1, \quad (2.12)$$

etc.

All the expansions, together with (2.4), are substituted into Eqs. (1.1)–(1.4) and the various orders in  $\epsilon$  are separated, giving rise to a sequence of problems that we describe separately. Many of the details of the calculation can be found in Appendix B. Only a streamlined summary of the procedure is described here.

### A. Order zero

At order zero the solution is

$$u_0 = 0, \quad (2.13)$$

with  $p_0$  and  $T_0 = T_{w0}(x)$  arbitrary but regarded as given.

### B. First order

The problem to order one was studied in an earlier paper (Karpov and Prosperetti, 1998) and the details will be omitted here. The equations are

$$\frac{\partial \rho_1}{\partial t} + \frac{1}{S} \frac{\partial}{\partial x} (S \rho_0 u_1) = 0, \quad (2.14)$$

$$\rho_0 \frac{\partial u_1}{\partial t} + \frac{\partial p_1}{\partial x} = 0, \quad (2.15)$$

$$\frac{\partial p_1}{\partial t} + \frac{\gamma p_0}{S} \frac{\partial (S u_1)}{\partial x} = 0, \quad (2.16)$$

and the pressure field is found to be

$$p_1 = A(\tau, \theta, \eta) P_1(x) \exp(i\omega t) + \text{c.c.}, \quad (2.17)$$

where  $A$  is the slowly varying amplitude, c.c. denotes the complex conjugate, and  $P_1$  is the solution of

$$\frac{1}{S} \frac{d}{dx} \left( c_0^2(x) S \frac{dP_1}{dx} \right) + \omega^2 P_1 = 0, \quad (2.18)$$

where

$$c_0^2(x) = \gamma \mathcal{R} T_{w0}(x), \quad (2.19)$$

is the local adiabatic sound speed.

The way in which  $T_{w0}$  is to be understood needs some clarifications, which amplify the comments made before Eq. (2.6). Let us consider first the case in which a certain temperature distribution  $T_w(x)$  is prescribed. If one plans to study only the linear problem,  $T_{w0}$  can be taken as the given  $T_w$ . On the other hand, if the plan is to carry the expansion to include nonlinear effects, as will be discussed below,  $T_{w0}$  must be determined (or, at least, constrained) using the results of the next step in the perturbation procedure. Finally, if the objective is the determination of the onset temperature distribution,  $T_w(x)$  is unknown at the outset, and the  $T_{w0}$  appearing in (2.18) is its first approximation that will be determined at the next step. Actually, the last two possibilities are one and the same as, if one plans to carry the expansion to include nonlinear effects, one needs to determine the onset temperature distribution first. A consequence of the indeterminacy of  $T_{w0}$  at this stage is an indeterminacy of  $\rho_0$ , which will be determined from the equation of state  $\mathcal{R} \rho_0 T_{w0} = p_0$  once  $T_{w0}$  is found.

We take the tube to be rigidly terminated at the ends, and therefore impose that

$$\frac{dP_1}{dx} = 0 \quad \text{at } x=0, x=L, \quad (2.20)$$

which ensure that the eigenvalue  $\omega^2$  is real and positive (see, e.g., Morse and Feshbach, 1953, p. 728; Naylor and Sell, 1982, p. 502). The eigenfunction  $P_1$  can therefore also be taken real and, for later convenience, we normalize it so that

$$\int_0^L S(x) P_1^2(x) dx = V p_0^2, \quad (2.21)$$

where  $V$  is the volume of the device.

In deriving an expression for the gas density to this order, we anticipate the fact that, at the next order, we will encounter a contribution  $T_{w1}$  that will affect the undisturbed value of  $\rho$ . Hence, we write

$$\rho_1 = \rho_{10} + \rho_{11}, \quad (2.22)$$

where  $\rho_{10}$  will be determined at the third step from the equation of state

$$\mathcal{R}(\rho_0 + \epsilon \rho_{10})(T_{w0} + \epsilon T_{w1}) = p_0. \quad (2.23)$$

The term  $\rho_{11}$  is instead associated to  $p_1$  and can be found from (2.14) to (2.16), together with  $u_1$ . The dependence of both quantities on the time variables and the perturbation amplitude  $A$  is the same as (2.17), and the spatial dependence is given by

$$R_1(x) = -\frac{1}{\omega^2 S} \frac{d}{dx} \left( S \frac{dP_1}{dx} \right), \quad (2.24)$$



$$U_1(x) = \frac{i}{\omega \rho_0} \frac{dP_1}{dx}. \quad (2.25)$$

### C. Second order

As shown in Appendix B, at second order one finds the following equation for  $p_2$ :

$$\frac{\partial^2 p_2}{\partial t^2} - \frac{1}{S} \frac{\partial}{\partial x} \left( c_0^2 S \frac{\partial p_2}{\partial x} \right) = (\text{RHS})_2, \quad (2.26)$$

where  $c_0^2$  is defined in (2.19) and

$$\begin{aligned} (\text{RHS})_2 = & (\gamma - 1) \frac{\partial \mathcal{T}_1}{\partial t} - 2 \frac{\partial^2 \rho_1}{\partial t \partial \tau} \\ & - \frac{\partial}{\partial t} \left[ u_1 \frac{\partial p_1}{\partial x} + \frac{\gamma p_1}{S} \frac{\partial}{\partial x} (S u_1) \right] \\ & + \frac{\gamma p_0}{S} \frac{\partial}{\partial x} S \left[ \frac{1}{\rho_0} \mathcal{D}_1 + u_1 \frac{\partial u_1}{\partial x} + \frac{\rho_1}{\rho_0} \frac{\partial u_1}{\partial t} \right]. \end{aligned} \quad (2.27)$$

The forcing at frequencies  $\pm \omega$  in the right-hand side will generate resonant terms proportional to  $t \exp(\pm i \omega t)$  in the solution for  $p_2$  which would lead to a breakdown of the approximation over times of order  $(\epsilon \omega)^{-1}$ . To avoid this resonance, as in the standard procedure (Kevorkian and Cole, 1996; Murdock, 1991; Hinch, 1991), we impose the solvability condition that the right-hand side of the equation be orthogonal to the solutions of the (adjoint) homogeneous equation, namely  $\exp(\pm i \omega t) P_1$

$$\int_0^{2\pi/\omega} dt \int_0^L S [\exp(\mp i \omega t) P_1] (\text{RHS})_2 dx = 0. \quad (2.28)$$

Multiplication by  $S(x)$  before integration is necessary so that the  $x$  operator in the left-hand side of (2.26) be self-adjoint with the boundary conditions (2.20). The two conditions (2.28) give

$$\frac{\partial A}{\partial \tau} - i \Omega_1 A = 0, \quad (2.29)$$

plus its complex conjugate, where

$$\begin{aligned} \Omega_1 = & - \frac{1}{2V\rho_0^2\omega} \int_0^L S \left[ c_0^2 \left( F_D + \frac{\rho_{10}}{\rho_0} \right) \left( \frac{dP_1}{dx} \right)^2 \right. \\ & \left. + (\gamma - 1) \omega^2 F_K P_1^2 + (\gamma - 1) c_p F_Q G_0 P_1 \frac{dP_1}{dx} \right] dx. \end{aligned} \quad (2.30)$$

The solution of (2.29) is

$$A(\tau, \theta, \eta) = B(\theta, \eta) \exp(i \Omega_1 \tau), \quad (2.31)$$

which increases exponentially if  $\text{Im} \Omega_1 < 0$ . It will be noted that  $\text{Im} \Omega_1$  exists only due to the  $F$  terms that account for momentum and energy exchange with the solid structure and is therefore not affected by the indeterminacy of  $\rho_{10}$  at this stage of the calculation. It will be seen below that, in order to continue the perturbation scheme into the nonlinear domain, for consistency it is necessary to have  $\text{Im} \Omega_1 = 0$ , which is a constraint on  $G_0$  and, therefore,  $T_{w0}$ . This constraint can of

course be met in an infinity of ways, although one would not expect any particular choice to have consequences for the remainder of the calculation provided the choice is such that  $T_w - T_{w0} = O(\epsilon)$ , in keeping with (2.6). Indeed, as will be shown below, this is in agreement with our numerical evidence.

In our earlier paper (Karpov and Prosperetti, 1998) we gave explicit expressions for the onset temperature gradient  $G_0$  on the assumption that it is constant in the stack. By means of the short-stack approximation, it was possible to derive simple formulas that generalize the concept of critical gradient of elementary theory to the presence of viscous effects and narrow gaps.

When (2.28) is satisfied, one can look for the solution of Eq. (2.26) for  $p_2$  in the form of the superposition of terms at frequencies 0,  $\pm \omega$ , and  $\pm 2\omega$ , respectively

$$p_2 = A P_{21}(x) e^{i \omega t} + A^2 P_{22}(x) e^{2i \omega t} + A^* A P_{20}(x) + \text{c.c.} \quad (2.32)$$

Here, the asterisk denotes the complex conjugate. The equations for  $P_{21}(x)$ ,  $P_{22}(x)$ , and  $P_{20}(x)$  are obtained from Eq. (2.26) upon separating the different frequency components in the right-hand side. However,  $P_{20}$  is more easily determined from the momentum equation (B5). The explicit form of these equations is given in (B7)–(B9) in Appendix B.

### D. Third order

Proceeding as before, at the third order we find

$$\frac{\partial^2 p_3}{\partial t^2} - \frac{1}{S} \frac{\partial}{\partial x} \left( c_0^2 S \frac{\partial p_3}{\partial x} \right) = (\text{RHS})_3, \quad (2.33)$$

where

$$\begin{aligned} (\text{RHS})_3 = & (\gamma - 1) \frac{\partial}{\partial t} (\mathcal{T}_2 + u_1 \mathcal{D}_1) - 2 \frac{\partial^2 p_1}{\partial t \partial \theta} - \frac{\partial^2 p_2}{\partial t \partial \tau} \\ & - \frac{\partial}{\partial t} \left( u_1 \frac{\partial p_2}{\partial x} + u_2 \frac{\partial p_1}{\partial x} + \frac{\gamma p_1}{S} \frac{\partial (S u_2)}{\partial x} \right. \\ & \left. + \frac{\gamma p_2}{S} \frac{\partial (S u_1)}{\partial x} \right) + \frac{1}{S} \frac{\partial (S c_0^2 \mathcal{D}_2)}{\partial x} + \frac{\gamma p_0}{S} \frac{\partial}{\partial x} \\ & \times \left( S \left[ \frac{\partial u_2}{\partial \tau} + \frac{\partial (u_1 u_2)}{\partial x} + \frac{\rho_1}{\rho_0} \left( \frac{\partial u_2}{\partial t} + \frac{\partial u_1}{\partial \tau} + u_1 \frac{\partial u_1}{\partial x} \right) \right. \right. \\ & \left. \left. + \frac{\rho_2}{\rho_0} \frac{\partial u_1}{\partial t} \right] \right). \end{aligned} \quad (2.34)$$

The solvability condition is again (2.28), which leads to

$$\frac{\partial A}{\partial \theta} - i(\Omega_2 + \Lambda_2 |A|^2) A = 0, \quad (2.35)$$

where

$$\begin{aligned} \Omega_2 = & \frac{1}{2\omega_0 V p_0^2} \int_0^L S c_0^2 \frac{dP_1}{dx} \left[ \left( F_D^2 + F_D \frac{\rho_{10}}{\rho_0} + \frac{T_{w2}}{T_{w0}} \right) \frac{dP_1}{dx} \right. \\ & - \left. \left( F_D + \frac{\rho_{10}}{\rho_0} \right) \frac{dP_{21}}{dx} \right] dx - \frac{\Omega_1^2}{2\omega} \\ & - \frac{\gamma-1}{2\omega V p_0^2} \int_0^L S P_1 \left[ 2\omega \Omega_1 F_K P_1 + \omega^2 F_K P_{21} \right. \\ & \left. + c_p F_Q G_1 \frac{dP_1}{dx} + c_p F_Q G_0 \left( \frac{dP_{21}}{dx} - F_D \frac{dP_1}{dx} \right) \right] dx, \end{aligned} \quad (2.36)$$

$$\begin{aligned} \Lambda_2 = & \frac{i}{2V p_0^2} \int_0^L S P_1 \left[ U_1 \left( 2 \frac{dP_{20}}{dx} - \frac{dP_{22}}{dx} \right) + U_{22} \frac{dP_1}{dx} \right. \\ & \left. + \frac{\gamma P_1}{S} \frac{d(SU_{22})}{dx} + \frac{\gamma}{S} (2P_{20} - P_{22}) \frac{d(SU_1)}{dx} \right] dx \\ & - \frac{\gamma p_0}{2\omega_0 V p_0^2} \int_0^L S \frac{dP_1}{dx} \left[ U_1 \frac{dU_{22}}{dx} + U_{22} \frac{dU_1}{dx} + \frac{R_1}{\rho_0} \right. \\ & \left. \times \left( U_1 \frac{dU_1}{dx} - 2i\omega U_{22} \right) - \left( \frac{R_{22}}{\rho_0} + 2 \frac{P_{20}}{p_0} \right) i\omega U_1 \right] dx. \end{aligned} \quad (2.37)$$

Since in  $\Lambda_2$  all the  $U$ 's are pure imaginary while the  $P$ 's are real, this quantity is real. On the other hand,  $\Omega_2$  is complex in general but becomes real in neutrally stable conditions

where  $\text{Im } \Omega_2 = 0$  gives the second approximation to the onset critical gradient.

Upon substituting into (2.35) Eq. (2.31) for  $A$ , we have

$$\frac{\partial B}{\partial \theta} - i[\Omega_2 + \Lambda_2 |B|^2 \exp(-2 \text{Im } \Omega_1 \tau)] B = 0. \quad (2.38)$$

Since, by hypothesis,  $B$  cannot depend on  $\tau$ , this equation is only consistent provided  $\text{Im } \Omega_1 = 0$ , which, as noted before, ensures that the temperature gradient corresponds to the onset value at this level of approximation. This consistency condition motivated the expansion of  $T_w$  into the series of terms (2.6) to be chosen so that the condition can be satisfied. When  $\text{Im } \Omega_1 = 0$ , then we have

$$B(\theta, \eta) = C(\eta) \exp(i[\Omega_2 + \Lambda_2 |C|^2] \theta), \quad (2.39)$$

The solution  $p_3$  is given by

$$\begin{aligned} p_3 = & (A P_{311}(x) + A^* A^2 P_{313}) e^{i\omega t} + A^2 P_{32}(x) e^{2i\omega t} \\ & + A^3 P_{33}(x) e^{3i\omega t} + A^* A P_{30}(x) + \text{c.c.} \end{aligned} \quad (2.40)$$

The equations for  $P_{311}(x)$ ,  $P_{313}(x)$ ,  $P_{32}(x)$ , and  $P_{30}(x)$  are given in (B28)–(B32) of Appendix B.

#### E. Fourth order

At fourth order we find

$$\frac{\partial^2 p_4}{\partial t^2} - \frac{1}{S} \frac{\partial}{\partial x} \left( c_0^2 S \frac{\partial p_4}{\partial x} \right) = (\text{RHS})_4, \quad (2.41)$$

where

$$\begin{aligned} (\text{RHS})_4 = & -2 \frac{\partial^2 p_1}{\partial t \partial \eta} - \frac{\partial^2 p_2}{\partial t \partial \theta} - \frac{\partial^2 p_3}{\partial t \partial \tau} + (\gamma-1) \frac{\partial}{\partial t} (\mathcal{T}_3 + u_1 \mathcal{D}_2 + u_2 \mathcal{D}_1) - \frac{\partial}{\partial t} \left( u_1 \frac{\partial p_3}{\partial x} + u_2 \frac{\partial p_2}{\partial x} + u_3 \frac{\partial p_1}{\partial x} \right. \\ & \left. + \frac{\gamma}{S} \left( p_1 \frac{\partial(Su_3)}{\partial x} + p_2 \frac{\partial(Su_2)}{\partial x} + p_3 \frac{\partial(Su_1)}{\partial x} \right) \right) + \frac{\gamma p_0}{S} \frac{\partial}{\partial t} \left( S \left[ \frac{1}{\rho_0} \mathcal{D}_3 + \frac{\partial u_3}{\partial \tau} + \frac{\partial u_2}{\partial \theta} \right. \right. \\ & \left. \left. + \frac{\rho_1}{\rho_0} \left( \frac{\partial u_3}{\partial t} + \frac{\partial u_2}{\partial \tau} + \frac{\partial u_1}{\partial \theta} + \frac{\partial(u_1 u_2)}{\partial x} \right) + \frac{\rho_2}{\rho_0} \left( \frac{\partial u_2}{\partial t} + \frac{\partial u_1}{\partial \tau} + u_1 \frac{\partial u_1}{\partial x} \right) + \frac{\rho_3}{\rho_0} \frac{\partial u_1}{\partial t} + \frac{\partial(u_1 u_3)}{\partial x} + u_2 \frac{\partial u_2}{\partial x} \right] \right). \end{aligned} \quad (2.42)$$

For the purpose of extracting the slow time evolution of the disturbance amplitude, we only need the part of  $(\text{RHS})_4$  proportional to  $\exp \pm i\omega t$ , on which the solvability condition (2.28) is to be imposed. The result is

$$\frac{dC}{d\eta} - i[\Omega_3 + \Lambda_3 |C|^2 \exp(-2 \text{Im } \Omega_2 \theta)] C = 0, \quad (2.43)$$

together with its complex conjugate. Here,

$$\begin{aligned} \Omega_3 = & \frac{i}{2\omega V p_0^2} \int_0^L S c_0^2 \frac{dP_1}{dx} [F_D U_{311} + (\Omega_1 \rho_0 + \rho_{10})(U_{311} + F_D U_{21}) + (\Omega_2 \rho_0 + \Omega_1 \rho_{10} + \omega \rho_{200})(U_{21} + F_D U_1) \\ & + (\Omega_2 \rho_{10} + \Omega_1 \rho_{200} + \omega \rho_{300}) U_1] dx + \frac{i(\gamma-1)}{2V p_0^2} \int_0^L S P_1 [i F_K (\omega P_{311} + \Omega_1 P_{21} + \Omega_2 P_1) \\ & + c_p F_Q (\rho_0 G_0 U_{311} + (\rho_0 G_1 + \rho_{10} G_0) U_{21} + (\rho_0 G_2 + \rho_{10} G_1 + \rho_{200} G_0) U_1)] dx, \end{aligned} \quad (2.44)$$

$$\begin{aligned}
\Lambda_3 = & \frac{i}{2V\rho_0^2} \int_0^L SP_1 \left[ 2U_1 \frac{d}{dx} (\text{Re } P_{30}) - U_1 \frac{dP_{32}}{dx} + U_{21}^* \frac{dP_{22}}{dx} + 2U_{21} \frac{dP_{20}}{dx} + U_{22} \frac{dP_{21}^*}{dx} + U_{32} \frac{dP_1^*}{dx} + 2U_{30} \frac{dP_1}{dx} \right] dx \\
& + \frac{i\gamma}{2V\rho_0^2} \int_0^L P_1 \left[ (2\text{Re } P_{30} - P_{32}) \frac{d(SU_1)}{dx} + 2P_{20} \frac{d(SU_{21})}{dx} + P_{22} \frac{d(SU_{21}^*)}{dx} + P_{21}^* \frac{d(SU_{22})}{dx} + 2P_1 \frac{d(SU_{30})}{dx} \right. \\
& + \left. P_1 \frac{d(SU_{32})}{dx} \right] dx + \frac{i\gamma}{2V\rho_0^2} \int_0^L SP_1 \left[ iF_K(\omega P_{313} + \Lambda_2 P_1) + c_p F_D G_0(\rho_0 U_{313} + \rho_{202} U_1) + (F_D^* - 2F_{2D}) \frac{dP_1}{dx} U_{22} \right] dx \\
& + \frac{i\gamma\rho_0}{2V\rho_0^2} \int_0^L S \frac{dP_1}{dx} \left[ \left( \frac{\Omega_1}{\omega} + \frac{\rho_{10}}{\rho_0} + F_D \right) U_{313} + 2 \frac{R_1}{\rho_0} U_{32} + \left( \frac{\Lambda_2}{\omega} + \frac{\rho_{202}}{\rho_0} \right) U_{21} - \frac{R_{22}}{\rho_0} U_{21}^* + 2 \left( \frac{R_{21}^*}{\rho_0} + \frac{R_1}{\rho_0} \frac{\Omega_1}{\omega} \right) U_{22} \right. \\
& + \left. \left( \frac{\rho_{302}}{\rho_0} + \frac{R_{32}}{\rho_0} + \left( \frac{\rho_{10}}{\rho_0} + F_D \right) \frac{\Lambda_2}{\omega} + \left( \frac{\rho_{202}}{\rho_0} + \frac{R_{22}}{\rho_0} \right) \frac{\Omega_1}{\omega} + \frac{\rho_{202}}{\rho_0} F_D \right) U_1 \right] dx \\
& + \frac{\gamma\rho_0}{2\omega V\rho_0^2} \int_0^L S \frac{dP_1}{dx} \left[ \frac{d(U_{21}^* U_{22})}{dx} - \frac{d(U_1 U_{32})}{dx} + \frac{R_1}{\rho_0} \left( U_1 \frac{dU_{21}^*}{dx} - U_{21} \frac{dU_1}{dx} \right) - \frac{\rho_{10}}{\rho_0} \frac{d(U_1 U_{22})}{dx} \right. \\
& + \left. 2U_1 \frac{dU_{30}}{dx} - \frac{R_{21}}{\rho_0} U_1 \frac{dU_1}{dx} \right] dx. \tag{2.45}
\end{aligned}$$

As before in (2.38), Eq. (2.43) is only consistent if  $\text{Im } \Omega_2 = 0$ .

### III. RESULTS

The key result of the calculation described in the previous section is a relation governing the growth of the amplitude of the perturbation. Upon introducing the unscaled amplitude  $\tilde{A} = \epsilon A \exp(i\omega t)$ , from (2.29), (2.38), (2.43), we have

$$\frac{d\tilde{A}}{dt} = i[(\omega + \tilde{\Omega}_1 + \tilde{\Omega}_2 + \tilde{\Omega}_3) + (\tilde{\Lambda}_2 + \tilde{\Lambda}_3)|\tilde{A}|^2]\tilde{A}, \tag{3.1}$$

where  $\tilde{\Omega}_1$  is obtained by multiplying (2.30) by  $\epsilon$ ,

$$\begin{aligned}
\tilde{\Omega}_1 = & -\frac{1}{2V\rho_0^2\omega} \int_0^L dx S \left[ c_0^2 \left( \frac{f_V}{1-f_V} + \frac{\rho_{10}}{\rho_0} \right) \left( \frac{dP_1}{dx} \right)^2 \right. \\
& + \left. (\gamma-1)\omega^2 f_K P_1^2 + \frac{(\gamma-1)(f_V-f_K)}{(1-\sigma)(1-f_V)} c_p G_0 P_1 \frac{dP_1}{dx} \right], \tag{3.2}
\end{aligned}$$

with similar expressions for  $\tilde{\Omega}_k = \epsilon^k \Omega_k$  obtained from (2.36) and (2.44) and  $\tilde{\Lambda}_k = \epsilon^k \Lambda_k$  from (2.37), (2.45).

Equation (3.1) contains both linear and nonlinear contributions that are better discussed separately. In all the numerical examples that follow, we use as reference case the geometry of the experiment of Atchley *et al.* (1990), the implementation of which in the context of the present mathematical model is described in detail in an earlier paper (Yuan *et al.*, 1997). Briefly, the system consists of a 38.2-mm-diameter tube with a length of 99.89 cm, a 35-mm-long stack of stainless-steel plates located 87.95 cm from the cold end, and two heat exchangers. The operating pressure was  $p_0 = 307$  kPa.

#### A. Linear regime

In the linear approximation, Eq. (3.1) is

$$\frac{d\tilde{A}}{dt} = i(\omega + \tilde{\Omega}_1 + \tilde{\Omega}_2 + \tilde{\Omega}_3)\tilde{A}. \tag{3.3}$$

A consideration of Eqs. (2.38) and (2.43) shows that, for the linear problem, one can allow nonzero imaginary parts for  $\Omega_1$  and  $\Omega_2$ . For this reason, it is not necessary to expand  $T_w$  in a perturbation series. Accordingly, in the previous expressions, one can set  $T_{w0} = T_w$ ,  $T_{w1} = T_{w2} = 0$ ,  $G_0 = dT_w/dx$ ,  $G_1 = G_2 = 0$ ; furthermore, all time-independent corrections to  $\rho$  due to the  $T_{wj}$ 's, i.e.,  $\rho_{10}$ ,  $\rho_{200}$ ,  $\rho_{300}$ , vanish. In addition to giving the growth rate of the perturbation for a given temperature distribution  $T_w$ , Eq. (3.3) can be used to determine onset conditions by adjusting the temperature distribution in such a way that  $\text{Im}(\tilde{\Omega}_1 + \tilde{\Omega}_2 + \tilde{\Omega}_3) = 0$ .

The results found in this way may be considered as a higher-order extension of those presented in our previous paper (Karpov and Prosperetti, 1998) where only the expression for  $\tilde{\Omega}_1$  was given. In that study, we showed that this result of relatively low order of accuracy is already in excellent agreement with Rott's exact linear theory prediction. Equation (3.3) can therefore be considered a higher-order approximation to Rott's theory and indeed we have confirmed this expectation by evaluating numerically  $\tilde{\Omega}_2$  and  $\tilde{\Omega}_3$ . A typical example is shown in Fig. 1, where the imaginary part of the exact linear eigenfrequency—i.e., minus the growth rate of the instability—(solid line) is compared with the successive approximations  $\text{Im } \tilde{\Omega}_1$  (dotted),  $\text{Im}(\tilde{\Omega}_1 + \tilde{\Omega}_2)$  (dashed), and  $\text{Im}(\tilde{\Omega}_1 + \tilde{\Omega}_2 + \tilde{\Omega}_3)$  (long dashes). Here, the horizontal axis is the normalized position  $x_S/L$  of the midpoint of the stack. The numbered points in this figure will be explained later. As mentioned before and in Eq. (B21) of Appendix B, here we evaluate  $f_{V,K}$  at the value of  $\omega$  given by

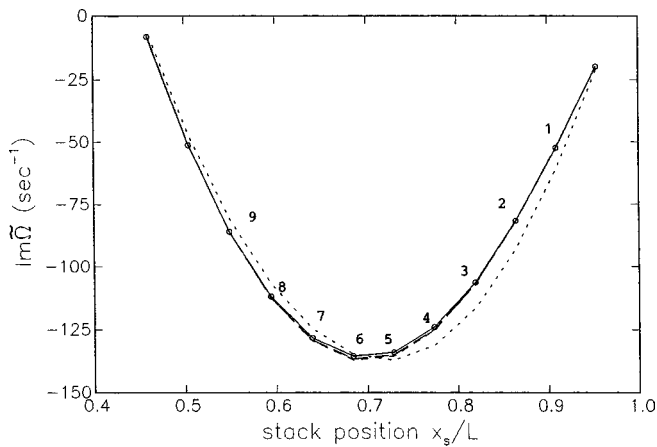


FIG. 1. Negative of the instability growth rate, i.e., imaginary part of the exact linear eigenfrequency (solid line), compared with the successive approximations  $\text{Im } \tilde{\Omega}_1$  (dotted line),  $\text{Im } \tilde{\Omega}_1 + \tilde{\Omega}_2$  (short dashes), and  $\text{Im } \tilde{\Omega}_1 + \tilde{\Omega}_2 + \tilde{\Omega}_3$  (long dashes) as functions of the normalized position  $x_s/L$  of the midpoint of the stack for the geometry described in the text. The temperature difference along the stack is 368 K. The numbered points are keyed to the lines in Fig. 6.

the first-order approximation (2.18) instead of updating it at each step of the perturbation procedure. We have verified that this simplification deteriorates the results only slightly.

We do not show any further results of this type as the earlier first-order approximation is already adequate for practical purposes and the added computational effort to go beyond it significant. Nevertheless, the fact bears mentioning as it confirms the correctness of the perturbation technique and the accuracy of the analysis. It is more interesting to focus on the chief novelty of this work, namely its nonlinear aspects.

## B. Nonlinear regime

Equation (3.1), and its complex conjugate, predict the evolution of the oscillation amplitude in time. These equations can be combined to find

$$\frac{d|\tilde{A}|^2}{d\hat{t}} = -2[\text{Im } \tilde{\Omega}_3 + \text{Im } \tilde{\Lambda}_3 |\tilde{A}|^2] |\tilde{A}|^2, \quad (3.4)$$

which is readily solved with the result

$$|\tilde{A}|^2 = \frac{|\tilde{A}|_{\text{sat}}^2}{1 + \exp(2\tilde{\Omega}_3 \hat{t}) (|\tilde{A}|_{\text{sat}}^2 / |\tilde{A}|_0^2 - 1)}, \quad (3.5)$$

where  $\tilde{A}_0$  denotes the initial value of the amplitude and

$$|\tilde{A}|_{\text{sat}}^2 = -\frac{\text{Im } \tilde{\Omega}_3}{\text{Im } \tilde{\Lambda}_3} = -\frac{\tilde{\Omega}_{3i}}{\tilde{\Lambda}_{3i}}. \quad (3.6)$$

Since  $\tilde{\Omega}_{3i} < 0$  for an unstable system, this equation shows that, for  $\hat{t} \rightarrow \infty$ ,  $|\tilde{A}|^2$  is asymptotic to the saturation value  $|\tilde{A}|_{\text{sat}}^2$  as also follows directly from the right-hand side of (3.1) since only  $\tilde{\Omega}_3$  and  $\tilde{\Lambda}_3$  have nonzero imaginary parts. In the stable case,  $\tilde{\Omega}_{3i}$  is positive and the amplitude decays to zero.

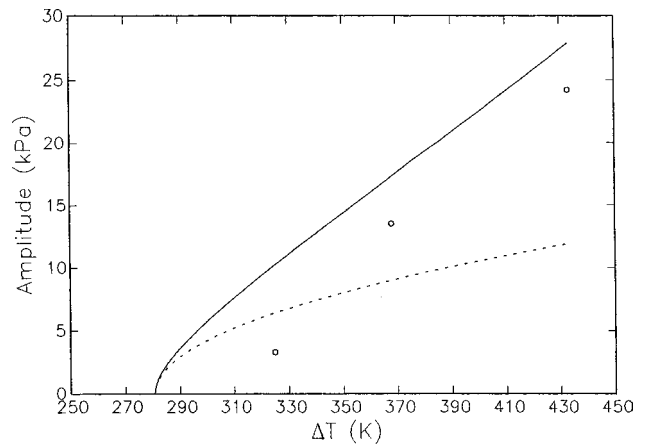


FIG. 2. Saturation amplitude of the fundamental mode (dotted line) and peak amplitude of the pressure perturbation at the cold end of the resonant tube (solid line) as a function of the temperature difference  $\Delta T$  along the stack. The symbols are the data points reported by Atchley *et al.* (1990).

We begin our discussion from the asymptotic saturation amplitude  $\tilde{A}_{\text{sat}}$ . This result is of particular interest as it contains the effect of a large number of variables such as the deviation of the temperature distribution from onset conditions, the shape of the resonator, the effect of stack geometry (through the values of the exchange parameters  $f_{V,K}$ ) and stack length, Prandtl number, and others.

We assume a linear temperature variation between a cold temperature  $T_C$  and a hot temperature  $T_H$  in the stack, and constant temperatures equal to  $T_C$  and  $T_H$  at the left and right of the stack, respectively. We have tested two different definitions of the terms in the perturbation expansion (2.6) of  $T_w$ . In the first one we kept  $T_C$  fixed and gradually raised  $T_H$ . In the second one, we kept the average temperature along the tube constant, adjusting the temperature difference  $T_H - T_C$ . For the same final  $T_H$  and  $T_C$ , the two procedures gave nearly identical results as expected on the basis of the considerations of Sec. II.

We consider first the experimental situation of Atchley *et al.* (1990). The solid line in Fig. 2 is the (positive) peak amplitude as a function of the temperature difference along the stack. The dotted line shows  $|\tilde{A}_{\text{sat}}|^2$ . For this case our calculations indicate an onset temperature of 573.5 K which, with  $T_C = 293$  K, corresponds to an onset temperature gradient of 8.01 K/mm. One recognizes here the typical structure of a supercritical bifurcation. In this figure the symbols represent the three data points reported by Atchley *et al.* (1990). In order to reconcile these data with our theoretical prediction, one would have to assume hot-end temperatures 36, 22, and 20 K lower than the reported values. We found a similar effect in our earlier numerical study (Yuan *et al.*, 1997). Two possible concurrent explanations come to mind. The first one is the temperature jump that is established between the ends of the stack and the heat exchangers, which has the effect of reducing the temperature gradient in the stack (Brewster *et al.*, 1997). Second, the temperature values reported in Atchley *et al.* were measured at the surface of the tube, rather than in the stack. As mentioned in Yuan *et al.* (1997), for the conditions of the experiment one may reasonably expect a temperature difference of this order between these

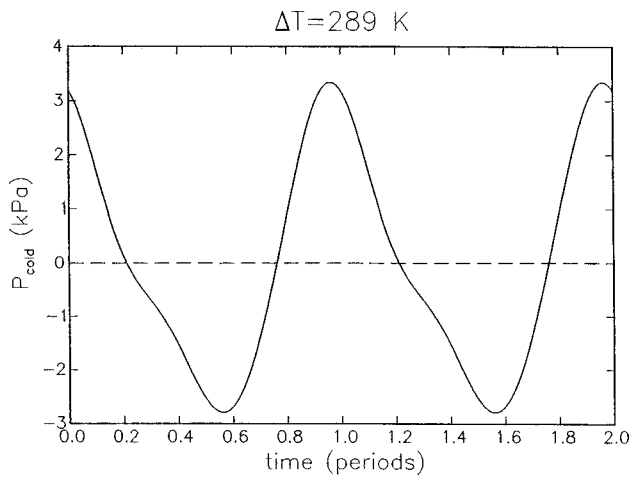


FIG. 3. Temporal waveform at the cold end of the tube for the lowest temperature case of Atchley *et al.* (1990). The temperature difference along the stack has been adjusted to 289 K so as to match the amplitude reported by Atchley *et al.* The measured period was 2.1 ms, to be compared with the present result of 1.96 ms.

two positions. In addition, the experimental setup will most likely include several losses (e.g., due to the lack of alignment between the heat exchanger and stack plates, which causes an increase of the effective blockage of the tube, heat conduction in the gas, and others) that are not included in our model. Hence, the difference with the data is in the expected direction and of a reasonable magnitude. The discrepancy at the lowest temperature difference is somewhat larger but, on the other hand, it is here that the effect of losses would be greatest. In our earlier numerical study, where we tried to accommodate additional loss mechanisms by smoothing the cross-sectional area and wall temperature distributions, we actually found that this low-temperature case was predicted to be stable.

The temporal waveform at the cold end of the tube corresponding to the case with the lowest temperature difference along the stack is shown in Fig. 3. Here, the temperature difference has been adjusted so as to match the amplitude reported by Atchley *et al.* Upon comparing with Fig. 2 of that reference, one sees the same steepening of the wave profile and an asymmetry in the curvatures of the peaks and valleys similar to the measured one. The second harmonic is, however, slightly more prominent in our result. Figure 4 shows the calculated spectrum. The ratio of the second and third harmonics to the first one is  $-12$  and  $-29$  dB, respectively, to be compared with the experimental values of  $-16$  and  $-32$  dB. The oscillation period measured in the experiment was 2.1 ms, which is in good agreement with our result of 1.96 ms.

The saturation amplitude for the case at the next higher temperature is much larger and the data show a substantial harmonic content extending to at least the sixth harmonic. Since our method only gives results up to the third harmonic, the comparison with the measured waveform is not as satisfactory (Fig. 5). In this case the second and third harmonics are 3 and 11 dB below the fundamental, to be compared with reported values of 7.6 and 14.5. Thus, in this case as in the previous one, the second harmonic is found to be larger than

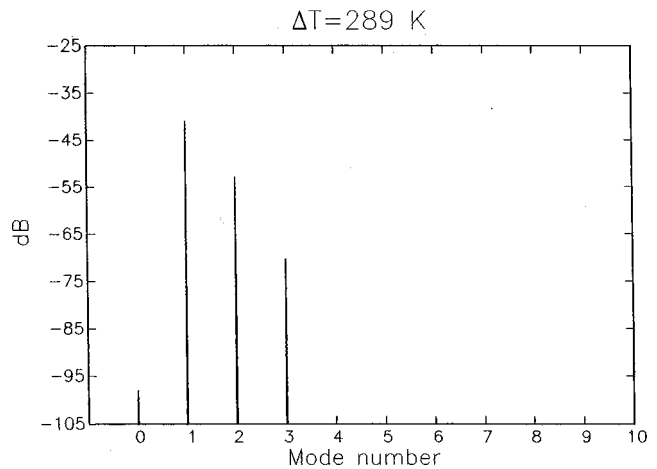


FIG. 4. Frequency spectrum for the waveform of the previous figure.

the measured one. In our experience, the amplitude of this harmonic is strongly dependent on the details of the resonator geometry and it is therefore also possible that this difference is due to a mismatch between the experimental arrangement and the way it is modeled here rather than to inaccuracies of the perturbation method. The measured period was 2.1 ms, to be compared with our result of 1.96 ms.

The expressions derived before contain a dependence on many variables, the effect of which can also be studied. To produce the results that follow, the stack length was taken as 30 mm while the geometry and other parameter values were the same as before, unless otherwise noted. To reduce the number of parameters, however, we removed the heat exchangers and took the stack plates to have zero thickness so that there is no blockage of the tube's cross section. Furthermore, momentum and energy transfer with the solid structure away from the stack region were neglected.

Figure 6 shows the dependence of the (negative) peak amplitude at the cold end of the tube upon the position of the stack as a function of the temperature difference along the stack. The positions are keyed to those marked by numbers

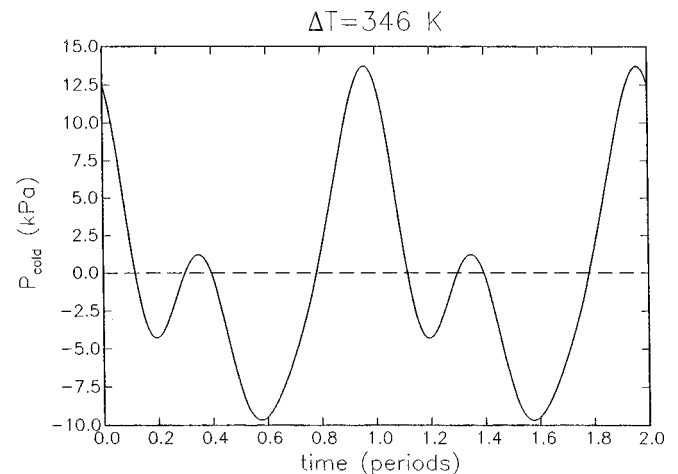


FIG. 5. Temporal waveform at the cold end of the tube for the intermediate temperature case of Atchley *et al.* (1990). The temperature difference along the stack has been adjusted to 346 K to match the amplitude reported by Atchley *et al.* The measured period was 2.1 ms, to be compared with the present result of 1.95 ms.

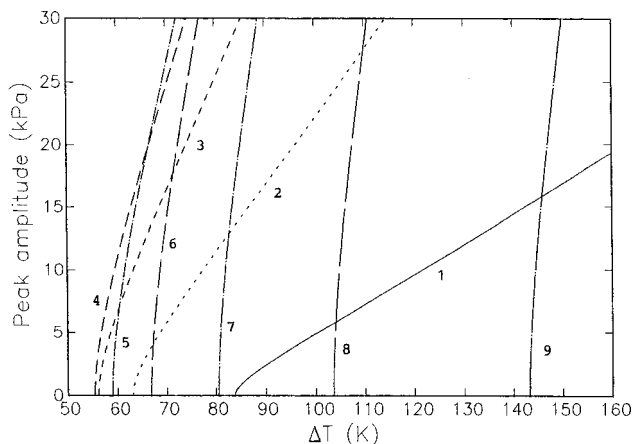


FIG. 6. Dependence of the peak amplitude at the cold end of the tube upon the position of the stack as a function of the temperature difference along the stack. The stack positions are keyed to the numbers marked in Fig. 1.

in Fig. 1. The strongest dependence on the temperature difference occurs near the three-quarters position along the tube as expected.

Figure 7 shows, for a stack positioned at  $x_S/L=0.729$ , the dependence of the (negative) peak amplitude at the cold end of the tube upon the stack-plate spacing. This is the same case considered in the linear approximation in Fig. 4 of Karpov and Prosperetti (1998). The solid line is for  $\delta_K/l=0.31$ , the dashed line for  $\delta_K/l=0.24$ , and the dash-and-dot line for  $\delta_K/l=0.40$ . The dotted lines are the amplitudes of the first harmonic. It can be seen that an optimal value of  $\delta_K/l \approx 0.31$  exists in this case. The value of  $\delta_K$  used here is based on the first eigenvalue of (2.18) and is 0.2 mm.

The dependence on the Prandtl number is shown in Fig. 8 for a stack placed at  $x_S/L=0.774$ . From the left, the results are for  $\sigma=0.5, 0.72$ , and 1.0. The dotted lines are the first-harmonic amplitude. As expected, the smaller the Prandtl number the lower the onset temperature difference, although the slope of the lines is not very sensitive to this quantity.

The effect of the tube length  $L$  is shown in Fig. 9. The

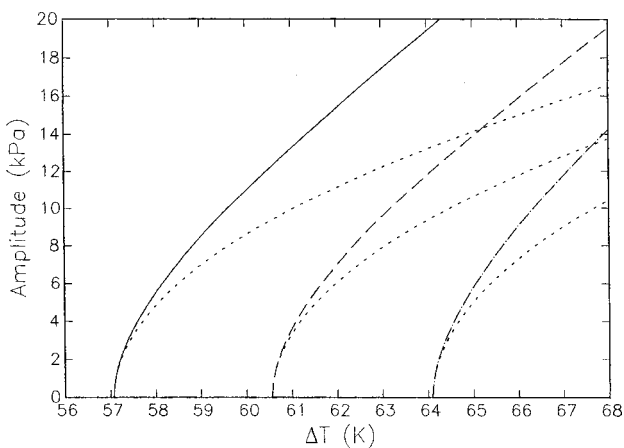


FIG. 7. Dependence of the peak amplitude at the cold end of the tube upon the dimensionless thermal penetration depth  $\delta_K/l$  for a stack positioned at  $x_S/L=0.729$ . The solid line is for  $\delta_K/l=0.31$ , the dashed line for  $\delta_K/l=0.24$ , and the dash-and-dot line for  $\delta_K/l=0.40$ . The dotted lines are the amplitudes of the first harmonic.

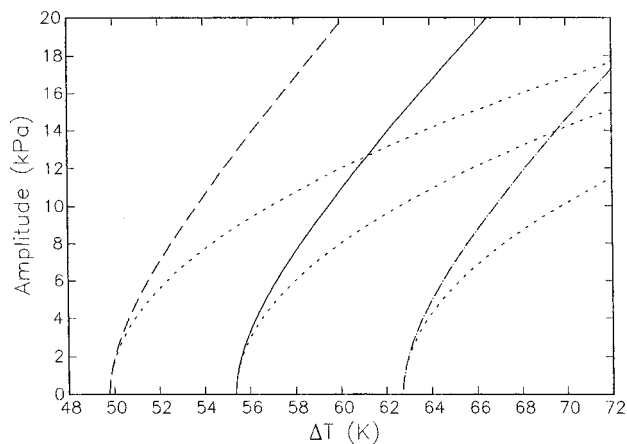


FIG. 8. Dependence of the peak amplitude at the cold end of the tube upon the gas Prandtl number  $\sigma$  for a stack position of  $x_S/L=0.774$ . From the left, the lines are for  $\sigma=0.5, 0.72$ , and 1.0. The dotted lines are the first-harmonic amplitude.

middle curve is for  $L=99.89$  cm (i.e., the length used by Atchley *et al.*, 1990). The lines on the left and the right are for tube lengths of 124.86 and 84.91 cm, respectively. For all three cases,  $x_S/L=0.774$ . An increase in  $L$  over the range considered decreases the onset temperature gradient and destabilizes the system. This result is mainly due to the lowering of the onset gradient as shown in Eq. (51) of our earlier paper (Karpov and Prosperetti, 1998).

The model we have used can accommodate tubes with a nonuniform cross section. To illustrate this effect, we consider a tube with cross-sectional area given by

$$\frac{S(x)}{S(0)} = \begin{cases} 1 & 0 \leq x \leq \frac{1}{4}L \\ [1 + Z \cos^2 \pi(2x/L - 1)]^2 & \frac{1}{4}L \leq x \leq \frac{3}{4}L \\ 1 & \frac{3}{4}L \leq x \leq L. \end{cases} \quad (3.7)$$

Figure 10 shows the peak amplitudes for several values of the parameter  $Z$  and  $x_S/L=0.774$ . For  $Z > 1.5\%$  the phase of the harmonics is such that the compression phase of the

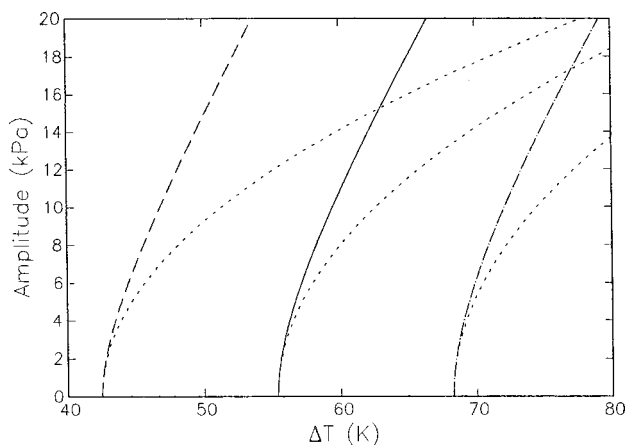


FIG. 9. Dependence of the peak amplitude at the cold end on the tube length  $L$ . The middle curve is for the length used by Atchley *et al.* (1990); the lines on the left and the right are for tube lengths 25% greater and 15% smaller. The dotted lines are the first-harmonic amplitude. In all cases, the stack is positioned at  $x_S/L=0.774$ .

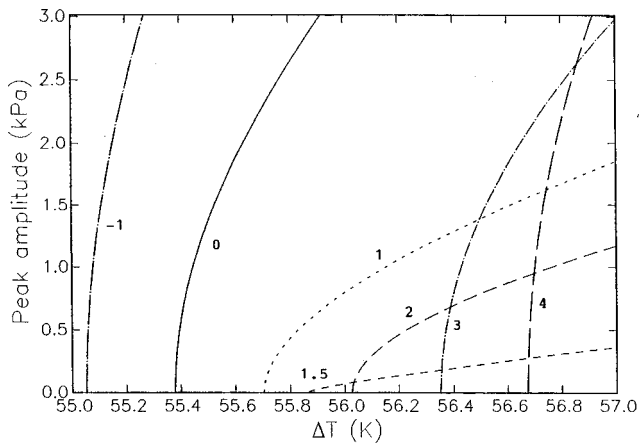


FIG. 10. Dependence of the peak amplitudes at the cold end of the tube upon the cross-sectional area variation as parametrized by the parameter  $Z$  in Eq. (3.7) expressed in percent. The stack is positioned at  $x_s/L=0.774$ .

wave has greater amplitude than the rarefaction phase. In this case, we show the positive peak amplitude. The reverse happens for other values of  $Z$ , for which we show the negative peak amplitude. The effect is very strong and we express  $Z$  in percent. As the cross section is widened there is a minimum occurring around  $Z=1.5\%$ . In these conditions the second harmonic is strongly resonant and its contribution to the spectrum is larger than that of the first harmonic. Narrowing cross sections destabilize the system, as already found earlier (Watanabe *et al.*, 1997).

The final parameter we consider is the tube diameter (Fig. 11). The middle curve is for a diameter of 38.2 mm, which is the same as used by Atchley *et al.* (1990); the lines on the left and on the right are for diameters of 45.8 and 30.6 mm, respectively. For these calculations of course we restore the exchange terms with the solid structure away from the stack region. The results show a destabilizing effect of an increasing radius. The observed effect is large because, even if the penetration depth is much less than the radius, the length of the tube is so large that the total fluid volume affected by the transfer of momentum and heat between the gas and the wall is comparable to the entire stack volume.

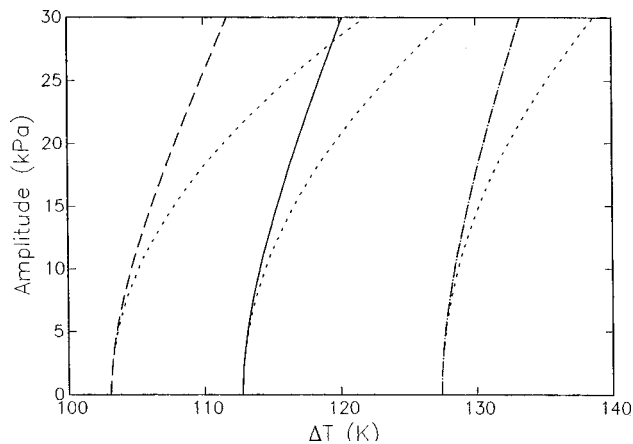


FIG. 11. Dependence of the peak amplitudes at the cold end of the tube upon the tube diameter. The middle curve is for the diameter used by Atchley *et al.* (1990); the lines on the left and the right are for diameters 20% bigger and 20% smaller. The dotted lines are the first-harmonic amplitude.

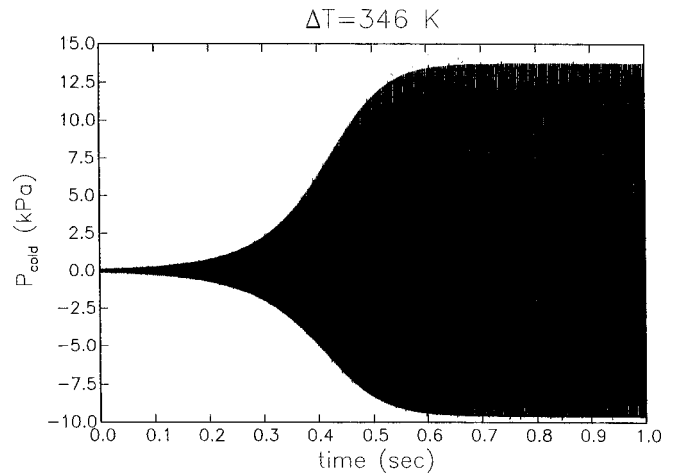


FIG. 12. Time dependence of the cold-end pressure for the middle-temperature case  $\Delta T=346$  K of Atchley *et al.* (1990) considered in Fig. 5 as predicted by Eq. (3.8).

If (3.5) is substituted into (3.1), the resulting equation is readily solved and the following complete solution for  $\tilde{A}(t)$  determined:

$$\frac{\tilde{A}(\hat{t})}{\tilde{A}_0} = \exp \left\{ i(\omega + \tilde{\Omega}_1 + \tilde{\Omega}_2 + \tilde{\Omega}_3)\hat{t} + i \frac{\tilde{\Lambda}_2 + \tilde{\Lambda}_3}{2\tilde{\Lambda}_{3i}} \right. \\ \left. \times \ln \left[ 1 + \frac{|\tilde{A}|_0^2}{|\tilde{A}|_{\text{sat}}^2} (\exp[-2\tilde{\Omega}_{3i}\hat{t}] - 1) \right] \right\}. \quad (3.8)$$

From this expression, the time dependence of all the fields can be calculated. As an example, Fig. 12 shows the cold-end pressure versus time for the middle-temperature case of Atchley *et al.* (1990) considered in Fig. 5. The general behavior of this result is in agreement with what is observed experimentally (see, e.g., Wheatley, 1986). The early-time behavior is shown on a much magnified time scale in Fig. 13, where the dashed line shows the envelope from Eq. (3.5).

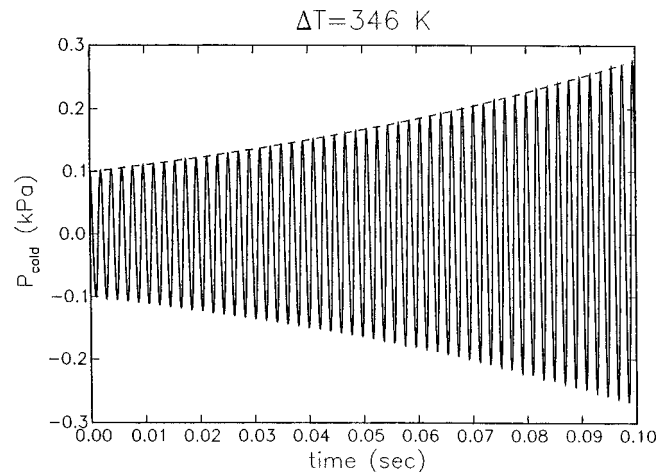


FIG. 13. Detail of the previous figure. The dashed line is the envelope from Eq. (3.5).

## IV. DISCUSSION

As mentioned in Sec. I, the form of the heat and momentum transfer operators that we have used in deriving the results discussed so far is somewhat uncertain in the nonlinear domain. Furthermore, it would be of great interest to know how accurate the asymptotic results are. In this section we address—albeit partially—these important questions with the help of the somewhat different heat-transfer model used in our earlier studies and by comparing the results with the fully nonlinear calculation of Yuan *et al.* (1997).

### A. Other form for the heat transfer operators

In our original model (Watanabe *et al.*, 1997), in the frequency domain the heat transfer operators  $\mathcal{H}$ ,  $\mathcal{Q}$  were given by

$$\mathcal{H}(T - T_w) = i\omega\rho_0 c_p \frac{f_K}{1 - f_K} (T - T_w), \quad (4.1)$$

$$\mathcal{Q}(u) = \frac{\rho_0 c_p}{1 - \sigma} \left( \frac{f_V}{1 - f_V} - \frac{\sigma f_K}{1 - f_K} \right) u. \quad (4.2)$$

The form (1.5) used before is found by eliminating the temperature difference  $T - T_w$  using the exact linear result given in Eq. (52) of Watanabe *et al.* (1997). The two formulations are therefore precisely equivalent if the linear problem is solved exactly. However, if, as in the present method, the linear problem is not solved exactly, the results are not necessarily the same.

The method of calculation is the same as that described in the previous sections. For example, the first-order result for  $\tilde{\Omega}_1$  is similar to (3.2), namely

$$\begin{aligned} \tilde{\Omega}_1 = & -\frac{1}{2V\rho_0^2\omega} \int_0^L dx S \left[ c_0^2 \left( \frac{f_V}{1 - f_V} + \frac{\rho_{10}}{\rho_0} \right) \left( \frac{dP_1}{dx} \right)^2 \right. \\ & + \gamma\omega^2 \frac{f_K}{1 - f_K} P_1 (P_1 - \mathcal{R}R_1 T_{w0}) \\ & \left. + \frac{\gamma - 1}{1 - \sigma} \left( \frac{f_V}{1 - f_V} - \frac{\sigma f_K}{1 - f_K} \right) c_p G_0 P_1 \frac{dP_1}{dx} \right]. \quad (4.3) \end{aligned}$$

The field  $P_1$  appearing here satisfies the same equation (2.18) found before. We omit further details, which can be found in Karpov (2000), and focus on the results.

Figure 14, where the horizontal axis is the position of the stack, compares the growth rates of the instability predicted by the model (4.1), (4.2) at different orders; the solid line is the exact linear result calculated from Rott's equation. The conditions are the same as in Fig. 1. The dotted, dashed, and long-dashed lines show the growth rate as predicted by the first-, second-, and third-order theory, respectively. Although the picture improves as the plate spacing becomes wider, a comparison with Fig. 1 shows that the model of Sec. I is superior to that of (4.1) and (4.2). The difference between the two models decreases as the plate spacing increases, but the model of Sec. I definitely appears to be more robust.

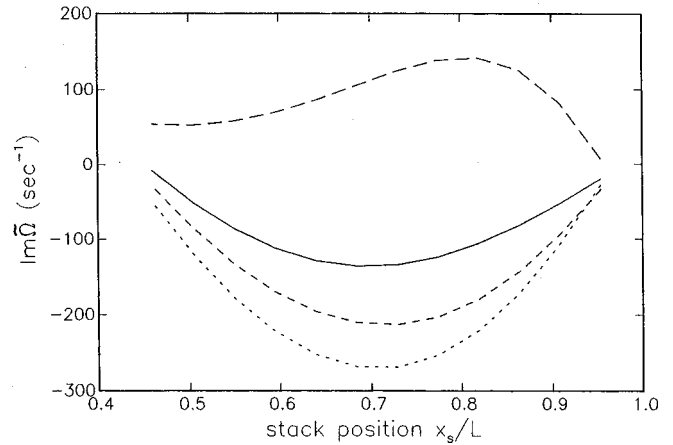


FIG. 14. Comparison of the exact linear growth rate of the instability given by Rott's model (solid line) with the first- (dotted), second- (dashed), and third-order (long dashes) results based on (4.1) and (4.2). Conditions as in Fig. 1. In this case the third-order approximation (top line) does not improve over the second-order one; it is possible that the asymptotic series does not converge. For wider gaps, however, we have found a systematic improvement as the order of the terms is increased.

### B. Comparison with numerical results

In order to assess the accuracy of our perturbation method, it would be desirable to compare the results of Sec. III B with fully nonlinear numerical results based on the same model. In principle, one could think of using the approach of our earlier paper (Yuan *et al.*, 1997), in which a numerical treatment of the model of Watanabe *et al.* (1997) was given. Unfortunately, this is not possible for the reasons that we now explain.

As noted before, the model of Watanabe *et al.* (1997) differs from the one of the present paper in the use of (4.1) and (4.2) in place of (1.5). In that work, in order to convert the forms (4.1) and (4.2) to the time domain, we fixed  $\omega$  to the (real) frequency of the lowest mode, and wrote the imaginary parts as time derivatives (see Yuan *et al.*, 1997 for details). If one attempts the same procedure on Eq. (1.5), the linear spectrum of the problem changes so severely that some of the higher modes become unstable. We had encountered a similar difficulty in Yuan *et al.* (1997), but there we were able to sidestep it by adding some artificial dissipation that damped this instability without much effect on the more important lower modes. We have found that it is much more difficult to reach this objective with the model of Sec. I.

For this reason, we will effect a comparison between the numerical and perturbation results based on (4.1) and (4.2). We have checked that, with a larger gap,  $l/\delta_K = 5.98$ , the nonlinear perturbation results based on (4.1), (4.2), and (1.5) are very close. Thus, it may be expected that the error introduced by the perturbation procedure would be comparable for the two models. Again, the perturbation calculation is quite analogous to that described before and leads to an equation for the amplitude of the same form as (3.1). Explicit expressions for the quantities  $\Omega_j$ ,  $\Lambda_k$  calculated with (4.1) and (4.2) can be found in Karpov (2000).

The results of a typical comparison are shown in Fig. 15, where the line is the perturbation result and the circles the numerically calculated peak amplitudes as function of the



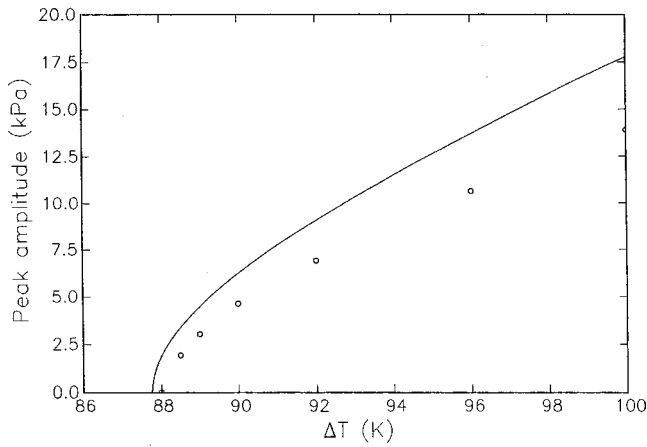


FIG. 15. Comparison between the numerically calculated peak amplitude (circles) and the perturbation results (line) as functions of the temperature difference along the stack. The dimensionless gap width is  $l/\delta_K=5.88$ .

temperature difference along the stack. Here, the stack length is 30 mm and its cold end is positioned at  $x_S/L=0.774$ . The gap in the stack has the value  $l/\delta_K=5.98$  and the results have been obtained by effecting the smoothing operations on  $S(x)$  and  $T_w(x)$  described in Yuan *et al.* (1997) 100 times. For both the analytical and the numerical results an artificial damping as described in Yuan *et al.* (1997) was used with  $k_n=0.15\text{ m}^2/\text{s}$ . The difference between the perturbation and numerical result increases with  $\Delta T$ , as expected, but the analytical prediction is nevertheless acceptably close to the results of the numerical computations.

One last item is the following. In the model of Watanabe *et al.* (1997) the time derivatives  $\partial/\partial t$  arising from the imaginary part of the heat and momentum transfer operators were transformed to a convective derivative  $\partial/\partial t + u(\partial/\partial x)$  to extend the definition to the nonlinear regime. We have compared numerically the results obtained with and without the convective term  $u(\partial/\partial x)$  finding very minor differences.

## V. CONCLUSIONS

We have presented a time-dependent weakly nonlinear theory of the build-up of unstable oscillations in a thermoacoustic prime mover. The expressions we have derived simulate the initial growth of the instability as well as its subsequent saturation. The results are in good agreement with the limited number of experimental observations reported in the literature for the configuration that we can model.

For all the examples that we have discussed, a solution of Rott's equation shows that the imaginary part of all the modes higher than the fundamental one is positive, so that they are all stable. Therefore, the second and higher harmonic amplitudes are nonzero only due to nonlinear energy transfers from the fundamental. In the present perturbation scheme, energy transfer to the second mode starts occurring at the second order. However, the energy loss of this mode due to the coupling with the solid structure only arises at the fourth order. Since the amplitude of the fundamental saturates at a level such that the energy transferred to the second mode can be dissipated by the mechanisms affecting this mode, in order to find the saturation amplitude of the funda-

mental, we had to carry the expansion to this order. The fact that the perturbation method only includes a small number of modes prevents energy transfer from the fundamental to the higher modes. As a consequence, our results tend to overestimate somewhat the saturation amplitude. Nevertheless, the comparison of Sec. IV shows that there is a parameter region where the error is acceptable.

Even if our results may not be sufficiently precise in absolute terms when the temperature gradient is too far from the onset value, they are still useful for comparative purposes as they indicate trends of the system performance when design parameters or operating conditions are varied. For example, we have demonstrated the very strong effect of cross-sectional area variations. Since the expressions we have derived contain a great amount of detail on the geometry and other characteristics of the system, similar sensitivity studies with respect to other parameter variations are possible and will be pursued in the future.

In spite of the algebraic complexity of the final form of the results, the present solution is certainly much easier to evaluate than carrying out a full-fledged multidimensional *ab initio* numerical calculation. In our earlier paper (Karpov and Prosperetti, 1998) it was possible to simplify the form of the solution by using the short-stack approximation. Here, this approximation is not very useful because the result depends on auxiliary fields such as  $P_{21}$ ,  $P_{22}$ , etc., the determination of which requires the solution of (B7), (B8), etc. Nevertheless, it might be possible to develop some comparable approximation, a topic that also will be pursued in future work.

## ACKNOWLEDGMENT

The authors express their gratitude to the Office of Naval Research for the support of this work.

## APPENDIX A: SCALING

In this Appendix we provide a justification of (2.2) and of the perturbation expansion used in the paper. For this purpose it is sufficient to consider the result (3.2) for  $\tilde{\Omega}_1$ ; since it will be sufficient to stop at this order in  $\epsilon$ , we omit the term  $\rho_{10}/\rho_0$  thus dealing, in effect, with the linear case already studied in our earlier paper [Karpov and Prosperetti, 1998, Eq. (32)]

$$\tilde{\Omega}_1 = -\frac{1}{2Vp_0^2\omega} \int_0^L dx S \left[ c_{01}^2 \frac{f_V}{1-f_V} \left( \frac{dP_1}{dx} \right)^2 + (\gamma-1) \times \omega^2 f_K P_1^2 + \frac{(\gamma-1)(f_V-f_K)}{(1-\sigma)(1-f_V)} c_p G_0 P_1 \frac{dP_1}{dx} \right]. \quad (\text{A1})$$

If we multiply Eq. (2.18) by  $SP_1$  and integrate over the tube length, by using the normalization (2.21) we find

$$\omega = \frac{1}{Vp_0^2\omega} \int_0^L dx S c_0^2 \left( \frac{dP_1}{dx} \right)^2. \quad (\text{A2})$$

The perturbation procedure that we have used is based on the assumption that  $\tilde{\Omega}_1/\omega$  is small. This ratio consists of the sum of three terms that we examine in turn. The first one is

$$N_1 = -\frac{1}{2} \frac{\int_0^L dx S c_0^2 [f_V / (1 - f_V)] (dP_1/dx)^2}{\int_0^L dx S c_0^2 (dP_1/dx)^2}. \quad (\text{A3})$$

In typical thermoacoustic engines,  $f_V$  is very small away from the stack, while it is of order 1 in the stack region of length  $L_S$ . If  $\bar{f}_V$  is the typical magnitude of  $f_V$  in this region, we thus have

$$N_1 \approx \frac{\bar{f}_V}{1 - \bar{f}_V} \frac{L_S}{L}. \quad (\text{A4})$$

Thus,  $N_1$  will be small when the relation (2.2) is verified.

For the second term,

$$N_2 = -\frac{1}{2} \frac{\int_0^L dx S (\gamma - 1) \omega^2 f_K P_1^2}{\int_0^L dx S c_0^2 (dP_1/dx)^2}, \quad (\text{A5})$$

we use the estimate

$$P_1 \sim L \frac{dP_1}{dx} \sim \frac{c_0}{\omega} \frac{dP_1}{dx}, \quad (\text{A6})$$

to conclude in a similar way that

$$N_2 \sim (\gamma - 1) \bar{f}_K \frac{L_S}{L}, \quad (\text{A7})$$

which, again, proves the smallness of  $N_2$  when (2.2) is verified.

Proceeding similarly in the case of the third term we note that, since  $(\gamma - 1)c_p = c_0^2/T_w$ , the estimate (A6) leads to

$$N_3 \sim \frac{(\bar{f}_V - \bar{f}_K)}{(1 - \sigma)(1 - \bar{f}_V)} \frac{G_0}{\bar{T}_w} L_S \sim \frac{\Delta T}{\bar{T}_w}. \quad (\text{A8})$$

Formally, this quantity is small when  $\Delta T \ll \bar{T}_w$ , which is seldom the case in real thermoacoustic engines. However, this is the only potentially destabilizing term and its imaginary part is therefore of the opposite sign of the contribution of the other two terms. The  $\Delta T$  in this result should therefore more properly be interpreted as the amount by which the temperature difference along the stack differs from its threshold value. This estimate confirms the previous statement that the present theory is applicable near onset conditions.

## APPENDIX B: DETAILS OF THE CALCULATION

Here, we supply some details on the calculation with intermediate expressions that can be useful to better understand the method and check the results. For clarity, it is useful to rewrite the equations of the model casting all the nonlinear terms into the right-hand sides. The result is

$$\frac{\partial \rho'}{\partial t} + \frac{1}{S} \frac{\partial}{\partial x} (S \rho_0 u') = -\frac{1}{S} \frac{\partial}{\partial x} (S \rho' u'), \quad (\text{B1})$$

$$\rho_0 \frac{\partial u'}{\partial t} + \frac{\partial p'}{\partial x} = -\rho' \frac{\partial u'}{\partial t} - (\rho_0 + \rho') u' \frac{\partial u'}{\partial x} - \mathcal{D}(u'), \quad (\text{B2})$$

$$\frac{\partial p'}{\partial t} + \frac{\gamma p_0}{S} \frac{\partial (S u')}{\partial x} = -u' \frac{\partial p'}{\partial x} - \frac{\gamma p'}{S} \frac{\partial (S u')}{\partial x} + (\gamma - 1) \times [\mathcal{T}(p', u') + u' \mathcal{D}(u')]. \quad (\text{B3})$$

Note that the gas temperature is not involved in these equations. At each step of the perturbation procedure this quantity can be determined from the equation of state (1.4).

As in Sec. II, we separate the different steps in the perturbation method.

### A. Second order

At second order, Eqs. (B1)–(B3) give

$$\frac{\partial \rho_2}{\partial t} + \frac{1}{S} \frac{\partial}{\partial x} (S \rho_0 u_2) = -\frac{\partial \rho_1}{\partial \tau} - \frac{1}{S} \frac{\partial (S \rho_1 u_1)}{\partial x}, \quad (\text{B4})$$

$$\rho_0 \frac{\partial u_2}{\partial t} + \frac{\partial p_2}{\partial x} = -\mathcal{D}_1 - \rho_0 \left( \frac{\partial u_1}{\partial \tau} + u_1 \frac{\partial u_1}{\partial x} \right) - \rho_1 \frac{\partial u_1}{\partial t}, \quad (\text{B5})$$

$$\frac{\partial p_2}{\partial t} + \frac{\gamma p_0}{S} \frac{\partial (S u_2)}{\partial x} = (\gamma - 1) \mathcal{T}_1 - \frac{\partial p_1}{\partial \tau} - u_1 \frac{\partial p_1}{\partial x} - \frac{\gamma p_1}{S} \frac{\partial}{\partial x} (S u_1). \quad (\text{B6})$$

The second and third equations can be combined to give the single equation (2.26) for  $p_2$  by taking the derivative with respect to time of (B6) and using (B5) to eliminate  $\partial u_2/\partial t$ .

The functions  $P_{21}$ ,  $P_{22}$ , and  $P_{20}$  appearing in the expansion (2.32) of  $p_2$  satisfy the equations

$$\begin{aligned} \frac{1}{S} \frac{d}{dx} \left( c_0^2 S \frac{dP_{21}}{dx} \right) + \omega^2 P_{21} \\ = -2\omega \Omega_1 P_1 - (\gamma - 1) \omega^2 F_K P_1 - (\gamma - 1) c_p F_Q G_0 \frac{dP_1}{dx} \\ + \frac{1}{S} \frac{d}{dx} \left[ S c_0^2 \left( F_D + \frac{\rho_{10}}{\rho_0} \right) \frac{dP_1}{dx} \right], \end{aligned} \quad (\text{B7})$$

$$\begin{aligned} \frac{1}{S} \frac{d}{dx} \left( c_0^2 S \frac{dP_{22}}{dx} \right) + 4\omega^2 P_{22} \\ = -\frac{\gamma p_0}{S \omega^2} \frac{d}{dx} \left[ \frac{1}{\rho_0^3} \frac{d(S \rho_0)}{dx} \left( \frac{dP_1}{dx} \right)^2 \right] \\ - 2 \left[ \frac{1}{\rho_0} \left( \frac{dP_1}{dx} \right)^2 - \omega^2 \frac{P_1^2}{\rho_0} \right], \end{aligned} \quad (\text{B8})$$

$$\frac{dP_{20}}{dx} = -\frac{1}{S \omega^2} \frac{d}{dx} \left[ \frac{S}{\rho_0} \left( \frac{dP_1}{dx} \right)^2 \right]. \quad (\text{B9})$$

The first two equations are found by substituting (2.32) into (2.26) and separating the frequency components. The last equation could also be found in the same way, but it is more convenient to obtain it from the momentum equation. The boundary conditions for all these equations are the same [(2.20)] as for  $P_1$ . Note that these conditions are automatically satisfied by  $P_{20}$ , which is then determined only up to a

constant. To make the solution unique, we impose the additional condition

$$\int_0^L S \frac{P_{20}}{T_{w0}} dx = 0, \quad (\text{B10})$$

which follows from the conservation of the total gas mass in the tube.

We write the  $O(\epsilon^2)$  contribution to the gas density as

$$\rho_2 = \rho_{20} + \rho_{21} + \rho_{22}, \quad (\text{B11})$$

where the first index denotes the order in  $\epsilon$  and the second one the frequency. The time-independent term  $\rho_{20}$ , however, consists of two contributions, one,  $\rho_{200}$ , due to the corrections  $T_{w1}$  and  $T_{w2}$  of the wall temperature, to be determined later, and the other,  $\rho_{202}$ , analogous to  $P_{20}$  in (2.32), that exists due to nonlinear effects.

$$\rho_{20} = \rho_{200} + A^* A \rho_{202}. \quad (\text{B12})$$

Both quantities follow from the equation of state in the form

$$\rho_{200} = -\rho_{10} \frac{T_{w1}}{T_{w0}} - \rho_0 \frac{T_{w2}}{T_{w0}}, \quad \rho_{202} = 2\rho_0 \frac{P_{20}}{p_0}. \quad (\text{B13})$$

The time-dependent parts of  $\rho_2$  and  $u_2$  have the same structure as (2.32), except that  $U_{20} = 0$ . The spatial parts are

$$R_{21} = -\frac{\Omega_1}{\omega} R_1 - \frac{1}{\omega^2 S} \frac{d}{dx} \left[ S \left( \frac{dP_{21}}{dx} - \left( \frac{\Omega_1}{\omega} + F_D \right) \frac{dP_1}{dx} \right) \right], \quad (\text{B14})$$

$$R_{22} = \frac{i}{2\omega S} \frac{d}{dx} S(\rho_0 U_{22} + R_1 U_1), \quad (\text{B15})$$

$$U_{21} = \frac{i}{\omega \rho_0} \left[ \frac{d\rho_{21}}{dx} - \left( F_D + \frac{\Omega_1}{\omega} + \frac{\rho_{10}}{\rho_0} \right) \frac{dP_1}{dx} \right], \quad (\text{B16})$$

$$U_{22} = \frac{i}{2\omega \rho_0} \left[ \frac{dP_{22}}{dx} + \frac{1}{\rho_0 \omega^2} \left( \frac{dP_1}{dx} \right)^2 \left( \frac{1}{S} \frac{dS}{dx} + \frac{1}{\rho_0} \frac{d\rho_0}{dx} \right) \right]. \quad (\text{B17})$$

The contributions of the drag and heat transfer operators necessary for the next order split into parts at frequency  $\omega$  and  $2\omega$  as

$$\mathcal{D}_2 = \mathcal{D}_{21} + \mathcal{D}_{22}, \quad \mathcal{T}_2 = \mathcal{T}_{21} + \mathcal{T}_{22}. \quad (\text{B18})$$

In writing  $\mathcal{D}_{21}$ , care must be exerted because  $\mathcal{D}$  acting on

$$\begin{aligned} u_1 &= B(\theta, \eta) \exp(i\omega t + i\Omega_1 \tau) U_1(x) + \text{c.c.} \\ &= B(\theta, \eta) \exp[i(\omega + \epsilon\Omega_1)t] U_1(x) + \text{c.c.} \end{aligned} \quad (\text{B19})$$

gives

$$\begin{aligned} \mathcal{D}_1 &= i(\omega + \epsilon\Omega_1)(\rho_0 + \epsilon\rho_{10}) F_D U_1 B(\theta, \eta) \\ &\quad \times \exp[i(\omega + \epsilon\Omega_1)t] + \text{c.c.} \end{aligned} \quad (\text{B20})$$

The terms containing  $\epsilon$  have properly been disregarded in writing the form (2.12) of  $\mathcal{D}_1$  for use with the second-order problem. However, this term contributes to the next order

and therefore it must be restored. We include it in the definition of  $\mathcal{D}_{21}$  and  $\mathcal{T}_{21}$ , and we thus find

$$\begin{aligned} \mathcal{D}_{21} &= iF_D [\rho_0 \omega U_{21} + (\rho_0 \Omega_1 + \rho_{10} \omega) U_1] B(\theta, \eta) \\ &\quad \times \exp[i(\omega + \epsilon\Omega_1)t] + \text{c.c.} \\ &= -iF_D \left[ \frac{dP_{21}}{dx} - F_D \frac{dP_1}{dx} \right] B(\theta, \eta) \\ &\quad \times \exp[i(\omega + \epsilon\Omega_1)t] + \text{c.c.} \end{aligned} \quad (\text{B21})$$

Strictly speaking, the definition of  $f_V$  should also be modified to contain  $\omega + \epsilon\Omega_1$ , rather than  $\omega$ . However, in the range of practical interest, the derivative of  $f_V$  with respect to  $\omega$  is small and the correction numerically not very important. We therefore disregard it for simplicity. The component at frequency  $2\omega$  is

$$\mathcal{D}_{22} = 2i\omega \rho_0 F_{2D} U_{22} A^2(\tau, \theta, \eta) \exp(2i\omega t) + \text{c.c.}, \quad (\text{B22})$$

where we write  $F_{2D}$  as a reminder that  $\omega$  should be replaced by  $2\omega$  in the definition (1.5). Similarly, we have

$$\begin{aligned} \mathcal{T}_{21} &= [-i\omega F_K P_{21} - i\Omega_1 F_K P_1 - c_p F_Q (\rho_0 G_0 U_{21} \\ &\quad + \rho_{10} G_0 U_1 + \rho_0 G_1 U_1)] B(\theta, \eta) \exp[i(\omega + \epsilon\Omega_1)t] \\ &\quad + \text{c.c.}, \end{aligned} \quad (\text{B23})$$

$$\begin{aligned} \mathcal{T}_{22} &= [-i2\omega F_{2K} P_{22} - \rho_0 c_p F_{2Q} G_0 U_{22}] A^2(\tau, \theta, \eta) \\ &\quad \times \exp(2i\omega t) + \text{c.c.} \end{aligned} \quad (\text{B24})$$

## B. Third order

At the third order Eqs. (1.1)–(1.3) give

$$\begin{aligned} \frac{\partial \rho_3}{\partial t} + \frac{1}{S} \frac{\partial}{\partial x} (S \rho_0 u_3) &= -\frac{\partial \rho_1}{\partial \theta} - \frac{\partial \rho_2}{\partial \tau} \\ &\quad - \frac{1}{S} \frac{\partial}{\partial x} (S \rho_1 u_2 + S \rho_2 u_1), \end{aligned} \quad (\text{B25})$$

$$\begin{aligned} \rho_0 \frac{\partial u_3}{\partial t} + \frac{\partial p_3}{\partial x} &= -\mathcal{D}_2 - \rho_0 \left( \frac{\partial u_1}{\partial \theta} + \frac{\partial u_2}{\partial \tau} + u_1 \frac{\partial u_2}{\partial x} + u_2 \frac{\partial u_1}{\partial x} \right) \\ &\quad - \rho_1 \left( \frac{\partial u_1}{\partial \tau} + \frac{\partial u_2}{\partial t} + u_1 \frac{\partial u_1}{\partial x} \right) - \rho_2 \frac{\partial u_1}{\partial t}, \end{aligned} \quad (\text{B26})$$

$$\begin{aligned} \frac{\partial p_3}{\partial t} + \frac{\gamma p_0}{S} \frac{\partial (S u_3)}{\partial x} &= (\gamma - 1)(\mathcal{T}_2 + u_1 \mathcal{D}_1) - \frac{\partial p_1}{\partial \theta} - \frac{\partial p_2}{\partial \tau} \\ &\quad - u_1 \frac{\partial p_2}{\partial x} - u_2 \frac{\partial p_1}{\partial x} - \frac{\gamma p_1}{S} \frac{\partial (S u_2)}{\partial x} \\ &\quad - \frac{\gamma p_2}{S} \frac{\partial (S u_1)}{\partial x}, \end{aligned} \quad (\text{B27})$$

from which (2.33) is obtained by eliminating  $\partial u_3 / \partial t$

between the second and third equations.

The equations for  $P_{311}(x)$ ,  $P_{313}(x)$ ,  $P_{32}(x)$ , and  $P_{30}(x)$  are obtained from Eq. (2.33) upon separating the different

frequency and amplitude components in the right-hand side;  $P_{20}$  is more easily determined from the momentum equation (B26). The results are

$$\begin{aligned} \frac{1}{S} \frac{d}{dx} \left( c_0^2 S \frac{dP_{311}}{dx} \right) + \omega^2 P_{311} = & -2\omega\Omega_2 P_1 - 2\omega\Omega_1 P_{21} - \Omega_1^2 P_1 - (\gamma - 1) \left[ \omega F_K (2\Omega_1 P_1 + \omega P_{21}) \right. \\ & \left. + c_p F_Q G_0 \left( \frac{dP_{21}}{dx} - F_D \frac{dP_1}{dx} \right) + c_p F_Q G_1 \frac{dP_1}{dx} \right] \\ & + \frac{1}{S} \frac{d}{dx} \left( S c_0^2 \left[ \left( F_D + \frac{\rho_{10}}{\rho_0} \right) \frac{dP_{21}}{dx} - \left( F_D^2 + F_D \frac{\rho_{10}}{\rho_0} + \frac{T_{w2}}{T_{w0}} \right) \frac{dP_1}{dx} \right] \right), \end{aligned} \quad (\text{B28})$$

$$\begin{aligned} \frac{1}{S} \frac{d}{dx} \left( c_0^2 S \frac{dP_{313}}{dx} \right) + \omega^2 P_{313} = & -2\omega\Lambda_2 P_1 + i\omega \left[ U_1 \left( 2 \frac{dP_{20}}{dx} - \frac{dP_{22}}{dx} \right) + U_{22} \frac{dP_1}{dx} + \frac{\gamma P_1}{S} \frac{d(SU_{22})}{dx} \right. \\ & \left. + \frac{\gamma}{S} (2P_{20} - P_{22}) \frac{d(SU_1)}{dx} \right] + \frac{\gamma P_0}{S} \frac{d}{dx} \left[ S \left( \frac{d(U_1 U_{22})}{dx} + \frac{R_1}{\rho_0} \left( U_1 \frac{dU_1}{dx} - 2i\omega U_{22} \right) \right. \right. \\ & \left. \left. - \left( \frac{R_{22}}{\rho_0} + 2 \frac{P_{20}}{p_0} \right) i\omega U_1 \right) \right], \end{aligned} \quad (\text{B29})$$

$$\begin{aligned} \frac{1}{S} \frac{d}{dx} \left( c_0^2 S \frac{dP_{32}}{dx} \right) + 4\omega^2 P_{32} = & -4\omega\Omega_1 P_{22} - (\gamma - 1) \left[ 4\omega^2 F_{2K} P_{22} - 2i\omega c_p \rho_0 F_{2Q} G_0 U_{22} + \frac{2}{\rho_0} F_D \left( \frac{dP_1}{dx} \right)^2 \right] \\ & - 2i\omega \frac{\gamma P_0}{S} \frac{d}{dx} (S F_{2D} U_{22}) + 2i\omega \left[ U_1 \frac{dP_{21}}{dx} + U_{21} \frac{dP_1}{dx} + \frac{\gamma P_1}{S} \frac{d(SU_{21})}{dx} + \frac{\gamma P_{21}}{S} \frac{d(SU_1)}{dx} \right] \\ & - \frac{1}{S} \frac{d}{dx} \left[ S c_0^2 \left( 2i\Omega_1 \rho_0 U_{22} + \rho_0 \frac{d(U_1 U_{21})}{dx} + i\omega R_1 U_{21} + 2i\omega \rho_{10} U_{22} + i\Omega_1 R_1 U_1 \right. \right. \\ & \left. \left. + \rho_{10} U_1 \frac{dU_1}{dx} + i\omega R_{21} U_1 \right) \right], \end{aligned} \quad (\text{B30})$$

$$\begin{aligned} \frac{1}{S} \frac{d}{dx} \left( c_0^2 S \frac{dP_{33}}{dx} \right) + 9\omega^2 P_{33} = & 3i\omega \left[ U_1 \frac{dP_{22}}{dx} + U_{22} \frac{dP_1}{dx} + \frac{\gamma P_1}{S} \frac{d(SU_{22})}{dx} + \frac{\gamma P_{22}}{S} \frac{d(SU_1)}{dx} \right] \\ & - \frac{1}{S} \frac{d}{dx} \left[ S c_0^2 \left( \rho_0 \frac{d(U_1 U_{22})}{dx} + 2i\omega R_1 U_{22} + R_1 U_1 \frac{dU_1}{dx} + i\omega R_{22} U_1 \right) \right], \end{aligned} \quad (\text{B31})$$

$$\frac{dP_{30}}{dx} = \rho_0 \frac{d(U_1 U_{21})}{dx} - i\omega R_1 U_{21} - i\Omega_1 R_1 U_1 + \rho_{10} U_1 \frac{dU_1}{dx} - i\omega R_{21} U_1. \quad (\text{B32})$$

$P_{30}(x)$  is complex in general, but we only need its real part  $\text{Re } P_{30}(x)$  as will be seen later. The solution for  $\text{Re } P_{30}(x)$  is made unique by means of the additional condition

$$\int_0^L \frac{S}{T_{w0}} \left( \text{Re } P_{30} + P_{20} \frac{\rho_{10}}{\rho_0} \right) dx = 0, \quad (\text{B33})$$

that follows from the conservation of mass in the tube as in (B10).

The velocity and density fields have expressions similar to (2.40) with the spatial parts given by

$$\begin{aligned} U_{311} = & \frac{i}{\rho_0 \omega} \frac{dP_{311}}{dx} - \left( \frac{\Omega_2}{\omega} + \frac{\Omega_1}{\omega} \frac{\rho_{10}}{\rho_0} + \frac{\rho_{200}}{\rho_0} \right) U_1 \\ & - \left( \frac{\Omega_1}{\omega} + \frac{\rho_{10}}{\rho_0} \right) U_{21} - \frac{iF_D}{\rho_0 \omega} \left( \frac{dP_{21}}{dx} - F_D \frac{dP_1}{dx} \right), \end{aligned} \quad (\text{B34})$$

$$\begin{aligned} U_{313} = & \frac{i}{\rho_0 \omega} \frac{dP_{313}}{dx} - \left( \frac{\Lambda_2}{\omega} + 2 \frac{P_{20}}{\rho_0} + \frac{R_{22}}{\rho_0} \right) U_1 \\ & - \frac{i}{\omega} \frac{d(U_1 U_{22})}{dx} - 2 \frac{R_1}{\rho_0} U_{22} - \frac{i}{\omega} \frac{R_1}{\rho_0} U_1 \frac{dU_1}{dx}, \end{aligned} \quad (\text{B35})$$

$$U_{32} = \frac{i}{2\omega\rho_0} \frac{dP_{32}}{dx} - \left( \frac{\Omega_1}{\omega} + \frac{\rho_{10}}{\rho_0} + F_D \right) U_{22} - \frac{1}{2} \frac{R_1}{\rho_0} U_{21} - \frac{1}{2} \left( \frac{\Omega_1}{\omega} \frac{R_1}{\rho_0} + \frac{R_{21}}{\rho_0} \right) U_1 + \frac{i}{2\omega} \frac{d(U_1 U_{21})}{dx} + \frac{i}{2\omega} \frac{\rho_{10}}{\rho_0} U_1 \frac{dU_1}{dx}, \quad (\text{B36})$$

$$U_{30} = -\frac{R_1}{\rho_0} \text{Re } U_{21} + i \frac{U_1}{\rho_0} \text{Im } R_{21}, \quad (\text{B37})$$

This result for  $U_{30}$  follows from the condition that the time-independent part of the flow in the tube vanishes at every order. Furthermore

$$R_{311} = -\frac{\Omega_1}{\omega} R_{21} - \frac{\Omega_2}{\omega} R_1 + \frac{i}{\omega S} \frac{d}{dx} \times [S(\rho_0 U_{311} + \rho_{10} U_{21} + \rho_{200} U_1)], \quad (\text{B38})$$

$$R_{313} = -\frac{\Lambda_2}{\omega} R_1 + \frac{i}{\omega S} \frac{d}{dx} [S(\rho_0 U_{313} + R_1 U_{22} + (\rho_{202} - R_{22}) U_1)], \quad (\text{B39})$$

$$R_{32} = -\frac{\Omega_1}{\omega} R_{22} + \frac{i}{2\omega S} \frac{d}{dx} [S(\rho_0 U_{32} + \rho_{10} U_{22} + R_1 U_{21} + R_{21} U_1)]. \quad (\text{B40})$$

Again,

$$\rho_3 = \rho_{30} + \rho_{31} + \rho_{32} + \rho_{33}, \quad (\text{B41})$$

and

$$\rho_{30} = \rho_{300} + A^* A \rho_{302}, \quad (\text{B42})$$

where

$$\rho_{300} = \rho_{10} \left( \frac{\rho_{200}}{\rho_0} - \frac{T_{w2}}{T_{w0}} \right), \quad \rho_{302} = \rho_0 \frac{P_{30}}{P_0} + \rho_{10} \frac{P_{20}}{P_0} + \text{c.c.} \quad (\text{B43})$$

For  $\mathcal{T}_{31}$  and  $\mathcal{D}_{31}$  we have, respectively,

$$\mathcal{D}_{31} = i\omega F_D \left\{ \rho_0 U_{311} + \left( \rho_{10} + \rho_0 \frac{\Omega_1}{\omega} \right) U_{21} + \left( \rho_{200} + \rho_{10} \frac{\Omega_1}{\omega} + \rho_0 \frac{\Omega_2}{\omega} \right) U_1 + A^* A \left[ \rho_0 U_{313} + \left( \rho_{202} + \rho_0 \frac{\Lambda_2}{\omega} \right) U_1 \right] \right\} A(\tau, \theta, \eta) \exp(i\omega t) + \text{c.c.}, \quad (\text{B44})$$

$$\mathcal{T}_{31} = -[iF_K(\omega P_{311} + \Omega_1 P_{21} + \Omega_2 P_1) + c_p F_Q(G_0(\rho_0 U_{311} + \rho_{10} U_{21} + \rho_{200} U_1) + G_1(\rho_0 U_{21} + \rho_{10} U_1) + \rho_0 G_2 U_1) + A^* A \{iF_K(\omega P_{313} + \Lambda_2 P_1) + c_p F_Q G_0(\rho_0 U_{313} + \rho_{202} U_1)\}] A(\tau, \theta, \eta) \exp(i\omega t) + \text{c.c.} \quad (\text{B45})$$

## C. Fourth order

At fourth order Eqs. (1.2)–(1.3) give

$$\rho_0 \frac{\partial u_4}{\partial t} + \frac{\partial p_4}{\partial x} = -\mathcal{D}_3 - \rho_0 \left( \frac{\partial u_1}{\partial \eta} + \frac{\partial u_2}{\partial \theta} + \frac{\partial u_3}{\partial \tau} + u_1 \frac{\partial u_3}{\partial x} + u_2 \frac{\partial u_2}{\partial x} + u_3 \frac{\partial u_1}{\partial x} \right) - \rho_1 \left( \frac{\partial u_1}{\partial \theta} + \frac{\partial u_2}{\partial \tau} + \frac{\partial u_3}{\partial t} + u_1 \frac{\partial u_2}{\partial x} + u_2 \frac{\partial u_1}{\partial x} \right) - \rho_2 \left( \frac{\partial u_2}{\partial t} + \frac{\partial u_1}{\partial \tau} + u_1 \frac{\partial u_1}{\partial x} \right) - \rho_3 \frac{\partial u_1}{\partial t}, \quad (\text{B46})$$

$$\frac{\partial p_4}{\partial t} + \frac{\gamma P_0}{S} \frac{\partial(Su_4)}{\partial x} = (\gamma - 1)(\mathcal{T}_3 + u_1 \mathcal{D}_2 + u_2 \mathcal{D}_1) - \frac{\partial p_1}{\partial \eta} - \frac{\partial p_2}{\partial \theta} - \frac{\partial p_3}{\partial \tau} - u_1 \frac{\partial p_3}{\partial x} - u_2 \frac{\partial p_2}{\partial x} - u_3 \frac{\partial p_1}{\partial x} - \frac{\gamma}{S} \left[ p_1 \frac{\partial(Su_3)}{\partial x} + p_2 \frac{\partial(Su_2)}{\partial x} + p_3 \frac{\partial(Su_1)}{\partial x} \right]. \quad (\text{B47})$$

Upon eliminating  $\partial u_4 / \partial t$  between the second and the third one we find (2.41).

<sup>1</sup>More precisely, one could determine a value of  $T_{w2}^*$  corresponding to marginal stability, after which the difference  $T_w(x) - [T_{w0}(x) + \epsilon T_{w1}(x) + \epsilon^2 T_{w2}^*]$  would be the driving force for the instability.

<sup>2</sup>Since the pressure perturbation is the sum of two complex conjugate quantities, the dotted line actually shows  $2|\text{Re } \bar{A}_{\text{sat}}|$ .

Akhavan, R., Kamm, R. D., and Shapiro, A. H. (1991). "An investigation of transition to turbulence in bounded oscillatory Stokes flows. I. Experiment," *J. Fluid Mech.* **225**, 395–422.

Atchley, A. A., Bass, H. E., and Hofer, T. J. (1990). "Development of nonlinear waves in a thermoacoustic prime mover," in *Frontiers in Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Elsevier, New York), pp. 603–608.

Brewster, J. R., Raspet, R., and Bass, H. E. (1997). "Temperature discontinuities between elements of thermoacoustic devices," *J. Acoust. Soc. Am.* **102**, 3355–3360.

Cao, N., Olson, J. R., Swift, G. W., and Chen, S. (1996). "Energy flux density in a thermoacoustic couple," *J. Acoust. Soc. Am.* **99**, 3456–3464.

Cooper, W. L., Yang, K. T., and Nee, V. W. (1993). "Fluid mechanics of oscillatory modulated flows and associated applications in heat and mass transfer—A review," *J. Energy Heat Mass Transfer* **15**, 1–19.

Gopinath, A., Tait, N. L., and Garrett, S. L. (1998). "Thermoacoustic streaming in a resonant channel: The time-averaged temperature distribution," *J. Acoust. Soc. Am.* **103**, 1388–1405.

Hinch, E. J. (1991). *Perturbation Methods* (Cambridge University Press, Cambridge).

Karpov, S. (2000). "Nonlinear Phenomena in Thermoacoustics and Bubbly Liquids," Ph.D. thesis, Johns Hopkins University.

Karpov, S., and Prosperetti, A. (1998). "Linear thermoacoustic instability in the time domain," *J. Acoust. Soc. Am.* **103**, 3309–3317.

Kevorkian, J., and Cole, J. D. (1996). *Perturbation Methods in Applied Mathematics*, 2nd ed. (Springer, New York).

Morse, P. M., and Feshbach, H. (1953). *Methods of Theoretical Physics* (McGraw-Hill, New York).

Murdock, J. A. (1991). *Perturbations* (Wiley, New York).

Naylor, A. W., and Sell, G. R. (1982). *Linear Operator Theory in Engineering and Science* (Springer, New York), p. 502.

Rott, N. (1969). "Damped and thermally driven acoustic oscillations in wide and narrow tubes," *Z. Angew. Math. Phys.* **20**, 230–243.

- Rott, N. (1976). "Thermally driven acoustic oscillations. IV. Tubes with variable cross section," *Z. Angew. Math. Phys.* **27**, 197–224.
- Rott, N. (1980). "Thermoacoustics," *Adv. Appl. Mech.* **20**, 135–175.
- Rott, N. (1983). "Thermally driven acoustic oscillations, VI. Excitation and power," *Z. Angew. Math. Phys.* **34**, 609–626.
- Swift, G. W. (1988). "Thermoacoustic engines," *J. Acoust. Soc. Am.* **84**, 1145–1180.
- Watanabe, M., Prosperetti, A., and Yuan, H. (1997). "A simplified model for linear and nonlinear processes in thermoacoustic prime movers. I. Model and linear theory," *J. Acoust. Soc. Am.* **102**, 3484–3496.
- Wheatley, J. (1986). "Intrinsically irreversible or natural heat engines," in *Frontiers in Physical Acoustics*, edited by D. Sette (North-Holland, Amsterdam), pp. 35–475.
- Worlikar, A. S., and Knio, O. M. (1996). "Numerical simulation of a thermoacoustic refrigerator. I. Unsteady adiabatic flow around the stack," *J. Comput. Phys.* **127**, 424–451.
- Worlikar, A. S., and Knio, O. M. (1999). "Numerical study of oscillatory flow and heat transfer in a loaded thermoacoustic stack," *Numer. Heat Transfer A35*, 49–65.
- Worlikar, A. S., Knio, O. M., and Klein, R. (1998). "Numerical simulation of a thermoacoustic refrigerator. II. Stratified flow around the stack," *J. Comput. Phys.* **144**, 299–324.
- Yuan, H., Karpov, S., and Prosperetti, A. (1997). "A simplified model for linear and nonlinear processes in thermoacoustic prime movers. II. Nonlinear oscillations," *J. Acoust. Soc. Am.* **102**, 3497–3506.

# A thermoacoustic-Stirling heat engine: Detailed study

S. Backhaus and G. W. Swift

*Condensed Matter and Thermal Physics Group, Los Alamos National Laboratory, Los Alamos,  
New Mexico 87545*

(Received 29 May 1999; accepted for publication 16 February 2000)

A new type of thermoacoustic engine based on traveling waves and ideally reversible heat transfer is described. Measurements and analysis of its performance are presented. This new engine outperforms previous thermoacoustic engines, which are based on standing waves and intrinsically irreversible heat transfer, by more than 50%. At its most efficient operating point, it delivers 710 W of acoustic power to its resonator with a thermal efficiency of 0.30, corresponding to 41% of the Carnot efficiency. At its most powerful operating point, it delivers 890 W to its resonator with a thermal efficiency of 0.22. The efficiency of this engine can be degraded by two types of acoustic streaming. These are suppressed by appropriate tapering of crucial surfaces in the engine and by using additional nonlinearity to induce an opposing time-averaged pressure difference. Data are presented which show the nearly complete elimination of the streaming convective heat loads. Analysis of these and other irreversibilities show which components of the engine require further research to achieve higher efficiency. Additionally, these data show that the dynamics and acoustic power flows are well understood, but the details of the streaming suppression and associated heat convection are only qualitatively understood. © 2000 Acoustical Society of America.

[S0001-4966(00)00206-X]

PACS numbers: 43.35.Ud [HEB]

## INTRODUCTION

Since their invention in 1815, Stirling engines and refrigerators have shown great promise as commercial heat engines.<sup>1</sup> Much of this promise is due to the Stirling cycle's inherent reversibility, and therefore, its high thermal efficiency. Despite this promise, the drawback of Stirling engines has always been the need for multiple sliding seals capable of withstanding high pressure and many cycles without failure. Recent advances in Stirling technology include free-piston machines, the use of linear alternators and motors, and orifice pulse tube refrigerators.<sup>2</sup> Unfortunately, these advances did not completely eliminate the unreliability and high cost of sliding seals.

The Fluidyne engine<sup>3</sup> was the first attempt to totally eliminate sliding seals, by using U-tube liquid pistons. Unfortunately, this solution is limited to low frequency by the high mass of the liquid pistons. Ceperley also attempted to totally eliminate sliding seals.<sup>4-6</sup> He realized that a traveling acoustic wave propagating through a regenerative heat exchanger (regenerator) undergoes a thermodynamic cycle similar to the Stirling cycle. As the wave travels up the temperature gradient through the regenerator (i.e., from ambient temperature toward high temperature), it carries the gas in the regenerator through a sequence of displacement toward higher temperature, depressurization, displacement toward lower temperature, and pressurization. The gas experiences thermal expansion during the displacement toward higher temperature and thermal contraction during the displacement toward lower temperature, so that acoustic power is amplified as it travels toward higher temperature; this is the source of the engine's net work. Simultaneously, the gas gains entropy from the regenerator solid during the depressurization and loses entropy to the solid during pressurization, so that the two displacement processes result in a net convection of

entropy down the temperature gradient; this is the mechanism by which the engine extracts heat from the high-temperature heat source and rejects heat to the ambient-temperature heat sink.

Ceperley's experimental engine was not able to amplify acoustic power. Much later, Yazaki *et al.*<sup>5</sup> first demonstrated such an engine, but at low efficiency. Ceperley and Yazaki realized that this was due to the low acoustic impedance of the working gas, which caused large viscous losses resulting from high acoustic velocities.<sup>6</sup> Also, they did not anticipate the presence of several forms of acoustic streaming, including "Gedeon streaming" which occurs whenever a traveling-wave component is present<sup>7</sup> and "Rayleigh streaming" which is due to boundary-layer effects in the gas space that provides the thermal buffer for the hot heat exchanger.<sup>8</sup> Although Gedeon streaming has been detected experimentally,<sup>9,10</sup> we have found no published description of a systematic method to suppress it.

Another engine without sliding seals is a standing-wave thermoacoustic engine.<sup>11</sup> Many engines of this type have been built, but because they are based on an intrinsically irreversible thermodynamic cycle their thermal efficiency (ratio of acoustic power produced to heater power consumed) has thus far been limited to 0.20.<sup>12</sup>

This paper describes a Stirling-cycle based thermoacoustic engine, which was introduced briefly in an earlier paper.<sup>13</sup> Since it employs the inherently reversible Stirling cycle, this engine's thermal efficiency can easily surpass that of standing-wave engines. To create the traveling-wave phasing and high acoustic impedance in the regenerator desired by Ceperley,<sup>6</sup> the engine uses a compact acoustic network. Unchecked, both Gedeon and Rayleigh streaming would seriously degrade the thermal efficiency of the engine. Gedeon streaming is suppressed by using hydrodynamic end effects

which produce an opposing time-averaged pressure difference. Rayleigh streaming is suppressed through the use of an appropriately tapered thermal buffer tube<sup>8</sup> (referred to as a pulse tube in the context of orifice pulse tube refrigerators). With the heat loads due to both types of streaming reduced to an acceptable level, this engine has demonstrated a thermal efficiency of 0.30 which is more than 50% higher than any previous thermoacoustic engine. For comparison, typical thermal efficiencies of internal combustion engines are 0.25 for gasoline and 0.40 for diesel.<sup>14</sup>

The next section describes the engine and its construction in detail. The following section presents two simplified models of the engine which provide some intuition about the creation of traveling-wave phasing with high acoustic impedance and how various components of the engine affect these properties. Then, preliminary measurements are shown which demonstrate our good understanding of the acoustic network and our qualitative understanding of the effects and suppression of acoustic streaming. Finally, data on the performance of the engine and on the various loss mechanisms are presented. These data demonstrate a quantitative understanding of the dynamics and acoustic power flows and a qualitative understanding of the acoustic streaming and associated heat convection.

## I. APPARATUS

Scale drawings of the apparatus used in these measurements are shown in Fig. 1(a) and (b). Essentially, it is composed of a 1/4-wavelength resonator filled with 30-bar helium. The torus-shaped section contains the heat exchangers, regenerator and other duct work necessary to force the helium to execute the Stirling cycle.<sup>15</sup> The rest of the hardware past the resonator junction forms the resonator and variable acoustic load.<sup>16</sup>

Near the top of the torus is the main cold heat exchanger. It is of shell-and-tube construction consisting of 299 2.5-mm inside-diameter, 20-mm-long stainless-steel tubes welded into two 1.6-mm-thick stainless-steel plates. The wall thickness of the tubes is 0.7 mm. The diameter of the heat exchanger, measured from the outer edge of the tubes on the periphery, is 9.5 cm. The tubes, which are parallel to the acoustic displacement, carry the helium gas. They are cooled by chilled water ( $\approx 15^\circ\text{C}$ ) passing through the shell. At high amplitude, the peak-to-peak gas displacement is roughly equal to the length of the tubes and the typical Reynolds number in the tubes is 3000.

Below the main cold heat exchanger is the regenerator, made from a 7.3-cm-tall stack of 120-mesh stainless-steel screen machined to a diameter of 8.89 cm. The diameter of the screen wire is  $65\ \mu\text{m}$ . The randomly stacked screens are contained within a thin-wall stainless-steel can for ease of installation and removal. On each end, three 3.2-mm-thick stainless-steel ribs are welded to the can to hold the screen axially and to provide spacing between the regenerator and adjacent heat exchangers. At the upper end, this spacing allows the flow to spread over the entire regenerator cross section after exiting the narrow passages of the main cold heat exchanger. Based on the total weight of screen in the

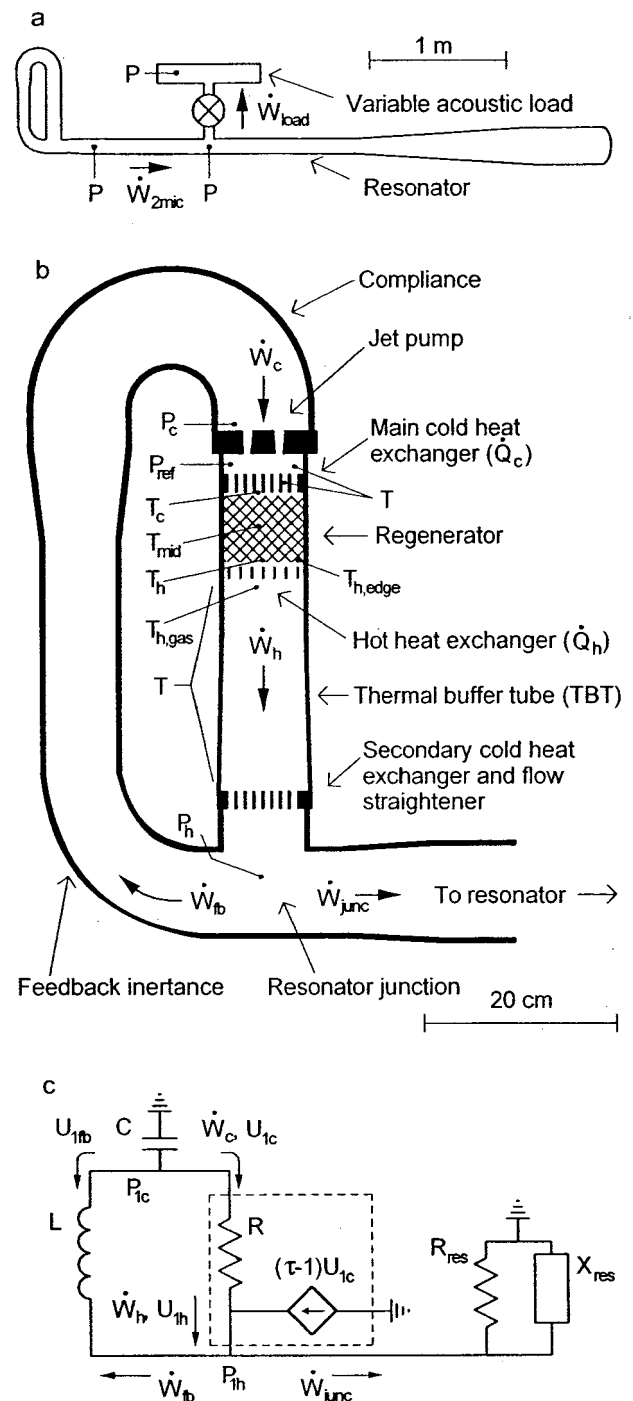


FIG. 1. (a) Scale drawing of the engine, resonator and variable acoustic load (load is not to scale). The 1/4-wavelength resonator is filled with 30-bar helium. The pressure antinode is located near the top of the torus section, and the node is near the inlet to the large cone. The variable acoustic load consists of an adjustable globe valve attached to a 2.2 l tank. (b) Scale drawing of the torus section. Locations labeled with T or P indicate where temperature and pressure are measured. The pressure measurement location labeled  $P_{ref}$  is the reference location for measurements of acoustic amplitude in the engine. (c) Lumped-element model of the engine. All elements in the torus are much shorter than 1/4 wavelength. Therefore, they can be modeled as lumped elements. Subscripts in (c) correspond to the locations in (b). The regenerator and adjacent heat exchangers are modeled by the elements inside the dashed box. See text for description of the model.

regenerator, we calculate its volume porosity  $\phi \approx 0.72$  and the hydraulic radius  $r_h \approx 42\ \mu\text{m}$ .<sup>15</sup> This is smaller than the helium's thermal penetration depth  $\delta_\kappa$ , which varies from  $140\ \mu\text{m}$  to  $460\ \mu\text{m}$  through the regenerator. The pressure



vessel around the regenerator, which is made from Inconel 625 (ASTM B446-Grade 2), has a wall thickness of 4.0 mm. This alloy was chosen for its high tensile strength at elevated temperatures as well as its comparatively low thermal conductivity;<sup>17</sup> the heat leak carried by the pressure-vessel wall would have been about a factor of 3 higher if we had used 316H stainless steel.

The lower end of the regenerator abuts the hot heat exchanger. It consists of a 0.64-cm-wide by 3.5-m-long Ni-Cr ribbon wound zigzag on an alumina frame. The ribbon is divided into 3 equal-length segments of 4.0  $\Omega$  each and driven with 208-V three-phase power in a delta configuration. Electrical leads enter perpendicular to the pressure-vessel axis just below the hot heat exchanger. This location is chosen so as to not disturb the flow in the thermal buffer tube. Power flowing into the hot heat exchanger is measured using a commercial wattmeter with a 0.2% accuracy.<sup>18</sup> Up to 5.5 kW is supplied to the heat exchanger with the surrounding gas at 725 °C.

The thermal buffer tube (TBT) is a tapered, 24-cm-long open cylinder made of the same Inconel 625. It provides a thermal buffer between the hot heat exchanger and room temperature. Its inner diameter is much larger than the helium's thermal penetration depth, and its length is much greater than the peak-to-peak gas displacement (4 cm) at high amplitudes. The inside surface is polished to a 1- $\mu\text{m}$  finish to ensure that the surface roughness is much less than the viscous and thermal penetration depths. The wall thickness is 4.0 mm for the entire length. The upper 8.0 cm of the TBT is a straight cylinder while the lower 16.0 cm is flared with a 1.35° half-angle taper. This shape is used to minimize boundary-layer driven streaming (Rayleigh streaming) in the TBT.<sup>8</sup>

At the lower end of the TBT is a flow straightener made of five layers of 22-mesh copper screen with a wire diameter of 0.4 mm. It ensures that the flow entering the bottom of the TBT is spatially uniform, not a jet flow due either to the geometry of the secondary cold heat exchanger or to flow separation at the resonator junction. A jet flow would cause streaming within the TBT to convect heat from the hot heat exchanger. Before the flow straightener was installed, up to 1 kW of heat was convected away in this fashion.

Next, a secondary, water-cooled, shell-and-tube heat exchanger anchors the lower end of the TBT at room temperature. It contains 109 4.6-mm inside-diameter, 10-mm-long stainless-steel tubes welded into two 1.6-mm-thick stainless-steel plates. If both forms of streaming are suppressed, this heat exchanger only intercepts the leak down the TBT wall and radiation from the hot heat exchanger, so it can be made from short, large-diameter tubes.

Several 1.6-mm-diameter type-K thermocouples allow the measurement of temperatures throughout the Stirling section of the engine. Starting at the top of Fig. 1(b), two thermocouples are located near the main cold heat exchanger. One is freely suspended within one of the tubes and the other is located in the helium, 1 cm above the heat exchanger assembly. Also, there are two thermocouples (not shown in the figure) in the inlet and outlet chilled water streams supplying this heat exchanger. These four thermocouples allow

us to assess the performance of this heat exchanger by measuring the heat load on the heat exchanger (via the temperature rise of the chilled water) and the temperature defect between the helium and the chilled water. Next are four thermocouples in contact with the regenerator screen bed. The first extends through a tube of the main cold heat exchanger and is pressed against the center of the regenerator's cold end. The other three pass through the pressure-vessel wall and are inserted into tight-fitting pockets drilled into the regenerator. Of these three, two are centered radially with one near the hot end and the other centered axially. The last thermocouple is near the hot end, but its pocket is only drilled 1 cm into the regenerator. The three radially centered thermocouples allow the measurement of the axial temperature profile within the regenerator. This profile is important in detecting and eliminating Gedeon streaming. The two thermocouples near the hot end allow measurement of the radial temperature profile. We believe that vigorous Rayleigh or jet-driven streaming in the TBT causes unequal heating of the hot face of the regenerator. In a previous version of this engine in which there was no attempt to suppress either type of streaming, up to 100 °C differences were measured between the center and edge of the regenerator's hot face. In the present engine, up to 90 °C differences were measured before the flow straightener was installed; a maximum of 30 °C was measured after the installation. Finally, there are eight thermocouples held against the outer wall of the TBT by set screws, spaced every 2.5 cm. After the bulk of the Gedeon streaming is suppressed, the axial temperature profile measured by these thermocouples is used to detect the presence of either Rayleigh or jet-driven streaming.

Clockwise farther around the torus are the resonator junction, feedback inertance, and compliance. The junction is a standard-wall, 3 1/2-in. nominal, stainless-steel tee. The inside diameter is roughly 9.0 cm. The feedback inertance consists of three separate sections of piping. The first section is a 3 1/2-in. to 3-in. nominal, long-radius reducing elbow. The centerline length of the elbow is 20.9 cm, and the final inside diameter of the elbow is 7.8 cm. All burrs and irregularities on the inside of the elbow have been removed by grinding and filing. The next section is a 3-in. nominal, stainless-steel pipe 25.6-cm long. The inside diameter of this pipe is honed to 2- $\mu\text{m}$  finish, to minimize viscous loss at high amplitudes. This segment of pipe includes a pressure-balanced sliding joint (not shown) which allows for the thermal expansion of the TBT and regenerator case. When the hot temperature rises from 25 °C to 725 °C, these components expand roughly 3 mm; without the sliding joint, the resulting stress would cause the hot components to distort. The last section of the inertance is a machined cone that adapts the 3-in. nominal pipe to the compliance. Its initial and final inside diameters are 7.8 cm and 10.2 cm, and its length is 10.2 cm, giving a taper angle of 13.5°. The gradual cross-sectional area changes, gentle bends, and smooth surfaces are used to minimize dissipation due to flow separation<sup>19</sup> and viscosity in the feedback inertance. The compliance consists of two 4-in. nominal, short-radius 90° elbows made from carbon steel. The inner surface of the elbows is sandblasted to remove a protective layer of paint.

Although this treatment left the surface quite rough, the acoustic velocities in the compliance are low, so the rough surface should not affect the dissipation significantly. The internal volume of the compliance is 2830 cm<sup>3</sup>, measured by filling it with a known volume of water.

Between the compliance and the main cold heat exchanger is a device we have termed a “jet pump,”<sup>20</sup> which stops Gedeon streaming. We will reserve discussion of this component until the section on streaming.

The resonator, which joins with the torus at the junction between the TBT and feedback duct, consists of three sections. The first section is a machined cone that adapts the 3 1/2-in. nominal tee to a 4-in. nominal, carbon-steel pipe. The initial and final inside diameters of the cone are 9.0 cm and 10.2 cm, and the length is 10.2 cm resulting in a 6.8° taper angle. The small angle is used to prevent flow separation at the walls of the cone.<sup>19</sup> The main section of the resonator is a 1.90-m length of 4-in. nominal, carbon-steel pipe. The inside diameter of the pipe is 10.2 cm and is honed to a 2- $\mu$ m surface finish. The resonator diameter is enlarged over the engine diameter to reduce the acoustic velocities near the velocity antinode at the end of the main resonator section. The increase in surface area tends to increase the dissipation, but the reduction in velocity is more important at high amplitude where the acoustic losses grow as the cube of the amplitude.<sup>21</sup> The final section of the resonator includes a 7° cone which enlarges the inside diameter of the resonator from 10.2 cm to 25.5 cm over a length of 1.22 m. The large end of the cone is closed with a 25.5-cm-diameter pipe with an approximate length of 52 cm, terminating in a 2:1 ellipsoidal cap. The main section of the resonator has three ports of 1.6-cm diameter. Two of these are located near the ends and the third is in the center. A variable, RC-type acoustic load<sup>16</sup> can be attached to either of the two ports nearest the resonator junction. The fill line is connected to the port nearest the 7° cone. Two piezoresistive pressure sensors<sup>22</sup> located in the resonator and a third in the variable acoustic load allow two-microphone power measurements of  $\dot{W}_{2mic}$  and  $\dot{W}_{load}$ .<sup>16</sup> All acoustic pressure measurements are made with a lock-in amplifier using the measured acoustic pressure immediately above the main cold heat exchanger as a reference signal.

In addition to the pressure sensors in the resonator, three additional pressure sensors are located in the torus: between the compliance and jet pump, between the jet pump and main cold heat exchanger, and at the resonator junction. Additionally, a differential pressure sensor is placed across the jet pump. A Bourdon-tube gauge connected through a fine capillary is used to measure the mean pressure in the engine.

## II. SIMPLIFIED MODELS

Figure 1(c) shows the simplest lumped-element model of the engine that captures all the most important features. Each component of the engine is replaced with its idealized lumped-element equivalent.<sup>23</sup> The resonator is modeled as resistive and reactive impedances in parallel,  $R_{res}$  and  $iX_{res}$ , respectively.  $R_{res}$  is controlled by the resonator dissipation and the variable acoustic load. At this point, no distinction is

made between these two mechanisms. Under the normal operating conditions  $X_{res} \ll R_{res}$ . The impedances of the regenerator  $R$  and feedback inertance  $L$  are much smaller than that of the compliance  $C$ . Therefore,  $\omega$  is determined primarily by the resonance condition between  $X_{res}$  and the compliance: i.e.,  $X_{res}(\omega) \approx 1/\omega C$ .

We use the conventional counterclockwise phasor notation,<sup>11</sup> so that time-dependent variables are expressed as

$$\xi(t) = \xi_m + \text{Re}[\xi_1 e^{i\omega t}], \quad (1)$$

with the mean value  $\xi_m$  real, and with  $\xi_1$  complex to account for both the magnitude and phase of the oscillation at angular frequency  $\omega$ . The acoustic pressure and volumetric velocity (i.e., volume flow rate) are represented by  $p_{1,x}$  and  $U_{1,x}$  at the points in Fig. 1(b) and (c) labeled with the subscript  $x$ . For example,  $p_{1,h}$  is the complex amplitude of the acoustic pressure oscillation at the location labeled with a subscript  $h$  in Fig. 1(b), between the regenerator and hot heat exchanger.

Following Ceperley,<sup>4</sup> we model the regenerator and adjacent heat exchangers by the resistance and volumetric-velocity source inside the dashed box. The resistance is due to viscous flow in the tightly-spaced screen bed that forms the regenerator.<sup>24</sup> The volumetric-velocity source is due to the temperature profile. The mean temperature  $T_m$  changes axially through the regenerator, and hence the mean density  $\rho_m$  changes according to  $\rho_m \propto 1/T_m$ . To enforce the conservation of first-order mass flux,  $\rho_m U_1 = \text{const}$  throughout an ideal regenerator with negligible internal compliance, the volumetric velocity  $U_1$  must grow as  $T_m$ . Therefore, the volumetric velocity  $U_{1,h}$  at the hot end of the regenerator would ideally be  $\tau U_{1,c}$ , where  $\tau = T_h/T_c$  is the ratio of absolute temperatures at the regenerator ends. The volumetric-velocity source is set to  $(\tau - 1)U_{1,c}$  to enforce this condition.

The study of this simple model provides some basic insights into the operation of the engine. We have  $i\omega L U_{1,fb} = U_{1,c} R$  and  $U_{1,c} + U_{1,fb} = -i\omega C p_{1,c}$ . Eliminating  $U_{1,fb}$  and solving for  $U_{1,c}$ , we find

$$U_{1,c} = \frac{\omega^2 LC}{R} \frac{p_{1,c}}{1 + i\omega L/R}. \quad (2)$$

If the magnitude of the impedance of the feedback inertance  $\omega L$  is small compared with the resistance  $R$  of the regenerator,  $p_{1,c}$  and  $U_{1,c}$  are nearly in phase corresponding to the traveling-wave phasing necessary for efficient regenerator operation.<sup>4</sup> The magnitude of  $U_{1,c}$  is controlled by the magnitudes of all three impedances:  $R$ ,  $\omega L$  and  $1/\omega C$ . The magnitude of the volumetric velocity passing through the parallel combination of the feedback inertance and the regenerator is given by the magnitude of the volumetric velocity into the compliance, or  $\omega C p_{1,c}$ . As the compliance becomes larger, this volumetric velocity grows as does  $U_{1,c}$ . Since  $U_{1,c}/U_{1,fb} = i\omega L/R$ , the volumetric velocity in the feedback inertance is always 90° out of phase from that in the regenerator, with increasing volumetric velocity through the regenerator as the impedance of the inertance increases relative to the regenerator resistance. The ability to tune the magni-

tude of the volumetric velocity in the regenerator is equivalent to tuning its power output. This degree of freedom is useful in the design process when trading off losses in the regenerator against resonator and feedback loss during optimization of the entire system.

Next, it is useful to look at the time-averaged acoustic power flows  $\dot{W} = \text{Re}[p_1 \tilde{U}_1]/2$  at various points in the engine. Using Eq. (2), the time-averaged power flow into the regenerator is given by

$$\dot{W}_c \approx \frac{|p_{1,c}|^2}{2} \frac{\omega^2 LC}{R}, \quad (3)$$

where we have dropped terms of order  $(\omega L/R)^2$ . Neither the compliance nor inertance absorb or generate time-averaged power. This implies that  $\dot{W}_c$  is the power fed back through the inertance, i.e.,  $\dot{W}_c = \dot{W}_{fb}$ . Since the impedance of the parallel combination of the regenerator and feedback inertance is small compared to that of the compliance or resonator,  $p_{1,h} \approx p_{1,c}$ . Therefore, the time-averaged power flowing out of the hot end is then given by

$$\dot{W}_h \approx \tau \dot{W}_c. \quad (4)$$

By conserving the time-averaged power flow at the resonator junction, the resonator is found to absorb an amount of power given by

$$\dot{W}_{\text{junc}} \approx (\tau - 1) \dot{W}_c. \quad (5)$$

To determine  $\tau$ ,  $\dot{W}_{\text{junc}}$  is expressed in terms of  $R_{\text{res}}$  and substituted into Eq. (5). Using Eq. (3),  $\tau$  is given by

$$\tau \approx 1 + \frac{R}{\omega^2 LCR_{\text{res}}}. \quad (6)$$

In an ideal regenerator that uses an ideal gas,  $\dot{W}_h = \dot{Q}_h$  where  $\dot{Q}_h$  is the heat input from the hot heat exchanger.<sup>25</sup> Substituting Eqs. (3) and (4),  $|p_{1,c}|^2$  is found to be

$$|p_{1,c}|^2 \approx \frac{2R}{\tau \omega^2 LC} \dot{Q}_h. \quad (7)$$

Equations (6) and (7) show how the operating point of the engine is set by the external controls. Setting  $R_{\text{res}}$ , by adjusting the variable acoustic load, sets  $T_h$  by fixing  $\tau$ . With  $T_h$  determined, the acoustic amplitude is set by adjusting the heat input  $\dot{Q}_h$ .

At fixed acoustic amplitude, Eq. (3) shows that power input to the regenerator only depends on  $\omega$  and the geometry of the inertance, compliance, and regenerator. Therefore, for a simple fixed-amplitude point of view, a constant power  $\dot{W}_c$  circulates around the torus. This power is amplified by the temperature gradient in the regenerator, and the added power  $(\tau - 1) \dot{W}_c$  is delivered to the resonator. If  $\dot{W}_{\text{junc}}$  is increased at *fixed acoustic amplitude*, the ‘‘power gain’’ of the regenerator, i.e.,  $\tau$ , must grow.

The results of this section, in particular Eq. (2), separate this engine from that originally envisioned by Ceperley.<sup>4</sup> A figure of merit for a regenerator in a traveling-wave engine is  $\beta = \dot{W}_c / \dot{E}_{\text{vis}}$  where  $\dot{E}_{\text{vis}}$  is the time-average power dissipated in the regenerator due to viscosity. Using  $\dot{E}_{\text{vis}} \approx R |U_{1,c}|^2/2$

and  $\dot{W}_c = |p_{1,c}| |U_{1,c}|/2$  (assuming nearly traveling-wave phasing), this figure of merit can be rewritten  $\beta \approx |Z|/R$  where  $Z$  is the acoustic impedance in the regenerator. If a pure traveling wave is used, i.e., one with  $Z = \rho_m c/A$  where  $c$  is the speed of sound and  $A$  is the cross-sectional area of the waveguide containing the wave,  $\beta \approx 1$  indicating that the regenerator is dissipating nearly all of the acoustic power it is trying to amplify. Ceperley<sup>6</sup> realized that viscous dissipation in the regenerator is a serious problem, but his proposed solution involved placing the regenerator in a looped resonator at least one wavelength long.<sup>5,6</sup> In an engine of this type, the viscous dissipation associated with the large amount of resonator surface area would cause a large reduction in efficiency. Also, it is not clear how to control the phase between  $p_{1,c}$  and  $U_{1,c}$  or  $|p_{1,c}/U_{1,c}|$ . The engine described in this article avoids much of the resonator dissipation by using an acoustic network that is much shorter than 1/4 of a wavelength. Also, the magnitudes of  $R$ ,  $L$ , and  $C$  are designed to create  $|Z|$  which is 15 to 30 times  $\rho_m c/A$  and  $p_{1,c}$  roughly in phase with  $U_{1,c}$ , thus avoiding large viscous losses in the regenerator.

The simplified model presented above provides an intuitive picture of how the power flows are distributed, how  $T_h$  and  $p_{1,c}$  are determined, and what sets  $\omega$ . One place where more accuracy is desired is the prediction of  $U_{1,c}$  and  $U_{1,h}$ . These two phasors determine whether the acoustic wave has the desired traveling-wave phasing. Detailed control of this phasing is advantageous as it allows the maximum acoustic power flow with the minimum acoustic velocity and therefore minimum viscous loss.<sup>2</sup> It is not intuitively obvious how these phasors are influenced by the impedances of the components, and it turns out they can be very sensitive to non-ideal behavior of the various components in the torus. To analyze these phenomena, Fig. 2(b) shows a detailed impedance model of the engine, which includes more of the acoustic elements in the engine, provides a more complete model of the regenerator, and includes the nonideal behavior of some the acoustic elements.

The regenerator is more complex than in the lumped-element model presented above. The volumetric-velocity gain occurs in a distributed manner. Therefore, the hotter sections of the regenerator have larger volumetric velocities and larger pressure drops. The pressure drops on the hot end are enhanced further due to the temperature dependence of the viscosity. Also, the porosity of the regenerator gives it compliance. These two effects are shown in the distributed regenerator model of Fig. 2(a). The regenerator of length  $l$  is split into  $N = l/\Delta x$  segments, each of length  $\Delta x$  and spanning a temperature difference  $\Delta T_n$ . With an ideal gas as the working fluid, each segment of regenerator has an isothermal compliance  $\Delta C = C_0 \Delta x/l$  independent of temperature, where  $C_0 = \phi S/l/p_m$  and  $\phi$  and  $S$  are the volume porosity and cross-sectional area of the regenerator. Across each segment, the volumetric velocity changes due to the temperature increase and the compliance, i.e.,  $U_{1,n+1} - U_{1,n} = U_{1,n} \Delta T_n / T_{m,n} - i \omega \Delta C p_{1,n}$ . Dividing through by  $\Delta x$  and letting  $\Delta x \rightarrow 0$  yields a differential equation for  $U_1(x)$ ,

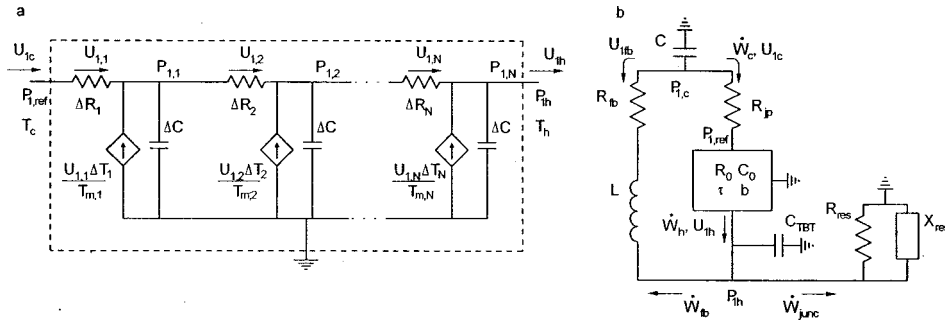


FIG. 2. (a) Distributed model of the regenerator. Each of the  $N$  sections spans  $\Delta T_n$  and represents a length  $\Delta x$  of the regenerator. The resistances  $\Delta R_n$  differ due to the temperature dependence of viscosity. The compliances are due to the void volume in the regenerator. (b) Detailed impedance model of the thermoacoustic-Stirling engine. The distributed model of the regenerator is included as the three-terminal network. Resistors  $R_{fb}$  and  $R_{jp}$  represent the resistances due to viscous and minor losses in the feedback inductance and jet pump respectively.  $C_{TBT}$  represents the compliance of the thermal buffer tube. Subscripts in (b) correspond to the locations in Fig. 1(b).

$$\frac{dU_1}{dx} = \frac{U_1}{T_m} \frac{dT_m}{dx} - i\omega(C_0/l)p_1(x). \quad (8)$$

Each segment of regenerator also has a resistance given by  $\Delta R_n = (R_0 \Delta x / l)(T_{m,n} / T_c)^b$ , where  $R_0 \approx 6\mu_c l / S r_h^2$  is the low-Reynolds-number-limit flow resistance of the regenerator when its entire length is at temperature  $T_c$ . At high amplitude,  $R_0$  would increase by up to 50%, but this effect is not included in the model. The factor  $(T_{m,n} / T_c)^b$  accounts for the temperature dependence of the viscosity  $\mu(T) = \mu_c (T / T_c)^b$ . The pressure drop across each segment is given by  $p_{1,n} - p_{1,n-1} = -\Delta R_n U_{1,n}$ . Dividing through by  $\Delta x$  and letting  $\Delta x \rightarrow 0$  yields a differential equation for  $p_1(x)$ ,

$$\frac{dp_1}{dx} = -\frac{R_0}{l} \left(\frac{T_m}{T_c}\right)^b U_1(x). \quad (9)$$

In the limit of low amplitude and  $r_h \ll \delta_c$ , Eqs. (8) and (9) are consistent with Eqs. (12) and (42) of Ref. 26.

Under typical operating conditions, the pressure drop across the regenerator  $|\Delta p_{1,regen}| = |p_{1,ref} - p_{1,h}| \approx 0.1 |p_{1,ref}|$ . Therefore, an approximate solution of Eqs. (8) and (9) is obtained by setting  $p_1(x) = p_{1,ref}$  in Eq. (8) and integrating. The result,

$$U_1(x) = \left[ 1 + (\tau - 1) \frac{x}{l} \right] \times \left\{ U_{1,c} - \frac{i\omega C_0 p_{1,ref}}{\tau - 1} \ln \left[ 1 + (\tau - 1) \frac{x}{l} \right] \right\}, \quad (10)$$

is then substituted into Eq. (9) and integrated again. In both integrations,  $T_m(x)$  is assumed to vary linearly from  $T_c$  to  $T_h$ . The volumetric velocity at the hot end of the regenerator and the pressure drop are then given by

$$U_{1,h} = \tau \left( U_{1,c} - \frac{i\omega C_0 \ln \tau}{\tau - 1} p_{1,ref} \right), \quad (11)$$

$$\Delta p_{1,regen} = U_{1,c} R_0 \frac{\tau + 1}{2} f(\tau, b) - \frac{i\omega C_0 R_0 p_{1,ref}}{2} g(\tau, b),$$

where  $f$  and  $g$  are given by

$$f(\tau, b) = \frac{2}{b+2} \left[ \frac{\tau^{b+2} - 1}{\tau^2 - 1} \right], \quad (12)$$

$$g(\tau, b) = \frac{2}{b+2} \left[ \frac{\tau^{b+2} \ln \tau - (\tau^{b+2} - 1)/(b+2)}{(\tau - 1)^2} \right].$$

Both of these functions approach 1 as  $\tau \rightarrow 1$ . Using  $b = 0.68$ , which is appropriate for helium gas, the calculated pressure drop at  $\tau = 3$  is approximately 70% bigger than that calculated with  $b = 0$ . Equations (11) and (12) for  $\Delta p_{1,regen}$  and  $U_{1,h}$  in terms of  $U_{1,c}$  and  $p_{1,ref}$  provide a simple model of a regenerator which includes its distributed properties such as axial temperature dependence and compliance.

The distributed model of the regenerator is included as a three-terminal network in the detailed impedance model of the engine shown in Fig. 2(b). Several new elements are also included. The resistance  $R_{fb}$  in the feedback path represents acoustic losses due to both viscous drag and flow separation.<sup>19</sup> The extra resistance  $R_{jp}$  appearing on the cold side of the regenerator represents the same type of losses which occur in the jet pump. Finally, the compliance  $C_{TBT}$  appearing below the regenerator models the compressibility of the gas in the TBT. A detailed analysis of the model in Fig. 2(b) shows that the volumetric velocity into the cold end of the regenerator is given by

$$U_{1,c} = \frac{p_{1,c}}{R_{fb} + \frac{i\omega L}{R_{St}}} \left[ \frac{\omega^2 LC}{R_{St}} \left( 1 - \frac{iR_{fb}}{\omega L} \right) + \frac{i\omega R_0 C_0 g}{2R_{St}} \right], \quad (13)$$

where  $R_{St} = R_{jp} + R_0(\tau + 1)f/2$  is the effective series resistance of the jet pump and regenerator referred to  $U_{1,c}$ . In Eq. (13), we have ignored the pressure drop across  $R_{jp}$  by taking  $p_{1,ref} = p_{1,c}$ . This equation provides a more accurate version of the phenomena discussed above than Eq. (2).

Equations (11) and (13) are important in the design of a thermoacoustic-Stirling engine. For efficient regenerator operation,  $U_{1,c}$  should lead  $p_{1,c}$ , and  $U_{1,h}$  should lag  $p_{1,c}$ .<sup>2</sup> This allows maximum acoustic power flow with the minimum volumetric velocity, and therefore a minimum of viscous loss in the regenerator. Also, for a given  $p_{1,c}$ , the size of the various components determines the magnitude of  $U_{1,c}$  and hence the amount of power flow through the regenerator.

If the engine must drive a very lossy system, the power output of the engine can be increased at the expense of increased regenerator loss, allowing the designer some freedom to balance the losses against one another.

The origin of the dominant part of Eq. (13),  $\omega^2 LC p_{1,c}/R_{St}$ , is understood by considering the pressure drop across the impedance of the parallel combination of the feedback and regenerator branch,  $Z_p$ . Ignoring the compliance in the regenerator branch for the moment, the total volumetric velocity through these two branches is  $U_{1,c} + U_{1,fb} = -i\omega C p_{1,c}$ . Therefore, the pressure drop  $\Delta p_1 = -i\omega C Z_p p_{1,c}$  appears across the temperature-dependent resistance  $R_{St}$  of the regenerator branch, giving  $U_{1,c} = -i\omega C Z_p p_{1,c}/R_{St}$ . If the impedance of the inertance dominates, i.e.,  $R_{St} \gg \omega L \gg R_{fb}$ , then  $Z_p \approx i\omega L$ . In this case,  $U_{1,c} = \omega^2 LC p_{1,c}/R_{St}$  and  $U_{1,c}$  is in phase with  $p_{1,c}$ .

If  $\omega L$  becomes comparable to either  $R_{St}$  or  $R_{fb}$ ,  $Z_p$  rotates into the first quadrant, causing  $U_{1,c}$  to lag  $p_{1,c}$ . This accounts for the  $1 - iR_{fb}/(\omega L)$  and  $1 + (i\omega L)/R_{St}$  terms in Eq. (13). Even in the most ideal case, i.e.,  $R_{St} \gg \omega L \gg R_{fb}$ ,  $U_{1,c}$  will at best be in phase with  $p_{1,c}$ . Some other effect is needed to force  $U_{1,c}$  to the desired condition leading  $p_{1,c}$ . Such an effect is found in the compliance of the regenerator. The oscillating pressure  $p_{1,c}$  causes a volumetric-velocity difference of  $i\omega C_0 p_{1,c}$  across the regenerator. Only part of this volumetric velocity originates from the cold end of the regenerator, but the fraction that does is responsible for the term  $(i\omega C_0 p_{1,c})gR_0/(2R_{St})$  in Eq. (13). This volumetric velocity is proportional to  $i p_{1,c}$  and causes  $U_{1,c}$  to lead  $p_{1,c}$ .

### III. MEASUREMENTS OF ACOUSTICAL PRESSURE PHASORS

Many thermoacoustic quantities in the engine, such as pressure drops, acoustic power output, heat flows, and streaming mass fluxes, typically scale as the square of the ratio of the acoustic pressure to mean pressure. Often these quantities also depend on the temperature ratio across the regenerator. To separate the dependencies, the pressure ratio is held constant by adjusting  $\dot{Q}_h$  while  $T_h$  is varied by adjusting the variable acoustic load. Therefore, we must pick a reference location where the pressure amplitude will be held constant. We choose the pressure  $p_{ref} = |p_{1,ref}|$  immediately above the main cold heat exchanger [see Figs. 1(b) and 2(b)] as it determines  $\dot{W}_c$  and therefore the acoustic power output and heat flow into the engine. Therefore, all acoustic amplitudes will be quoted in terms of  $p_{ref}/p_m$ .

Analysis of the engine begins with the acoustics within the torus section. In addition to the pressure amplitude, the important quantities are the magnitude and phase of the volumetric velocity at the hot and cold ends of the regenerator. Although these quantities are not directly measurable, related quantities such as the pressure drop across the regenerator and jet pump are readily measured. Since  $U_1$  and the local resistance depend on location in the regenerator, the regenerator pressure drop  $\Delta p_{1,regen}$  comes from a weighted average of  $U_1$  within the regenerator. Since the jet pump has negligible compliance,  $U_1$  is constant there, and the pressure drop across the jet pump is determined by  $U_{1,c}$ .

We will compare the pressure-drop data to the results of two different models, the impedance model of Fig. 2(b) and a more exact DeltaE model.<sup>27,28</sup> The DeltaE model is based on the as-built dimensions of the engine, and it also includes ‘‘minor losses’’ where the flow expands in a diffuser, turns in an elbow, jets from an orifice, or merges with another flow. The dissipation associated with these minor losses is based on steady-flow correlations,<sup>19</sup> and will be discussed further in the later section on the jet pump. The term ‘‘minor loss’’ originates in the analysis of (long) piping systems where entry effects, valve constrictions, bends, etc., are ‘‘minor’’ compared to the energy loss due to flow through long, straight piping sections (e.g., Ref. 29). In devices such as the engine described here, these minor losses have significant consequences.

Starting with the impedance model of Fig. 2(b), the pressure drop across the jet pump plus regenerator is given by  $\Delta p_{1,jp+regen} = R_{jp} U_{1,c} + \Delta p_{1,regen}$ . Substituting Eqs. (11) and (13) we find

$$\frac{\Delta p_{1,jp+regen}}{p_{1,c}} = \omega^2 LC \frac{(1 + gR_0 C_0 / 2R_{St} C)(1 - iR_{fb}/\omega L)}{1 + (i\omega L/R_{St})(1 - iR_{fb}/\omega L)}. \quad (14)$$

Before making a comparison, we make several simplifications. In the low-amplitude limit,  $R_{fb} \ll \omega L$  because the radius of the inertance is much larger than a viscous or thermal penetration depth. At high acoustic amplitudes,  $R_{fb}$  increases due to minor losses and turbulent flow,<sup>30</sup> but the gentle transitions in the feedback flow path keep  $R_{fb} \leq 0.1\omega L$  even at  $p_{ref}/p_m = 0.10$ . Therefore, we will ignore  $R_{fb}/\omega L$  terms in Eq. (14) since they contribute very little to the magnitude and only a few degrees of phase to the right hand side of Eq. (14). However, if a particular design results in  $R_{fb} \approx \omega L$ , these terms (especially the one in the numerator) cannot be ignored. The low-amplitude limit also allows us to take  $R_{St} \approx R_0(\tau + 1)f/2$ , ignoring  $R_{jp}$ . This is not such a good approximation at  $p_{ref}/p_m = 0.10$ , where  $R_{jp} \approx R_0(\tau + 1)f/2$ . Using these approximations, Eq. (14) becomes

$$\frac{\Delta p_{1,jp+regen}}{p_{1,c}} = \omega^2 LC \frac{1 + gC_0/fC(\tau + 1)}{1 + 2i\omega L/fR_0(\tau + 1)}. \quad (15)$$

Finally, we note that the volume of the inertance must be included when calculating  $C$  as it contributes a significant compliance.

The lower and upper crosses in Fig. 3 are the result of Eq. (15) with  $\tau = 2$  and 3, respectively. The filled symbols are the measured  $\Delta p_{1,jp+regen}/p_{1,c}$  at  $p_{ref}/p_m$  ranging from 0.026 to 0.10 with the acoustic amplitude higher for the data nearer the real axis. All data are taken with  $\tau \approx 3$  except for  $p_{ref}/p_m = 0.026$  which is taken at  $\tau \approx 2$ . Over this range of amplitudes, the agreement between the detailed impedance model of Fig. 2(b) and the measured data is excellent; better than 10% in magnitude and a few degrees in phase. The open symbols are  $\Delta p_{1,jp+regen}/p_{1,c}$  calculated using DeltaE. The numerical model, which is more accurate in computing the effective values of  $L$  and  $C$  in the feedback inertance and includes dissipation due to turbulence and minor losses, shows even better agreement.

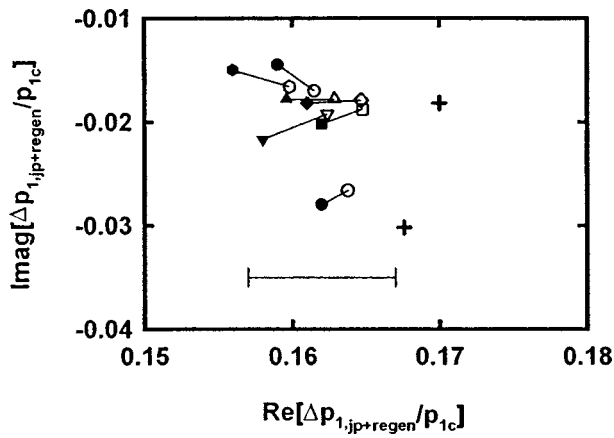


FIG. 3. Real and imaginary components of  $\Delta p_{1,jp+regen}/p_{1,c}$ , the pressure drop across the jet pump plus regenerator normalized by the acoustic amplitude. The upper cross is the result of the impedance model of Fig. 2(b) given in Eq. (15) with  $\tau=3$ . The lower cross is calculated with  $\tau=2$ . The filled symbols are the measured values of  $\Delta p_{1,jp+regen}/p_{1,c}$  at  $p_{ref}/p_m = 0.026, 0.038, 0.051, 0.061, 0.069, 0.088,$  and  $0.10$  with larger amplitudes nearer the top of the figure. All measurements are taken with  $\tau \approx 3$  except for  $p_{ref}/p_m = 0.026$  which is taken at  $\tau \approx 2$ . This low-amplitude measurement should be compared with the lower cross. The open symbols are calculated with DeltaE using the same value of  $\tau$  as the experimental data. At each amplitude, the experimental data and DeltaE result are connected with a line to ease the comparison. Over this span of amplitudes, there is good agreement between the measured values and the impedance model indicating the approximations used in deriving Eq. (15) are valid. The DeltaE model, which is more accurate in computing  $L$  and  $C$  and takes into account the nonlinear resistances of the feedback inertance and jet pump due to minor losses and turbulence, shows even better agreement as the amplitude is varied. The error bar near the bottom of the figure shows the worst-case uncertainty in the real part of the measured  $\Delta p_{1,jp+regen}/p_{1,c}$ . The uncertainty in the real part is mostly due to changes in the pressure sensors' sensitivity caused by thermal drifts. The uncertainty in the imaginary part is negligible on this scale.

Even though the approximations used in reaching Eq. (15) are far from being valid at  $p_{ref}/p_m = 0.10$ , the agreement in Fig. 3 is still quite good. The explanation of this robustness is similar to the discussion that follows Eq. (13). Since  $\omega L \ll R_0(\tau+1)f/2$  and the feedback inertance has been designed so that  $R_{fb} \ll \omega L$  even at high amplitude,  $\Delta p_{1,jp+regen}$  is mostly determined by  $L$  and  $C$ . Errors in the impedance of the regenerator and jet pump do not have a large effect. If the terms in Eq. (15) other than  $\omega^2 LC$  are ignored,  $\Delta p_{1,jp+regen}/p_{1,c} = 0.166$ , which differs from the measured  $\Delta p_{1,jp+regen}/p_{1,c}$  by less than 5% in magnitude and  $10^\circ$  in phase. Ignoring the nonideal properties of the acoustic elements in the torus, such as viscous flow resistance, thermal relaxation losses, and minor losses, does not lead to large errors in  $\Delta p_{1,jp+regen}$ . Therefore, the approximate expressions for the minor losses used in the DeltaE model, which are discussed in the section on Gedeon streaming, are expected to result in better agreement between the measurement and calculations, but errors in these approximations will not effect the results significantly.

As a check of how well the properties of the jet pump and regenerator are known, Fig. 4 presents the pressure drops across these individual elements as a function of acoustic amplitude and  $T_h$ . The filled and open symbols are measured values of  $\Delta p_{1,regen}$  and  $\Delta p_{1,jp}$ , respectively. The two data points for each acoustic amplitude are taken at the high and

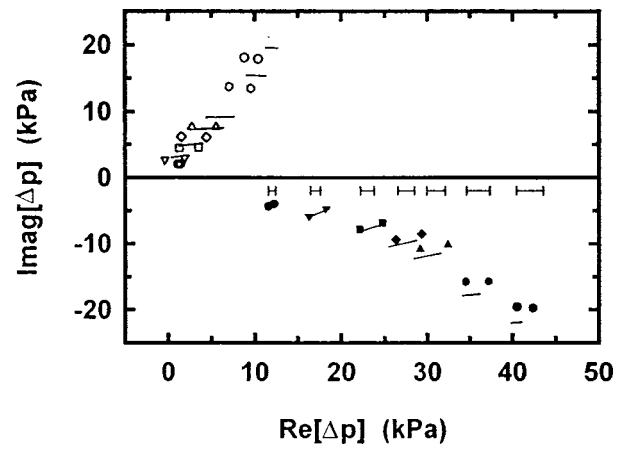


FIG. 4. Real and imaginary parts of  $\Delta p_{1,jp}$  (open symbols) and  $\Delta p_{1,regen}$  (filled symbols) measured at the same values of  $p_{ref}/p_m$  as shown in Fig. 3. The two symbols at each amplitude represent the upper and lower extremes of  $T_h$ . The upper limit is typically  $700^\circ\text{C}$  while the lower limit is amplitude dependent. The solid lines are DeltaE calculations at the same amplitudes and between the same temperature extremes. The error bars near the middle of the figure indicate the typical uncertainty in the real part of the  $\Delta p_{1,regen}$  measurements. Each individual error bar applies to the two equal-amplitude data points directly below the error bar and to the real part of the corresponding  $\Delta p_{1,jp}$  data points in the upper half of the figure. The uncertainty in the imaginary parts of  $\Delta p_{1,jp}$  and  $\Delta p_{1,regen}$  is negligible on this scale.

low extremes of  $T_h$ . The solid lines are calculated with DeltaE between the same values of  $T_h$ . The only adjustable parameter used in this part of the model is the relative surface roughness  $\epsilon$  of the piping used in the feedback inertance.<sup>29</sup> The DeltaE results in Fig. 4 are generated with  $\epsilon = 1.8 \times 10^{-3}$  corresponding to an absolute roughness of  $140 \mu\text{m}$ . This absolute roughness is consistent with what we have used in DeltaE models to obtain agreement with measurements in other thermoacoustic devices. Increasing or decreasing the surface roughness used in the calculation by a factor of 2 does not effect the pressure-drop predictions significantly. The agreement here is quite good, indicating that the acoustical phenomena in the torus section are well understood. In particular, the agreement between measured and calculated  $\Delta p_{1,jp}$  provides evidence that the model used for the minor loss acoustic resistance given in Eq. (24) is reasonably accurate. If the acoustic resistance predicted by this model were ignored, the predicted magnitude of the jet-pump impedance would be approximately 3.3 times smaller at  $p_{ref}/p_m = 0.10$ .

With confidence in the DeltaE model, it is used to calculate the volumetric-velocity phasors in the regenerator at  $\tau = 3.3$ . In Table I, the results of the DeltaE calculations are compared with those of the detailed impedance model of Fig. 2(b) performed in two ways, setting  $R_{fb} = 0$  or  $R_{fb} = R_{jp} = 0$ . When  $R_{jp}$  is taken into account, there is reasonable agreement in  $|U_{1,c}|$  and  $|U_{1,h}|$  between DeltaE and the detailed impedance model. If  $R_{jp}$  is ignored at high amplitude where  $R_{jp} \approx f(\tau+1)R_0/2$ , there are large errors in the magnitudes due to the  $\omega^2 LC/R_{st}$  term in Eq. (13). However, in neither case do the impedance model calculations reproduce the phase of  $U_{1,c}$  or  $U_{1,h}$ . Much of this phase error can be attributed to the open spaces between the jet pump and regenerator which contribute additional compliances not included

TABLE I. Calculated values of acoustic admittances  $U_{1,c}/p_{1,c}$  and  $U_{1,h}/p_{1,c}$  at  $p_{\text{ref}}/p_m=0.026$  and  $0.10$  with  $\tau=3.3$  using three different methods: DeltaE, the impedance model of Fig. 2(b) with  $R_{\text{fb}}=0$ , and the impedance model of Fig. 2(b) with  $R_{\text{fb}}=R_{\text{jp}}=0$ . The values of  $U_{1,c}/p_{1,c}$  are given in  $\text{m}^3/\text{Pa s}$ . At low amplitude,  $R_{\text{jp}} \ll (\tau + 1)fR_0/2$  and  $R_{\text{fb}} \ll \omega L$ . Therefore, all three methods are in reasonable agreement at  $p_{\text{ref}}/p_m=0.026$ . At  $p_{\text{ref}}/p_m=0.10$ ,  $R_{\text{jp}} \approx (\tau + 1)fR_0/2$  and ignoring  $R_{\text{jp}}$  leads to significant error in the magnitudes of  $U_{1,c}$  and  $U_{1,h}$ . There are phase errors of  $\approx 15^\circ - 30^\circ$  in the impedance model, much of which is due to ignoring compliances of components between the jet pump and regenerator.

$p_{\text{ref}}/p_m$		DeltaE		$R_{\text{fb}}=0$		$R_{\text{fb}}=R_{\text{jp}}=0$	
0.026	$U_{1,c}/p_{1,c}$	$5.0 \times 10^{-8}$	$30^\circ$	$5.5 \times 10^{-8}$	$16^\circ$	$5.8 \times 10^{-8}$	$15^\circ$
0.026	$U_{1,h}/p_{1,c}$	$1.5 \times 10^{-7}$	$-30^\circ$	$1.8 \times 10^{-7}$	$-14^\circ$	$1.9 \times 10^{-7}$	$-13^\circ$
0.100	$U_{1,c}/p_{1,c}$	$3.5 \times 10^{-8}$	$46^\circ$	$3.8 \times 10^{-8}$	$17^\circ$	$5.8 \times 10^{-8}$	$15^\circ$
0.100	$U_{1,h}/p_{1,c}$	$9.6 \times 10^{-8}$	$-45^\circ$	$1.3 \times 10^{-7}$	$-25^\circ$	$1.9 \times 10^{-7}$	$-13^\circ$

in the impedance model. Although the impedance model does not predict the phases of  $U_{1,c}$  or  $U_{1,h}$  with great accuracy, we believe it is useful for making estimates in the initial design of an engine. Also, the intuitive picture it provides is indispensable when optimizing the engine dimensions while performing more precise, numerical calculations such as with DeltaE.

#### IV. STREAMING HEAT TRANSPORT AND ITS SUPPRESSION

We now discuss two types of streaming that occur within the engine, their effect on the engine, and methods of suppression.

##### A. Gedeon streaming

Gedeon streaming—acoustic streaming around the torus—is a major problem to be overcome.<sup>20</sup> Gedeon<sup>7</sup> has discussed how a second-order, time-averaged mass flux  $\dot{M}_2$  can exist in Stirling systems which contain a closed-loop path. The acoustic network of the torus in Fig. 1(b) clearly provides such a path. Gedeon argues that  $\dot{M}_2$  is given by

$$\dot{M}_2 = \text{Re}[\rho_1 \tilde{U}_1]/2 + \rho_m U_{2,0}, \quad (16)$$

where  $U_{2,0}$ , the second-order time-independent volumetric velocity, is the next term in the expansion given in Eq. (1). Since  $\rho_1 \propto p_1$ , the first term on the right hand side of Eq. (16) is nonzero wherever the acoustic power flux is nonzero, such as around the torus of Fig. 1(b). If no measures are taken to impose a  $U_{2,0}$  to cancel the first term, the resulting mass flux convects heat away from the hot heat exchanger and deposits it at the secondary cold heat exchanger, creating an unwanted heat leak of  $\dot{Q}_{\text{leak}} = \dot{M}_2 c_p (T_h - T_c)$ . Here,  $c_p$  is the constant-pressure heat capacity of the helium. At the cold end of the regenerator, the first term on the right hand side of Eq. (16) can be rewritten  $\text{Re}[\rho_1 \tilde{U}_1]/2 = \rho_{m,c} \dot{W}_c / p_m$  where  $\rho_{m,c}$  is the mean density evaluated at  $T_c$  and  $T_1$  is assumed to be zero. Using this result and Eq. (5),  $\dot{Q}_{\text{leak}}$  can be rewritten

$$\frac{\dot{Q}_{\text{leak}}}{\dot{W}_{\text{junc}}} = \frac{\rho_{m,c} c_p T_c}{p_m} = \frac{\gamma}{\gamma - 1} \quad (17)$$

if  $U_{2,0} = 0$ . For a monatomic ideal gas, this ratio is 2.5. The scale of this heat leak can be understood by considering a reversible engine operating with  $T_h \rightarrow \infty$ , and therefore having Carnot efficiency<sup>14</sup>  $\eta_C \rightarrow 1$ . The heat leak of Eq. (17)

would reduce the maximum possible efficiency to 0.29. Therefore, this streaming mechanism is a considerable load on the hot heat exchanger and must be suppressed.

To enforce  $\dot{M}_2 = 0$ , a nonzero  $U_{2,0}$  must flow around the feedback loop and up through the regenerator. The only significant steady-flow resistance in this path is the regenerator. Swift *et al.*<sup>20</sup> have shown that the pressure drop across the regenerator that is needed to drive the correct  $U_{2,0}$  in the low-Reynolds-number limit is given by

$$\Delta p_{2,0} \approx \frac{6}{S r_h^2 p_m} \int_{\text{reg}} \mu_m(x) \dot{W}(x) dx, \quad (18)$$

where  $\mu_m(x)$  and  $\dot{W}(x)$  are the spatially dependent coefficient of viscosity and acoustic power flow in the regenerator. Assuming a constant temperature gradient throughout the regenerator (only valid when the streaming is suppressed) and  $\dot{W}(x) = \dot{W}_c T_m(x)/T_c$ , the integral in Eq. (18) yields

$$\Delta p_{2,0} = \frac{6 \mu_c l}{(b+2) S r_h^2 p_m} \left[ \frac{\tau^{b+2} - 1}{\tau - 1} \right] \dot{W}_c \approx \frac{R_0(\tau+1)f}{2} \frac{\dot{W}_c}{p_m}. \quad (19)$$

This required pressure drop is typically in the range 300–1500 Pa, with the higher pressure on the hot side of the regenerator.

One way<sup>20</sup> to impose such a  $\Delta p_{2,0}$  across the regenerator is to use the asymmetry of hydrodynamic end effects. When high-Reynolds-number flow makes an abrupt transition from a pipe of small cross-sectional area to a larger area, the transition is accompanied by jet flow and turbulence,<sup>30</sup> Bernoulli's equation is not expected to hold, and the transition generates an additional pressure drop and dissipation. This phenomenon is often referred to as ‘‘minor loss.’’ The additional pressure drop  $\Delta p_{\text{ml}}$  is

$$\Delta p_{\text{ml}} = K \rho v^2 / 2, \quad (20)$$

where  $K$  depends on the flow geometry and is termed the ‘‘minor loss coefficient.’’ For an abrupt transition where steady flow *expands* into a much larger cross-sectional area,  $K = K_{\text{exp}} \approx 1$ .<sup>19,30,29</sup> The value of  $K_{\text{exp}}$  is approximately constant and independent of the geometry of the transition. For steady flow in the opposite direction, *contracting* into the smaller pipe,  $K = K_{\text{con}}$  is strongly dependent on geometrical details. If the edge of the entrance is sharp,  $K_{\text{con}} \approx 0.5$ . As the edge of the entrance is rounded,  $K_{\text{con}}$  falls until the radius of the rounding,  $r$ , is such that  $r/D \geq 0.15$ .<sup>19,30,29</sup> Here,  $D$  is the

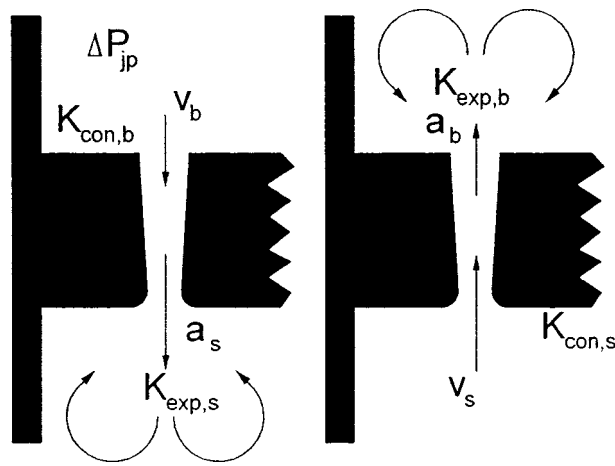


FIG. 5. Schematic diagram of the jet pump. The actual jet pump has two adjustable slots, only one of which is shown here. It is shown for both directions of flow. Areas  $a_s$  and  $a_b$  refer to total areas (both slots). The big opening has area  $a_b = 3 \text{ cm}^2$  and the small opening area  $a_s$  is adjustable, from outside the pressure vessel, over the range  $0\text{--}3 \text{ cm}^2$ . Subscripts  $s$  and  $b$  refer to the small and big openings, respectively. The two flow patterns represent the first-order acoustic velocity during the two halves of the acoustic cycle. Approximate values for the minor loss coefficients are  $K_{\text{exp},b} = K_{\text{exp},s} = 1.0$ ,  $K_{\text{con},b} = 0.5$ , and  $K_{\text{con},s} = 0.04$ . The time-averaged pressure drop  $\overline{\Delta p_{\text{jp}}}$  is generated such that the pressure is higher above the jet pump.

diameter of a circular opening which forms the entrance. At these values of  $r$ ,  $K_{\text{con}} \approx 0.04$  and the flow resistance of an abrupt transition displays a nearly perfect asymmetry with respect to flow direction.

Figure 5 shows a schematic of the ‘‘jet pump’’ which is located between the compliance and main cold heat exchanger in Fig. 1(b). The two different flow patterns shown represent the first-order acoustic velocity during the two halves of the acoustic cycle. The jet pump utilizes minor losses by channeling the flow above the main cold heat exchanger through two parallel tapered rectangular channels that open abruptly into the larger space below the jet pump. The long dimension (into the page) of the outlets is fixed at  $51 \text{ mm}$ . The short dimension is adjustable and typically set near  $\delta g = 1\text{--}2 \text{ mm}$ . The edges of the outlets are rounded with a  $0.8\text{-mm}$  radius, which is estimated to be enough to reduce  $K_{\text{con},s}$  to nearly zero. The subscripts  $s$  and  $b$  correspond to the small and big openings at opposite ends of the jet pump. If the channels leading up to the expansion were not tapered, nearly the same minor loss would take place at both ends and the jet pump would show little asymmetry. The taper increases the cross-sectional area at the large end thereby decreasing the velocity. From Eq. (20), the reduction in velocity results in a much smaller pressure drop.

To estimate the time-averaged pressure drop across the jet pump, the coefficients  $K_{\text{exp}}$  and  $K_{\text{con}}$  are assumed to have the same values in oscillating flow as they do in steady flow. The accuracy of this assumption is unknown, although the large displacement amplitudes ( $|x_1|/\delta g \approx 200$ ) give us hope that at each instant of time, the flow has little memory of its time dependence. Taking the instantaneous velocities to be  $v_s(t) = v_{1,s} \sin \omega t$  and  $v_b(t) = (a_s/a_b)v_s(t)$  with positive velocities downward, the instantaneous pressure drop due to the first-order velocity is given by

$$\Delta p_{\text{jp}}(0 \leq t \leq \pi/\omega) = \frac{\rho_m}{2} (K_{\text{con},b} v_b^2 + K_{\text{exp},s} v_s^2), \quad (21)$$

$$\Delta p_{\text{jp}}(\pi/\omega \leq t \leq 2\pi/\omega) = -\frac{\rho_m}{2} (K_{\text{exp},b} v_b^2 + K_{\text{con},s} v_s^2).$$

Again, we assume that the coefficients  $K_{\text{exp}}$  and  $K_{\text{con}}$  have the same values in oscillatory flow as they do in steady flow, and that Eq. (20) can be applied instantaneously throughout the acoustic cycle. The time-averaged pressure drop across the jet pump is found to be

$$\overline{\Delta p_{\text{jp}}} = \frac{\rho_m |U_{1,\text{jp}}|^2}{8a_s^2} \left[ (K_{\text{exp},s} - K_{\text{con},s}) + \left( \frac{a_s}{a_b} \right)^2 (K_{\text{con},b} - K_{\text{exp},b}) \right], \quad (22)$$

where  $U_{1,\text{jp}}$  is the amplitude of the first-order volumetric velocity through the jet pump. For the orientation of the jet pump in Fig. 1(b),  $\Delta p_{\text{jp}}$  has the correct sign to suppress the streaming around the feedback loop.<sup>20</sup> In this derivation, we have ignored the pressure drop due to the inertance and linear resistance of the jet pump as they do not contribute to  $\overline{\Delta p_{\text{jp}}}$ .

In addition to generating  $\overline{\Delta p_{\text{jp}}}$ , the jet pump also dissipates acoustic power.<sup>20</sup> The instantaneous power dissipation is given by  $\Delta p_{\text{jp}}(t)U_{\text{jp}}(t)$ . Averaging over one acoustic cycle, the time-averaged power dissipation is

$$\dot{E} = \frac{\rho_m |U_{1,\text{jp}}|^3}{3\pi a_s^2} \left[ (K_{\text{exp},s} + K_{\text{con},s}) + \left( \frac{a_s}{a_b} \right)^2 (K_{\text{con},b} + K_{\text{exp},b}) \right]. \quad (23)$$

For a pure acoustic resistance  $R$ , the time-averaged power dissipation is given by  $\dot{E} = R|U_1|^2/2$ . Using this relation, we can assign a resistance to the jet pump of

$$R_{\text{jp}} = \frac{2\rho_m |U_{1,\text{jp}}|}{3\pi a_s^2} \left[ (K_{\text{exp},s} + K_{\text{con},s}) + \left( \frac{a_s}{a_b} \right)^2 (K_{\text{con},b} + K_{\text{exp},b}) \right]. \quad (24)$$

This resistance is in addition to that due to viscous and thermal relaxation losses. The minor loss resistance of other elements in the engine is calculated in a similar way.

Figure 6(a) and (b) demonstrate the effectiveness of the jet pump. Figure 6(b) shows the time-dependent  $\Delta p_{\text{jp}}(t)$ , measured using a differential piezoresistive pressure sensor,<sup>22</sup> displayed as a function of the time phase  $\phi(t)$  of the acoustic cycle. The sensor itself is located outside the pressure vessel, and approximately  $10 \text{ cm}$  of  $1.5\text{-mm}$ -diameter copper tubing is used to connect the front and back of the sensing membrane to the top and bottom of the jet pump. The ends of the copper tubing are located close to the corners formed by the plate that houses the jet pump and the pressure-vessel wall. We believe these locations are best because they are nearly stagnation zones of the secondary flows generated by the jet pump, but we have not investigated how



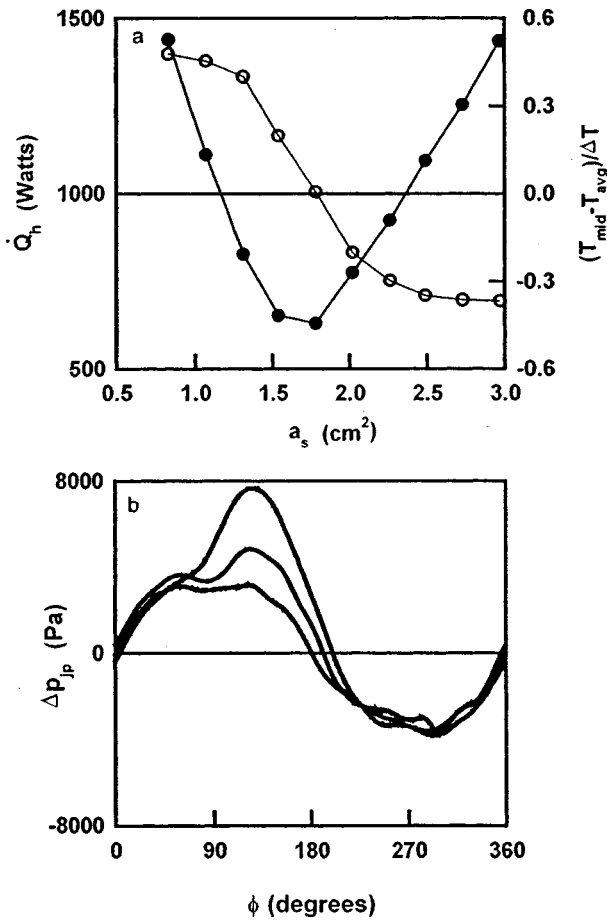


FIG. 6. (a)  $\dot{Q}_h$  (filled symbols) and  $(T_{mid} - T_{avg})/\Delta T$  (open symbols) vs the jet area  $a_s$ . Here,  $\Delta T = T_h - T_c$  and  $T_{avg} = (T_h + T_c)/2$ . As  $a_s$  is varied,  $\dot{Q}_h$  is adjusted to maintain  $p_{ref}/p_m = 0.036$ , and hence  $Re[\rho_1 \bar{U}_1]$  nearly constant within the regenerator. With  $a_s = 1.8$  cm<sup>2</sup>,  $T_{mid} = T_{avg}$  indicating that  $\dot{M}_2$  is completely suppressed. With  $\dot{M}_2 = 0$ ,  $\dot{Q}_{leak} = 0$  and the heat input  $\dot{Q}_h$  needed to maintain the oscillations is at a minimum. Typical uncertainties are  $\pm 20$  W in  $\dot{Q}_h$ ,  $\pm 0.04$  cm<sup>2</sup> in  $a_s$ , and  $\pm 1\%$  in  $(T_{mid} - T_{avg})/\Delta T$ . (b) Measurements of  $\Delta p_{jp}(t)$  as a function of the time phase of the acoustic cycle  $\phi(t)$ . The lowest curve (at  $\phi = 120^\circ$ ), obtained with  $a_s = 3.0$  cm<sup>2</sup>, has  $\overline{\Delta p_{jp}} = 120$  Pa; the middle curve, obtained with  $a_s = 1.8$  cm<sup>2</sup>, has  $\overline{\Delta p_{jp}} = 630$  Pa; the highest curve, obtained with  $a_s = 0.8$  cm<sup>2</sup>, has  $\overline{\Delta p_{jp}} = 1160$  Pa. The middle setting suppresses  $\dot{M}_2$ , as shown in (a). The sinusoidal background pressure drop is due to the inertance and linear resistance of the jet pump.

accurately this technique determines  $\Delta p_{jp}$ . Therefore, the magnitudes of  $\Delta p_{jp}$  cannot be trusted, but changes in  $\Delta p_{jp}$  between different settings of  $a_s$  allow a qualitative measure of the action of the jet pump.

The lowest curve (at  $\phi = 120^\circ$ ) in Fig. 6(b) displays  $\Delta p_{jp}(t)$  with  $a_s = a_b = 3.0$  cm<sup>2</sup>. Time averaged, this curve does not yield a significant  $\overline{\Delta p_{jp}}$ , so a large  $\dot{M}_2$  flows down through the regenerator and TBT causing a large  $\dot{Q}_{leak}$ . To maintain the oscillation at  $p_{ref}/p_m = 0.036$ , Fig. 6(a) shows that nearly 1500 W of  $\dot{Q}_h$  is necessary, with more than half of this  $\dot{Q}_h$  carried away by  $\dot{M}_2$ . Also, the temperature  $T_{mid}$  at the axial midpoint of the regenerator is held significantly below its expected temperature  $T_{avg} = (T_h + T_c)/2$  due to the flux of cold gas entering the cold end of the regenerator. The middle curve in Fig. 6(b) corresponds to  $a_s = 1.8$  cm<sup>2</sup> in Fig.

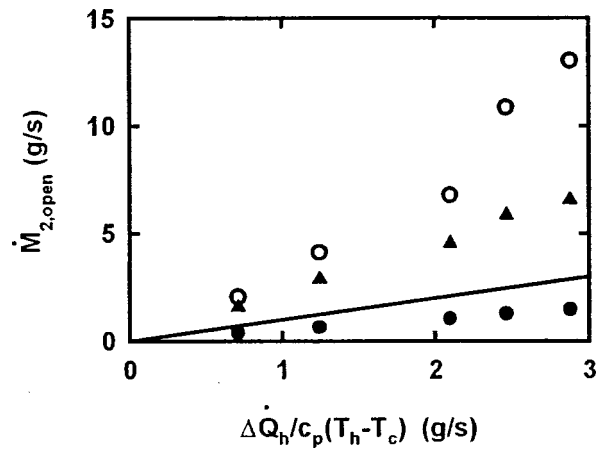


FIG. 7. Estimates of the Gedeon streaming mass flux  $\dot{M}_{2,open}$ . The solid line has slope 1 and represents  $\Delta \dot{Q}_h / c_p (T_h - T_c)$ . The filled circles are  $\rho_{m,c} \bar{W}_c / p_m$ . The filled triangles and open circles are the result of estimating  $\dot{M}_{2,open}$  from the measured and calculated values of  $\overline{\Delta p_{jp,0}}$ , respectively.

6(a). The resulting  $\overline{\Delta p_{jp}}$  generated by the jet pump enforces  $\dot{M}_2 = 0$ , so that  $T_{mid} = T_{avg}$  and  $\dot{Q}_{leak}$  is eliminated. Therefore,  $\dot{Q}_h$  needed to maintain  $p_{ref}/p_m = 0.036$  is at its minimum of 630 W. The highest curve in Fig. 6(b) corresponds to  $a_s = 0.8$  cm<sup>2</sup> in Fig. 6(a). The resulting  $\Delta p_{jp}$  exceeds the required  $\Delta p_{2,0}$ , so  $\dot{M}_2 < 0$ , implying that the time-averaged mass flux flows up through the TBT and regenerator. With  $\dot{Q}_h \approx 1500$  W, the convective heat leak has returned, and  $T_{mid} > T_{avg}$  due to the hot gas entering the hot end of the regenerator. These data demonstrate that  $\dot{M}_2$  causes a significant and unwanted heat load that can be completely eliminated by  $\Delta p_{jp}$ , and that  $T_{mid}$  can be used to indicate when  $\dot{M}_2 = 0$ . Also, the broad minimum in  $\dot{Q}_h$  in Fig. 6(a) shows that the cancellation of  $\dot{M}_2$  is insensitive to small variations in  $a_s$ .

The experiment described above demonstrates that the jet pump can stop the Gedeon streaming mass flux. A different experiment is used to determine the magnitude of the mass flux so it can be compared with several equations in this section. To avoid confusion, the streaming mass flux with  $a_s$  at its fully open position of 3 cm<sup>2</sup> is referred to as  $\dot{M}_{2,open}$ . At several fixed values of  $p_{ref}/p_m$ , two measurements of  $\dot{Q}_h$  are performed; one with  $a_s$  adjusted so that  $\dot{M}_2 = 0$  and a second with  $a_s$  fully open, i.e.,  $a_s = 3.0$  cm<sup>2</sup>. With  $a_s$  fully open, nearly all of the extra heat input  $\Delta \dot{Q}_h$  is carried away by  $\dot{M}_{2,open}$  which is calculated as  $\dot{M}_{2,open} = \Delta \dot{Q}_h / c_p (T_h - T_c)$ . If  $U_{2,0} = 0$  when  $a_s = 3.0$  cm<sup>2</sup>, Eq. (16) shows that  $\dot{M}_{2,open} = \rho_{m,c} \bar{W}_c / p_m$ . Using the DeltaE model developed in the previous section to compute  $\bar{W}_c$ , the measured  $\dot{M}_{2,open}$  and  $\rho_{m,c} \bar{W}_c / p_m$  are compared. The horizontal axis in Fig. 7 is given by the measured values of  $\Delta \dot{Q}_h / c_p (T_h - T_c)$ . The solid line has slope 1 and represents  $\dot{M}_{2,open} = \Delta \dot{Q}_h / c_p (T_h - T_c)$ . The filled circles are  $\dot{M}_{2,open} = \rho_{m,c} \bar{W}_c / p_m$  and are roughly a factor of 2 lower than  $\Delta \dot{Q}_h / c_p (T_h - T_c)$ . We believe that the combined uncertainty in the  $\Delta \dot{Q}_h$  measurement of  $\dot{M}_{2,open}$  and the computation of

$\dot{W}_c$  cannot account for this difference. This result suggests that there are additional sources of  $\Delta p_{2,0}$  in the torus, probably at locations of high velocity with asymmetry in minor loss coefficients, which cause nonzero  $U_{2,0}$  even when  $a_s$  is fully open. The velocity in the secondary cold heat exchanger is high enough to generate a significant  $\Delta p_{2,0}$ , but we believe there is very little asymmetry in its minor losses. The resonator junction, on the other hand, has both large velocities and significantly asymmetric minor losses. When  $\Delta \dot{Q}_h/c_p(T_h - T_c) = 2.9$  g/s, we estimate that the resonator junction generates  $\Delta p_{2,0} \approx 900$  Pa, which would drive roughly  $\rho_m U_{2,0} = 2.8$  g/s in the same direction as  $\rho_{m,c} \dot{W}_c/p_m$ . This would more than account for the 1.4 g/s difference between  $\rho_{m,c} \dot{W}_c/p_m$  and  $\dot{M}_{2, \text{open}}$  obtained from  $\Delta \dot{Q}_h$ . Given the uncertainty in the use of steady-flow minor loss correlations for oscillating flow, it seems likely that the resonator junction is responsible for the additional  $\dot{M}_{2, \text{open}}$ .

Next, the values of  $\dot{M}_{2, \text{open}}$  obtained from  $\Delta \dot{Q}_h$  are compared with those inferred from  $\Delta p_{\text{jp},0}$ , the time-averaged pressure drop required to force  $\dot{M}_2 = 0$ . Rewriting Eq. (19) in terms of  $\dot{M}_{2, \text{open}}$ , we find  $\dot{M}_{2, \text{open}} = [2\rho_m/R_0(\tau + 1)f]\Delta p_{\text{jp},0}$ . If the measured values of  $\Delta p_{\text{jp},0}$  are used in this expression, the result is the filled triangles in Fig. 7. Alternately, Eq. (22) and a DeltaE estimate of  $U_{1,\text{jp}}$  can be used to calculate the expected  $\Delta p_{\text{jp},0}$  which is then used to estimate  $\dot{M}_{2, \text{open}}$ . This result is shown as the open circles in Fig. 7. The discrepancy between the open circles and the filled triangles may be due to systematic errors in the measurement of  $\Delta p_{\text{jp}}$  or may be due to our imperfect knowledge of minor loss coefficients in oscillating flow. The cause of the even larger discrepancy between these estimates of  $\dot{M}_{2, \text{open}}$  and the solid line is unknown and a subject of future research. At  $\Delta \dot{Q}_h/c_p(T_h - T_c) = 2.9$  g/s, the  $\dot{M}_{2, \text{open}}$  obtained from the calculated  $\Delta p_{\text{jp},0}$  is 4.5 times higher than  $\Delta \dot{Q}_h/c_p(T_h - T_c)$ , roughly consistent with observations made by Swift *et al.*<sup>20</sup> in a similar device.

To summarize, when no measures are taken to suppress Gedeon streaming, the magnitude of  $\dot{M}_2 = \dot{M}_{2, \text{open}}$  is due to  $\rho_{m,c} \dot{W}_c/p_m$  and other generators of  $\Delta p_{2,0}$  within the torus. When these are taken into account, we find rough agreement between the calculated  $\dot{M}_{2, \text{open}}$  and  $\dot{M}_{2, \text{open}}$  measured from  $\Delta \dot{Q}_h$ . More precise comparisons are not possible due to a lack of experimental data on minor loss coefficients in oscillating flow. If  $\rho_{m,c} \dot{W}_c/p_m$  or the measured  $\dot{M}_{2, \text{open}}$  is used to estimate  $\Delta p_{\text{jp},0}$ , we find large, unexplained discrepancies with Eq. (22).

## B. Rayleigh (boundary-layer TBT) streaming

Olson and Swift<sup>8</sup> have shown that boundary-layer driven streaming in an orifice pulse tube refrigerator's pulse tube can convect significant heat,<sup>8</sup> and that an appropriately tapered pulse tube suppresses this streaming. The analogous component is the thermal buffer tube (TBT) in this engine, where we expect similar streaming effects. The streaming creates a time-averaged mass flux density about a thermal

penetration depth away from the TBT wall. Due to the viscosity of the gas, this annular sheet of mass-flux density drags the nearby gas along with it. The suppression of Gedeon streaming enforces zero total mass flux across the total cross section of the TBT. Therefore, the annular sheet of mass-flux density sets up a parabolic velocity profile across the TBT, with the gas near the center moving in the opposite direction to the gas at large radius. This flow pattern resembles an annular convection roll, and generates an unwanted heat leak by convecting heat from the hot end of the TBT to the ambient heat exchanger.

The boundary-layer analysis used in Ref. 8 should be valid if the oscillatory flow within the TBT is at low enough amplitude to be in the weakly turbulent regime, where turbulence is only generated outside the viscous and thermal boundary layers, leaving the velocity<sup>31</sup> and temperature profiles within the boundary layers nearly the same as in laminar flow. As  $p_{\text{ref}}/p_m$  increases, the flow at the ambient end of the TBT moves out of the weakly turbulent regime into the transitionally turbulent regime, and approaches the fully turbulent regime, as shown in Fig. 8(a). The solid lines in Fig. 8(a) represent the values of  $R/\delta_v$  and  $\text{Re}_D$  as a function of position within the TBT at the same values of  $p_{\text{ref}}/p_m$  as in Fig. 3. Here  $R$  is the local radius of the TBT and  $\text{Re}_D$  is the Reynolds number based on local peak velocity and TBT diameter. Larger values of  $p_{\text{ref}}/p_m$  are closer to the bottom of the figure. The hot end of the TBT is on the left, and the ambient end is on the right. The dashed lines indicate the boundaries between the various flow regimes. At  $p_{\text{ref}}/p_m = 0.10$ , much of the TBT lies within or very near the transitionally turbulent regime. The boundaries drawn between these various regimes are not sharp and were determined for incompressible flow.<sup>31</sup> The acoustic flow in the TBT is compressible, so the location of the boundaries in Fig. 8(a) falls into question. Therefore, at high amplitudes, turbulence may be generated in the boundary layer, which would invalidate the analysis of Ref. 8.

After optimizing a model of the engine using DeltaE, the resulting spatial dependencies of  $p_1$  and  $U_1$  are used to calculate the desired local taper angle via Eq. (12) of Ref. 8. The resulting ideal shape could be CNC machined into the inside surface of the TBT. For ease of manufacture, we have instead designed a two-segment taper that closely represents this ideal shape over a broad range of  $p_{\text{ref}}/p_m$ . The upper 8.0 cm is a straight-wall cylinder with a 8.9-cm inside diameter. The lower 16.0 cm is flared with a 1.35° half-angle taper.

To study the effects of the tapered TBT, we have performed measurements with the tapered TBT and with a thin-wall stainless-steel insert that converts the tapered TBT to a straight TBT with nearly the same inside diameter as the upper 8.0 cm of tapered TBT. With the resonator as the only load on the engine, the heat delivered to the secondary cold heat exchanger is measured as a function of amplitude. We subtract away the heat loads due to radiation (discussed below), thermal conduction (discussed below), minor loss at the secondary cold heat exchanger, and boundary-layer thermoacoustic transport along the TBT walls, and assume that the remaining heat load is due to streaming within the TBT. This subtraction ranges from about 30% of the total heat load

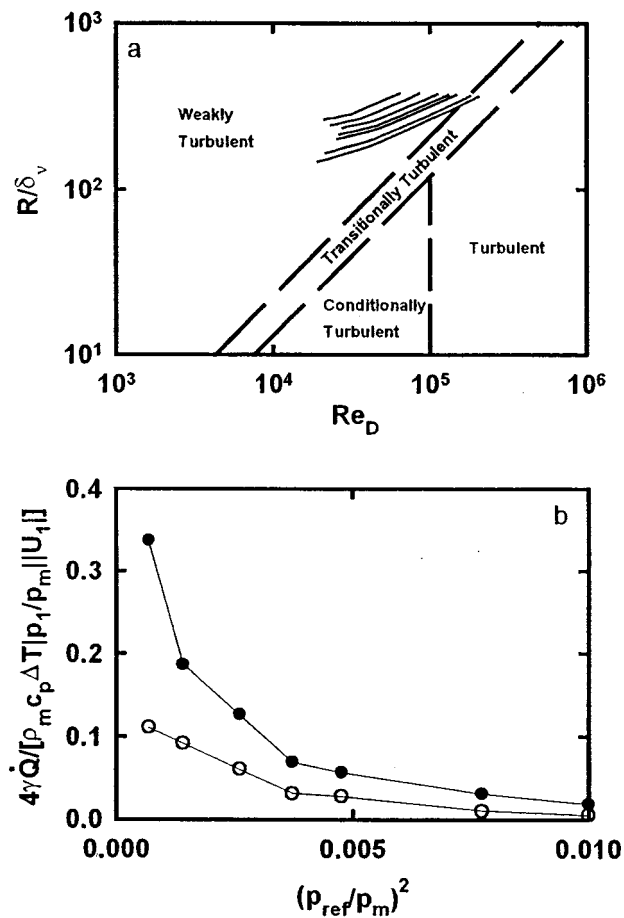


FIG. 8. (a) Flow regimes in the TBT as a function of  $p_{\text{ref}}/p_m$  and position within the TBT. The solid lines represent the values of  $R/\delta_v$  and  $Re_D$  within the TBT at the same values of  $p_{\text{ref}}/p_m$  as in Fig. 3. The curves are generated using the DeltaE model developed previously. Larger values of  $p_{\text{ref}}/p_m$  are closer to the bottom of the figure. The value of  $T_h$  is different for each curve and is given by the  $T_h$  necessary to create oscillations at the corresponding  $p_{\text{ref}}/p_m$  with the resonator as the only load on the engine. The hot end of the TBT is on the left; the ambient end is on the right. The dashed lines indicate the boundaries between various oscillatory flow regimes determined for incompressible flow (Ref. 31). The transition in the flow at the boundaries is not sharp. The boundary-layer analysis in Ref. 8 is probably valid only in the weakly turbulent regime. (b) Residual heat leak, normalized by  $\rho_m c_p \Delta T |p_1/p_m| |U_1| / 4\gamma$  vs  $(p_{\text{ref}}/p_m)^2$ . Open circles—tapered TBT; filled circles—straight TBT. Here,  $\Delta T = T_{h,\text{gas}} - T_c$ .  $p_1$  varies by  $\approx 1\%$  along the TBT and is taken to be the calculated  $p_1$  at the hot end of the TBT. Both  $\rho_m$  and  $|U_1|$  are taken to be the axial average of the calculated values in the TBT. Each data point is taken at the same values of  $p_{\text{ref}}/p_m$  and  $T_h$  as the corresponding curve in (a). If the boundary-layer analysis is valid, the normalized heat leak remains constant if  $p_1/U_1$  remains constant in the TBT as the amplitude is varied. As discussed in the text, the majority of the uncertainty in the measurement of the normalized heat leak is due to thermoacoustic effects on either side of the resonator junction.

at low amplitude to 80% of the total at high amplitude. To minimize the effect of acoustic power dissipated in other parts of the engine and resonator from being conducted or thermoacoustically transported back to the secondary cold heat exchanger, six turns of water-cooled, 1/4-in. copper tubing are wrapped around each of the ducts leading to the resonator and feedback inertance. This keeps the section of duct near the resonator junction near ambient temperature even at high  $p_{\text{ref}}/p_m$ . Without this tubing, this section of duct would approach  $50^\circ\text{C}$  at  $p_{\text{ref}}/p_m = 0.10$ .

Figure 8(b) shows the results of these measurements. The open and filled circles are the normalized residual heat leak to the secondary cold heat exchanger for the tapered and straight TBT, respectively. The heat leak is normalized such that if  $p_1/U_1$  in the TBT is constant as the amplitude is varied, the normalized heat leak would be constant and would depend only on the geometry of the TBT. This assumes that the boundary-layer analysis of Ref. 8 is valid. At the lowest amplitude, we expect a difference of 0.18 in normalized heat leak between the tapered and straight TBT based on the calculations of Ref. 8. This is roughly consistent with the measured difference of 0.23. As the amplitude increases, the difference between the straight and tapered TBTs decreases as well as the magnitude of the normalized heat leak. In absolute numbers for the tapered TBT, the residual heat leak is approximately 115 W and 36 W at  $p_{\text{ref}}/p_m = 0.061$  and 0.10, respectively, when  $T_{h,\text{gas}} = 725^\circ\text{C}$ . This amounts to 5% and 1% of the  $\dot{Q}_h$ , respectively.

Measurements of the temperature profile along the TBT wall are consistent with the dependence of the heat leak on amplitude. At low amplitude ( $p_{\text{ref}}/p_m = 0.026$ ), we expect the time-averaged mass-flux density at the wall to be directed downward from hot to cold, which would hold the upper half of the TBT wall near the temperature of the hot heat exchanger. In the lower half, the temperature should decrease sharply to the temperature of the secondary cold heat exchanger. This behavior is indeed observed in the tapered TBT at low amplitude. At  $p_{\text{ref}}/p_m = 0.026$ , the temperature at the axial midpoint of the TBT wall has decreased by only 20% of the total temperature difference spanned by the TBT. However, as the amplitude is increased, the temperature distribution becomes nearly linear with only small deviations from linearity near the ends of TBT. This indicates that the convective heat flux carried along the TBT wall is reduced.

There are several possibilities which are consistent with these observations. The first is that turbulence in the boundary layer destroys the velocity<sup>31</sup> and temperature profiles responsible for the streaming, making a straight TBT equivalent to a tapered one. As more of the TBT becomes turbulent at higher amplitude, the streaming is eliminated in more of the TBT, resulting in a decreasing normalized heat leak. Alternately, the turbulence may drive radial mixing of the gas in the TBT due to changes in the velocity profile through the acoustic cycle.<sup>31</sup> This radial mixing would cause heat exchange between the oppositely moving streaming flows and effectively “short out” the streaming heat load. As the amplitude increases and more of the TBT becomes turbulent, this mechanism would cause a decrease in normalized heat leak and make a straight TBT perform as well as a tapered one. Another possibility is that thermoacoustic effects on either side of the resonator junction are transporting heat into or out of the secondary cold heat exchanger. The additional heat flux would mask the effects of TBT boundary-layer streaming on the measurements, but it should not affect the temperature profile in the TBT which correlates well with the normalized heat leak. All of these explanations are merely conjectures and further research in this area is required to provide a definitive answer.

## V. THERMOACOUSTIC MEASUREMENTS

With Gedeon streaming completely suppressed and Rayleigh streaming reduced to an acceptable level, we now continue with basic thermoacoustic measurements. The general approach is to move sequentially through the engine from heat input to resonator dissipation, ultimately using exergy<sup>14</sup> to account quantitatively for all of the irreversibilities due to heat leaks and acoustic power dissipation. Some of these quantities are directly measured. Others require one step of calculation, typically using DeltaE,<sup>27</sup> after the measurement.

### A. Heat leaks

Not all of the heat input  $\dot{Q}_h$  to the hot heat exchanger is delivered to the thermoacoustic processes. Some is convected away by residual streaming that is not completely suppressed. The diagnostic measurements discussed above show that these residual convective heat leaks can be reduced to an acceptable level. Additional routes for heat leaks include the metal of the regenerator screen bed, surrounding pressure-vessel wall, and TBT wall; the ceramic fiber insulation<sup>32</sup> surrounding the TBT and regenerator; and radiation down the TBT. The heat leak caused by ordinary conduction in the helium gas is estimated to be small and is ignored in the rest of this discussion.

To determine the magnitude of these heat leaks, the engine is filled with 30-bar helium and heavily loaded with the variable acoustic load. This keeps the engine from oscillating, even at the highest operating temperature of 725 °C. Electric heat  $\dot{Q}_h$ , measured with a commercial wattmeter,<sup>18</sup> is applied at the hot heat exchanger and returns to ambient temperature  $T_c$  at the main cold heat exchanger or secondary cold heat exchanger, or through the ceramic fiber insulation, denoted  $\dot{Q}_{\text{main}}$ ,  $\dot{Q}_{\text{sec}}$ , and  $\dot{Q}_{\text{ins}}$ , respectively. The heats  $\dot{Q}_{\text{main}}$  and  $\dot{Q}_{\text{sec}}$  are determined by measuring the temperature rise and flow rate of the cooling water in the respective heat exchangers. The flow rate, which is measured with a commercial flowmeter, is restricted with a needle valve so that the temperature rise is easily measurable (5–15 °C). The heat passing through the ceramic fiber insulation is determined from the energy balance  $\dot{Q}_{\text{ins}} = \dot{Q}_h - (\dot{Q}_{\text{main}} + \dot{Q}_{\text{sec}})$ .

Before presenting the heat leak data, the spatial dependence of the hot temperature requires some discussion. With no acoustic oscillations present, the majority of  $\dot{Q}_h$  is carried by the TBT wall and pressure-vessel wall surrounding the regenerator, but  $\dot{Q}_h$  is injected evenly across the hot face of the regenerator. This leads to a radially outward flux of heat and a radial temperature dependence along the hot face of the regenerator. Therefore, the heat leaks along the TBT wall, regenerator pressure-vessel wall, and out through the ceramic fiber insulation are not driven by  $T_h$ , but by the temperature at the edge of the regenerator,  $T_{h,\text{edge}}$  [see Fig. 1(b)]. The other major sources of heat leak, conduction along the regenerator screen and radiation down the TBT, would be determined by integrating the equation for the local heat flux over the hot face of the regenerator. This leads to an expression for the total heat flux in terms of a spatial average of the hot temperature raised to some power, e.g.,  $\langle T^4(r) \rangle_r$  in the case

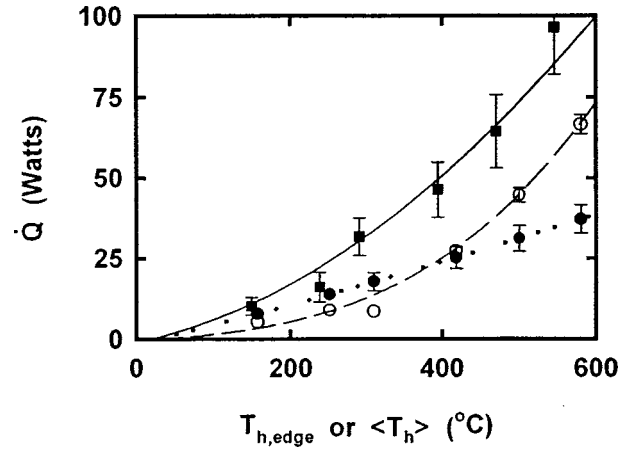


FIG. 9. Heat leaks vs  $T_{h,\text{edge}}$  or  $\langle T_h \rangle$ . The radiation heat leak  $\dot{Q}_{\text{rad}}$  (open circles) down the TBT and the conduction heat leak  $\dot{Q}_{\text{regen}}$  (filled circles) up the regenerator screen bed are given in terms of  $\langle T_h \rangle$ . The heat leak  $\dot{Q}_{\text{ins}}$  (filled squares) through the ceramic fiber insulation is given in terms of  $T_{h,\text{edge}}$ . The dashed, dotted, and solid lines are given by Eqs. (26), (27), and (28), respectively. For the lower temperature data, the error bars are smaller than the symbols.

of radiation. For the measured  $T_h$  and  $T_{h,\text{edge}}$ , the difference between  $\langle T^4(r) \rangle_r$  and  $\langle T(r) \rangle_r^4$  is not noticeable within the experimental error. Therefore, we will quote the radiation and screen-bed-conduction heat leaks in terms of the average hot temperature  $\langle T_h \rangle = \langle T(r) \rangle_r$ . The conduction heat leaks along the TBT wall, regenerator pressure-vessel wall, and through the ceramic fiber insulation will be given in terms of  $T_{h,\text{edge}}$ . With the acoustic oscillations present, the difference between  $T_h$  and  $T_{h,\text{edge}}$  is reduced due to mixing as the helium exits the hot heat exchanger and by the increased thermoacoustic heat loading.

The heat  $\dot{Q}_{\text{sec}}$  is due to radiation and metallic conduction down the TBT wall. Using the published thermal conductivity data for Inconel 625,<sup>17</sup> which are fit by  $K = K_0 + K_1 T$  (°C) with  $K_0 = 9.7 \text{ W/m}^\circ\text{C}$  and  $K_1 = 0.015 \text{ W/m}^\circ\text{C}^2$ , the metallic conduction can be reliably calculated as

$$\dot{Q}_{\text{TBT}}(T_{h,\text{edge}}) = \frac{(T_{h,\text{edge}} - T_c)[K_0 + K_1(T_{h,\text{edge}} + T_c)/2]}{l_{\text{st}}/A_0 + l_{\text{tap}} \ln(A_f/A_0)/(A_f - A_0)}, \quad (25)$$

where  $l_{\text{st}}$  and  $l_{\text{tap}}$  are the lengths of the straight and tapered sections and  $A_0$  and  $A_f$  are the cross-sectional areas of the TBT wall at the beginning and end of the taper. The TBT wall has a constant wall thickness for its entire length. The radiation heat leak is then extracted using  $\dot{Q}_{\text{rad}}(\langle T_h \rangle) = \dot{Q}_{\text{sec}} - \dot{Q}_{\text{TBT}}(T_{h,\text{edge}})$  and is shown as the open circles in Fig. 9. The appropriate  $\langle T \rangle^4$  temperature dependence is shown by  $\dot{Q}_{\text{rad}}$ , but the geometry of the hot zone is complicated and the calculation of the prefactor of  $\langle T \rangle^4$  (including emissivity) would be quite involved and not very enlightening. For comparison, the dashed line is given by

$$\dot{Q}_{\text{rad}} = 0.36\sigma S(\langle T_h \rangle^4 - T_c^4), \quad (26)$$

where  $\sigma$  is the Stephan–Boltzman constant,  $S$  is the area of the regenerator face, and the temperatures are expressed in Kelvin. In the absence of the TBT wall, Eq. (26) represents

36% of the total energy that would be radiated into the hemisphere seen by the hot face of the regenerator.

Turning our attention to  $\dot{Q}_{\text{main}}$ , the heat leak down the pressure-vessel wall is calculated from Eq. (25) by setting  $l_{\text{st}}=l$ =length of the regenerator and  $l_{\text{tap}}=0$ . Subtracting this from  $\dot{Q}_{\text{main}}$  leaves the heat leak down the regenerator screen bed  $\dot{Q}_{\text{regen}}$  which is given as the filled circles in Fig. 9.  $\dot{Q}_{\text{regen}}$  is expected to have a temperature dependence of the form

$$\dot{Q}_{\text{regen}}(\langle T_h \rangle) = \frac{(1-\phi)\epsilon_s S}{l} \left[ K_0(\langle T_h \rangle - T_c) + \frac{K_1}{1.68} (\langle T_h \rangle^{1.68} - T_c^{1.68}) \right], \quad (27)$$

where  $\epsilon_s$  is a factor which accounts for the poor thermal contact between the layers of the screen,<sup>33</sup> and  $S$  and  $\phi$  are the cross-sectional area and volume porosity of the regenerator. The thermal conductivity of the 304-stainless-steel screen<sup>34</sup> which forms the bed is given by  $K=K_0+K_1[T(^{\circ}\text{C})]^{0.68}$  where  $K_0=13.2$  W/m $^{\circ}\text{C}$ ,  $K_1=0.112$  W/m $^{\circ}\text{C}^{1.68}$ . Using  $\epsilon_s$  as an adjustable parameter, the data in Fig. 9 are best fit by  $\epsilon_s=0.13$  in reasonable agreement with the current literature.<sup>33</sup> The fit is shown as the dotted line in Fig. 9.

Again, the difference  $\dot{Q}_{\text{ins}}=\dot{Q}_h-(\dot{Q}_{\text{main}}+\dot{Q}_{\text{sec}})$  is assumed to pass through the insulation to ambient and is shown as the filled squares in Fig. 9. To calculate this heat leak, we assume linear temperature distributions along the inner surface of the insulation starting at  $T_{h,\text{edge}}$  at the hot heat exchanger and decreasing in both directions to  $T_c$  at the main and secondary cold heat exchangers. Also, we consider conduction only in the radial direction. Using the published thermal conductivity data<sup>32</sup> for the ceramic insulation packed at 12 lb/ft<sup>3</sup>, the heat leak takes the form

$$\dot{Q}_{\text{ins}}(T_{h,\text{edge}}) = \frac{\pi L_{\text{case}}(T_{h,\text{edge}}-T_c)}{\ln(R_o/R_i)} \times \left[ K_0 + K_1 \frac{T_{h,\text{edge}} + 2T_c}{3} \right], \quad (28)$$

where  $R_i$  and  $R_o$  are the inner and outer radii of the insulation and  $L_{\text{case}}$  is the distance from the main cold heat exchanger to the secondary exchanger. The thermal conductivity of the insulation is fit by  $K=K_0+K_1T(^{\circ}\text{C})$  with  $K_0=0.014$  W/m $^{\circ}\text{C}$  and  $K_1=0.17$  mW/m $^{\circ}\text{C}^2$ .<sup>32</sup> The result of Eq. (28) would fall well below the data in Fig. 9. Since the aspect ratio of the regenerator is near 1, the discrepancy is mostly likely due to axial conduction in the insulation around the regenerator. To compare the expected temperature dependence, the result of Eq. (28) is multiplied by 1.6 and plotted as the solid line in Fig. 9.

Although there are some discrepancies, the results of this section show that the overall heat leak and individual sources are reasonably well understood. The total uncertainty in the radiation and insulation heat leaks only amount to 3%–5% of  $\dot{Q}_h$  in the range  $p_{\text{ref}}/p_m=0.05$ –0.10. Therefore, if both Gedeon and Rayleigh streaming are suppressed, the amount of heat entering the thermoacoustic process is well

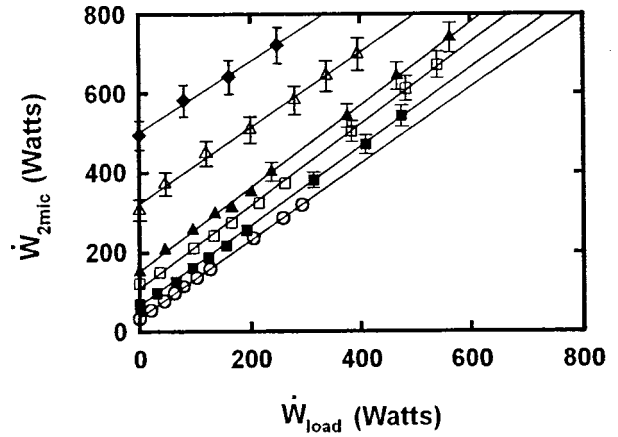


FIG. 10.  $\dot{W}_{2\text{mic}}$  vs  $\dot{W}_{\text{load}}$  for the same values of  $p_{\text{ref}}/p_m$  as in Fig. 3.  $p_{\text{ref}}/p_m=0.026$  has been left out for clarity. Acoustic amplitude is higher for larger values of  $\dot{W}_{2\text{mic}}$ . The solid lines are least-squares fits to the data. For all data except  $p_{\text{ref}}/p_m=0.10$ , the slopes are scattered in the range 0.96–1.05 consistent with Eq. (29). At  $p_{\text{ref}}/p_m=0.10$ , the slope is 0.89. This may be due to heating of the gas in the resonator causing errors in the physical parameters of the gas used in Eq. (23) of Ref. 16. The y-axis intercepts of the fits are used to extract  $\dot{W}_{\text{res}}$  and are summarized in Fig. 11. Typical uncertainties in  $\dot{W}_{2\text{mic}}$  range from  $\pm 2$  W at the lowest amplitude to  $\pm 45$  W at  $p_{\text{ref}}/p_m=0.10$ . At smaller values of  $\dot{W}_{2\text{mic}}$ , the error bars are smaller than the symbols. The uncertainty in  $\dot{W}_{\text{load}}$  is  $\pm 0.5\%$  and is independent of amplitude.

known. This section has emphasized heat leaks other than metallic conduction along the TBT and regenerator pressure-vessel wall, but one should not lose sight of the fact that these two sources account for over half of the total heat leak.

## B. Acoustic dissipation

Analysis of acoustic dissipation begins with the resonator where direct measurements are performed. The microphones placed in the resonator in Fig. 1(a) are used to make two-microphone measurements of the acoustic power flowing past their midpoint.<sup>16</sup> Since the resonator is a major contributor to the overall dissipation and two-microphone (two-mic) measurements can be sensitive to microphone placement, the technique is verified by comparing it with the power dissipated in the variable acoustic load  $\dot{W}_{\text{load}}$ . The load is placed at the middle port in the resonator, and by varying the load setting, different amounts of acoustic power are delivered to the load through the two-mic midpoint. If the two-mic measurements are accurate, the two-mic power is given by

$$\dot{W}_{2\text{mic}} = \dot{W}_{\text{res}} + \dot{W}_{\text{load}}, \quad (29)$$

where  $\dot{W}_{\text{res}}$  represents the dissipation in the resonator that occurs to the right of the two-mic midpoint. Accurate measurements of  $\dot{W}_{\text{load}}$  are easily made with the pair of microphones straddling the valve of the variable acoustic load in Fig. 1(a).<sup>16</sup> If  $p_{\text{ref}}/p_m$  is held constant (by varying  $\dot{Q}_h$ ) while  $\dot{W}_{\text{load}}$  is varied,  $\dot{W}_{\text{res}}$  is constant and a plot of  $\dot{W}_{2\text{mic}}$  vs  $\dot{W}_{\text{load}}$  should be a line of slope one whose intercept at  $\dot{W}_{\text{load}}=0$  is  $\dot{W}_{\text{res}}$  at  $p_{\text{ref}}/p_m$ . These data are presented in Fig. 10 for

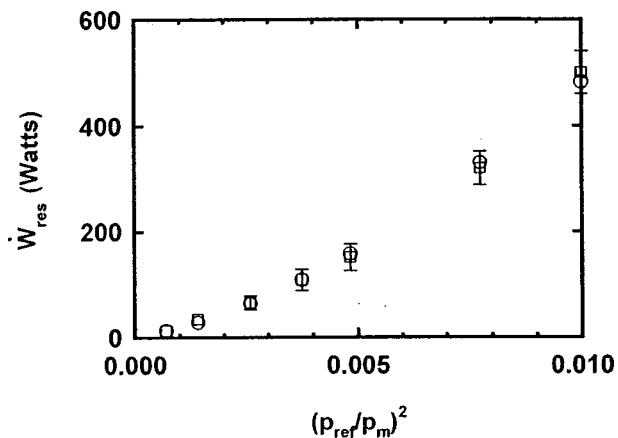


FIG. 11.  $\dot{W}_{res}$  measured (open squares) and calculated (open circles) vs  $(p_{ref}/p_m)^2$ . The DeltaE calculations use a relative surface roughness for the resonator of  $7 \times 10^{-5}$ . At low amplitudes, the error bars are smaller than the symbols.

values of  $p_{ref}/p_m$  ranging from 0.026 to 0.10. The solid lines are least-squares fits used to extract the slopes and  $\dot{W}_{res}$ . The slopes are scattered in the range 0.96–1.05 except at  $p_{ref}/p_m = 0.10$  where it takes the value 0.89. This low value may be due to resonator heating caused by the large dissipation of acoustic power.

The extracted values of  $\dot{W}_{res}$  are plotted as the open squares in Fig. 11. The data are clearly not linear in  $(p_{ref}/p_m)^2$  indicating enhanced dissipation in the resonator due to turbulence. This is to be expected because, as a function of  $p_{ref}/p_m$ , the peak Reynolds number in the resonator ranges from 0.4 to  $1.5 \times 10^6$  and the steady-flow friction factor is nearly constant over this range.<sup>29</sup> At such large Reynolds numbers, the roughness of the resonator's inner surface is expected to play a significant role in the resonator dissipation. The open circles in Fig. 11 are DeltaE calculations<sup>27</sup> of  $\dot{W}_{res}$  using a relative surface roughness of  $7 \times 10^{-5}$  which corresponds to an absolute roughness of approximately 7  $\mu\text{m}$ . This is considerably larger than the estimated 2- $\mu\text{m}$  roughness of the metal surface. A DeltaE calculation of  $\dot{W}_{res}$  using the estimated surface roughness of 2  $\mu\text{m}$  results in only 30 W less dissipation at  $p_{ref}/p_m = 0.10$ , or less than 6% of the total measured dissipation. This extra dissipation could also be due to unknown minor losses in the resonator. Given the approximate nature of DeltaE's turbulence algorithm and the uncertainty in using steady-flow minor loss correlations for oscillating flow, the discrepancy between the estimated surface roughness and that used in the DeltaE calculations is beyond the limit of our current knowledge.

Using this single 7- $\mu\text{m}$  surface roughness for all acoustic amplitudes, we obtain good agreement between the calculated and measured  $\dot{W}_{res}$ . Therefore, we are confident that DeltaE can be used to account for the dissipation between the resonator junction and the two-mic midpoint. This small additional dissipation is added to  $\dot{W}_{res}$ , and for the rest of this article, the acoustic power flow into the resonator is reported in terms of acoustic power passing the resonator junction  $\dot{W}_{junc}$ .

Next, we consider the acoustic power dissipated in the

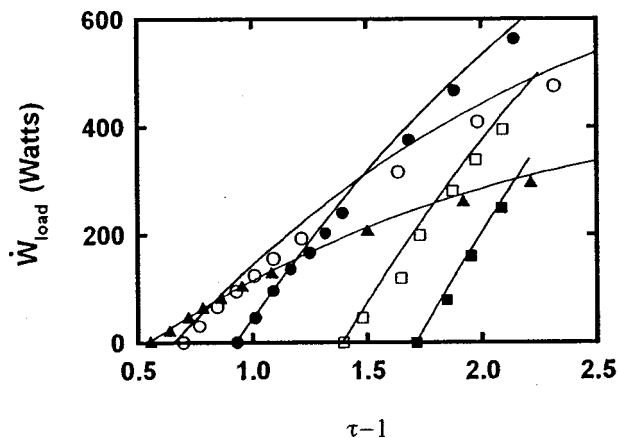


FIG. 12.  $\dot{W}_{load}$  vs  $(\tau - 1)$ . Measured values are given by the symbols and values calculated with DeltaE are shown as solid lines. Acoustic amplitudes are the same as those in Fig. 3 except that  $p_{ref}/p_m = 0.026$  and 0.061 have been left out for clarity. Acoustic amplitude increases from left to right along the  $(\tau - 1)$  axis. Agreement between the DeltaE model and the measurements is evidence that the overall dissipation in the engine is correctly accounted for in the model. The uncertainty in  $\dot{W}_{load}$  is  $\pm 0.5\%$  and is independent of amplitude and  $(\tau - 1)$ . Typical uncertainty in  $(\tau - 1)$  is  $\pm 1\%$ .

torus where direct measurement of the dissipation is far more difficult. The major contributors to the dissipation are viscous and thermal losses in the regenerator, inertance, and compliance; and minor losses in the inertance, compliance, jet pump, and secondary cold heat exchanger. The main cold heat exchanger, hot heat exchanger, and TBT do not dissipate a significant amount of acoustic power, but the lost energy due to irreversible heat transfer<sup>14</sup> at the heat exchangers will be taken into account in the next section. The previous discussion of the pressure-drop phasors in the torus showed that the DeltaE model makes fairly accurate predictions of the real and imaginary parts of the pressure drops across the jet pump, regenerator, and inertance. Therefore, we will use DeltaE to compute the acoustic power dissipation in the torus elements.

### C. Performance

Before using the DeltaE model to assign irreversibilities to various components, we begin this section with a summary of how the model was built up. There are two adjustable parameters, the surface roughnesses of the resonator and feedback inertance. Using the measurements of  $\dot{W}_{res}$ , a single surface roughness is assigned to the resonator which reproduces the data at all values of  $p_{ref}/p_m$  (see Fig. 11). Then, this roughness is used to extrapolate from the two-mic midpoint to the resonator junction and calculate the acoustic power flowing out of the torus  $\dot{W}_{junc}$ . The values of  $\dot{W}_{junc}$  are then used to fix the acoustic power generated by a DeltaE model of the torus. Using a single surface roughness in the feedback inertance, the torus model reproduces the pressure-drop phasors for both the jet pump and regenerator (see Fig. 4). In addition to the regenerator dissipation, the resonator and feedback dissipation provide the majority of the load on the engine, and therefore play a large role in setting  $T_h$ . As a final check of the DeltaE model, Fig. 12 presents the calculated and measured load curves, i.e.,  $\dot{W}_{load}$  vs  $\tau - 1$ , for the

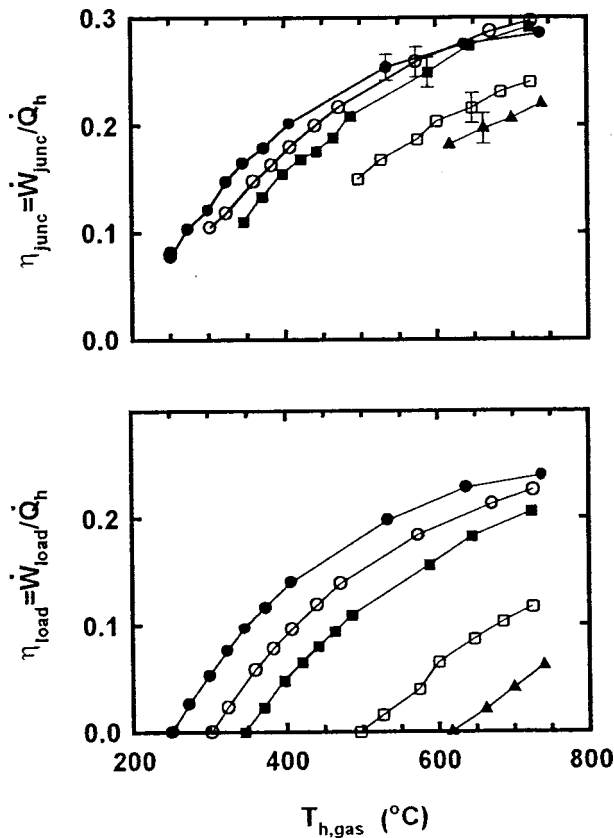


FIG. 13. Thermal efficiencies vs  $T_{h,gas}$ , the temperature of the helium immediately below the hot heat exchanger, for the same values of  $p_{ref}/p_m$  as in Fig. 3.  $p_{ref}/p_m=0.026$  and  $0.038$  have been left out for clarity. Acoustic amplitude increases from left to right. The thermal efficiency is measured in two ways, i.e., in terms of the acoustic power delivered past the resonator junction ( $\eta_{junc}$ ) or the acoustic power delivered to the load ( $\eta_{load}$ ). The single error bar shown on each constant-amplitude  $\eta_{junc}$  dataset is typical of the uncertainty of every measurement in the dataset. Typical uncertainties in  $\eta_{load}$  and  $T_{h,gas}$  are both  $\pm 1\%$ .

same values of  $p_{ref}/p_m$  as in Fig. 3. The good agreement provides further confidence in the ability of the DeltaE model to account for the various irreversibilities.

We now present the thermal efficiency of the engine at various acoustic amplitudes as a function of  $T_{h,gas}$ , the temperature of the helium immediately below the hot heat exchanger.  $T_{h,gas}$  is used instead of  $T_h$  because it more closely represents the temperature which limits a practical engine due to weakening of materials in the hot zone. Figure 13 presents the engine efficiency in two ways. If the assembly is to be used in its present configuration where it is driving a load attached to the resonator, the relevant efficiency is  $\eta_{load} = \dot{W}_{load}/\dot{Q}_h$ . If it is scaled up to a size where the resonator dissipation is a small fraction of the total dissipation or if it is one of many engines driving a single resonator, the relevant efficiency is  $\eta_{junc} = \dot{W}_{junc}/\dot{Q}_h$ . We will focus on the latter, as our future plans include multiple scaled-up engines on a single large resonator.

The  $\eta_{junc}$  data show that for  $p_{ref}/p_m \approx 0.06$  and  $T_{h,gas} \approx 725$  °C, the engine is converting  $\dot{Q}_h$  to acoustic power with a conversion efficiency of 0.30. Expressed in a different way, the engine is operating with 41% of the theoretical maximum efficiency, the Carnot efficiency. Here, the Carnot

efficiency is based on the estimated temperature of the Ni–Cr in the hot heat exchanger and the average temperature of the cooling water in the main cold heat exchanger. This represents more than a 50% increase over the best standing-wave thermoacoustic engines.<sup>12</sup> At this operating point, the engine is delivering 710 W past the resonator junction. As the acoustic amplitude is increased to  $p_{ref}/p_m=0.10$  where the power density is higher,  $\eta_{junc}$  falls off to 0.22 at  $T_{h,gas}=725$  °C with the engine delivering 890 W past the resonator junction.

Next, we use a combination of direct measurements and DeltaE calculations to determine where the major irreversibilities occur and why  $\eta_{junc}$  falls off at higher acoustic amplitude. To place all irreversibilities on an equal footing, we use the concept of exergy or available work,<sup>14</sup> which is a measure of the work-equivalent “value” of energy when a thermal reservoir at ambient temperature is freely available. Heat  $\dot{Q}$  at temperature  $T$  carries exergy  $\dot{X} = \dot{Q}(1 - T_0/T)$ , where  $T_0$  is the ambient temperature. A thermoacoustic wave carries  $\dot{X} = (T_0/T)\dot{W} + (1 - T_0/T)\dot{H}$ ,<sup>35</sup> where  $\dot{H}$  is the total power flow.<sup>11</sup> In a wide duct  $\dot{X} = \dot{W}$ . In an ideal regenerator ( $r_h \ll \delta_\kappa$ ),  $\dot{H} = 0$ , so  $\dot{X} = (T_0/T)\dot{W}$ . In a realistic regenerator,  $\dot{H}$  must be taken into account. In a thermoacoustic-Stirling engine, there are three basic mechanisms for destroying exergy: heat flow  $\dot{Q}$  across a finite temperature difference ( $T_1 \rightarrow T_2$ ) with  $\dot{X}_{lost} = \dot{Q}T_0(T_1 - T_2)/(T_1T_2)$ ; dissipation of acoustic power  $\dot{W}_{diss}$  at temperature  $T$  with  $\dot{X}_{lost} = (T_0/T)\dot{W}_{diss}$ ; and regenerator ineffectiveness which is a combination of acoustic power dissipation at temperature  $T$  and “thermoacoustic heat pumping” due to  $\dot{H} \neq 0$ . In the last case,  $\dot{X}_{lost} = \dot{X}_h - \dot{X}_c$  where  $\dot{X}_h$  and  $\dot{X}_c$  are the thermoacoustic exergy fluxes at the hot and cold faces of the regenerator.

Table II lists different elements of the engine, what physical process causes the loss of exergy, how it was determined, and what fraction of the total input exergy this process is responsible for destroying. The final two entries in the table list the total input exergy and what fraction cannot be accounted for in the analysis. The use of exergy allows direct comparison of the importance of different irreversibilities. The increase in amplitude from  $p_{ref}/p_m=0.061$  to 0.10 shows several expected trends. Losses due to conduction heat leaks decrease in their relative importance as the acoustic power density grows. The losses in the regenerator increase, but not as fast as  $(p_{ref}/p_m)^2$ . Since the Reynolds number in the regenerator is low, the dissipation is proportional to a weighted average of  $|U_1|^2$  within the regenerator. As  $p_{ref}/p_m$  increases,  $|U_{1,c}|$  remains roughly constant due to the amplitude dependent  $R_{jp}$ . However,  $|U_{1,h}|$  increases with amplitude due to the compliance of the regenerator, and the magnitude of the loss scales roughly as  $(|U_{1,c}|^2 + |U_{1,h}|^2)/2$ . Losses in the feedback inductance grow in their relative importance because, at these amplitudes, the minor losses and turbulence cause the acoustic dissipation to grow as  $(p_{ref}/p_m)^3$ . Also, the relative importance of residual streaming in the TBT seems to decrease, but the cause of this is presently unknown.

Table II also shows the areas where basic research into the loss mechanisms can provide the fastest improvements in

TABLE II. Major causes of lost exergy, sorted according to engine element and physical process. The losses are either directly measured or calculated using DeltaE. The DeltaE model is built up according to the principles outlined at the beginning of Sec. V C. The final two entries in the table are the total input exergy at the hot heat exchanger and what fraction of this cannot be accounted for in the analysis.

Element	Process	Method	$p_{\text{ref}}/p_m = 0.061$ $T_{h,\text{gas}} = 725\text{ }^\circ\text{C}$		$p_{\text{ref}}/p_m = 0.10$ $T_{h,\text{gas}} = 725\text{ }^\circ\text{C}$	
			$\dot{X}_{\text{lost}}(W)$	Fraction	$\dot{X}_{\text{lost}}(W)$	Fraction
Regenerator	Viscous/Thermal loss	DeltaE	238	0.14	393	0.13
	Heat leak	Measured	163	0.09	172	0.06
Feedback inertance	Minor loss	DeltaE	62	0.04	296	0.10
	Viscous/Thermal loss	DeltaE	36	0.02	145	0.05
Thermal buffer tube	Residual streaming	Measured	82	0.05	25	0.01
	Radiation	Measured	69	0.04	78	0.03
	Metallic conduction	Measured	52	0.03	55	0.02
	Boundary-layer transport	DeltaE	7	<0.01	18	<0.01
	Flow straightener	DeltaE	2	<0.01	10	<0.01
	Heat leak	Measured	82	0.05	87	0.03
Insulation	Heat leak	Measured	82	0.05	87	0.03
	Temperature difference	Measured	57	0.03	181	0.06
Main cold heat exchanger	Viscous loss	DeltaE	4	<0.01	12	<0.01
	Temperature difference	Measured	43	0.02	114	0.04
Hot heat exchanger	Viscous loss	DeltaE	4	<0.01	11	<0.01
	Temperature difference	Measured	43	0.02	114	0.04
Sec. cold heat exchanger	Minor/Viscous loss	DeltaE	34	0.02	144	0.05
Jet pump	Minor/Viscous loss	DeltaE	30	0.02	130	0.04
Resonator and load	Delivered power	Measured	710	0.41	890	0.30
Input exergy	$(1 - T_{\text{hhx}}/T_{\text{water}})\dot{Q}_h$	Measured	1724	1.00	2968	1.00
Unaccounted $\dot{W}_{\text{lost}}$			44	0.03	200	0.07

efficiency. The screen-bed regenerator used in this engine could be replaced by a parallel-plate regenerator which would have lower viscous losses while still providing the necessary regenerative heat exchange. The smaller viscous resistance would also allow the feedback inertance and compliance to be redesigned in such a way that viscous and minor losses in those parts would be reduced. The technical challenge is the construction of parallel plates with uniform spacing on the order of  $75\text{ }\mu\text{m}$  and an acceptable level of heat leak. Also, basic research into minor losses in oscillating flows might provide insight into how these losses occur and methods to reduce them.

## VI. CONCLUSION

We have constructed a new type of thermoacoustic engine that is based on traveling-wave phasing and a Stirling-like thermodynamic cycle. In designing and building this engine, we have followed the basic traveling-wave principle described by Ceperley,<sup>4,6</sup> but with many improvements and additional elements that are necessary to achieve high efficiency. Our engine uses an ideally reversible acoustic network, which is much shorter than  $1/4$  wavelength, to feed acoustic power to the cold end of the regenerator and convert standing-wave phasing in the acoustic resonator to the necessary traveling-wave phasing while maintaining  $|p_{1,c}/U_{1,c}| \gg \rho c/A$ . In comparison with Ceperley, our network naturally creates a large value of  $|p_{1,c}/U_{1,c}|$  without large changes in duct diameter. To achieve the high efficiencies reported in this article, two types of acoustic streaming that were not foreseen by Ceperley have been suppressed. Gedeon streaming is eliminated by the use of a jet pump that relies on the asymmetry of hydrodynamic end effects, and Rayleigh streaming is reduced by appropriately tapering the TBT.

Measurements show that the first-order dynamics in the torus section of the engine are well understood, but that we have only a qualitative understanding of many of the nonlinear effects in the engine. To allow for quantitative design in the future, several areas require further research. Basic research into minor losses in oscillating flow is required to place the incorporation of these losses on firmer ground. The magnitude of Gedeon streaming needs further study so that jet pumps, or other means of suppressing this time-averaged flow, can be appropriately designed. Also, more careful measurements on Rayleigh streaming in the thermal buffer tube are required to determine how turbulence modifies this type of streaming.

At the most efficient operating point, the engine delivers 710 W to its resonator with an efficiency of 0.30 which corresponds to 41% of the Carnot efficiency. At the most powerful operating point, the engine delivers 890 W to its resonator with an efficiency of 0.22. We have identified the irreversible processes and sources of lost work. This analysis suggests the path of future research on thermoacoustic-Stirling hybrid engines. Although some of the phenomena are not yet fully understood, and further improvement is possible, the efficiency is already excellent: more than a 50% increase over the best standing-wave thermoacoustic engines.

## ACKNOWLEDGMENTS

We thank D. L. Gardner, C. Espinoza, and R. Rockage for their expert assistance in the construction of the engine, and the Office of Basic Energy Sciences in the U.S. DOE for financial support.



- <sup>1</sup>G. Walker, *Stirling Engines* (Clarendon, Oxford, 1960).
- <sup>2</sup>R. Radebaugh, "A review of pulse tube refrigeration," *Adv. Cryog. Eng.* **35**, 1191–1205 (1990).
- <sup>3</sup>C. D. West, *Liquid Piston Stirling Engines* (Van Nostrand Reinhold, New York, 1983).
- <sup>4</sup>P. H. Ceperley, "A pistonless Stirling engine—the traveling wave heat engine," *J. Acoust. Soc. Am.* **66**, 1508–1513 (1979).
- <sup>5</sup>T. Yazaki, A. Iwata, T. Maekawa, and A. Tominaga, "Traveling wave thermoacoustic engine in a looped tube," *Phys. Rev. Lett.* **81**, 3128–3131 (1998).
- <sup>6</sup>P. H. Ceperley, "Gain and efficiency of a short traveling wave heat engine," *J. Acoust. Soc. Am.* **77**, 1239–1244 (1985).
- <sup>7</sup>David Gedeon, "DC gas flows in Stirling and pulse-tube cryocoolers," in *Cryocoolers 9*, edited by R. G. Ross (Plenum, New York, 1997), pp. 385–392.
- <sup>8</sup>J. R. Olson and G. W. Swift, "Acoustic streaming in pulse tube refrigerators: Tapered pulse tubes," *Cryogenics* **37**, 769–776 (1997).
- <sup>9</sup>Y. L. Ju, C. Wang, and Y. Zhou, "Dynamic experimental study of the multi-bypass pulse tube refrigerator with two-bypass tubes," *Adv. Cryog. Eng.* **43**, 2031–2037 (1998).
- <sup>10</sup>A. Ravex, J. M. Poncet, I. Charles, and P. Bleuze, "Development of low frequency pulse tube refrigerators," *Adv. Cryog. Eng.* **43**, 1957–1964 (1998).
- <sup>11</sup>G. W. Swift, "Thermoacoustic engines," *J. Acoust. Soc. Am.* **84**, 1145–1180 (1988).
- <sup>12</sup>To our knowledge, the highest thermal efficiency in a standing-wave engine is 0.18, reported by C. Jin, based on acoustic power delivered from the engine to the resonator, using the apparatus described by K. M. Godshalk *et al.*, *Adv. Cryog. Eng.* **41**, 1411–1418 (1996). Comparable efficiency has been obtained with the much larger standing-wave engine at Cryenco, Inc. described in <http://lib-www.lanl.gov/la-pubs/00412750.pdf>.
- <sup>13</sup>S. Backhaus and G. W. Swift, "A thermoacoustic-Stirling heat engine," *Nature (London)* **399**, 335–338 (1999).
- <sup>14</sup>Adrian Bejan, *Advanced Engineering Thermodynamics*, 2nd ed. (Wiley, New York, 1997).
- <sup>15</sup>A. J. Organ, *Thermodynamics and Gas Dynamics of the Stirling Cycle Machine* (Cambridge University Press, Cambridge, U.K., 1992).
- <sup>16</sup>A. M. Fusco, W. C. Ward, and G. W. Swift, "Two-sensor power measurements in lossy ducts," *J. Acoust. Soc. Am.* **91**, 2229–2235 (1992).
- <sup>17</sup>W. D. Klopp, Inconel Alloy 625. *Aerospace Structural Metals Handbook*, Code 4117, July 1993, Metals and Ceramics Information Center (1997).
- <sup>18</sup>Ohio Semitronics, Inc., 4242 Reynolds Dr., Hilliard, OH 43026.
- <sup>19</sup>I. E. Idelchik, *Handbook of Hydraulic Resistance*, 3rd ed. (Begell House, New York, 1996).
- <sup>20</sup>G. W. Swift, D. L. Gardner, and S. Backhaus, "Acoustic recovery of lost power in pulse tube refrigerators," *J. Acoust. Soc. Am.* **105**, 711–724 (1999).
- <sup>21</sup>M. Iguchi, M. Ohmi, and K. Maegawa, "Analysis of free oscillating flow in a U-shaped tube," *Bull. JSME* **25**, 1398 (1982).
- <sup>22</sup>Endevco Corporation, 30700 Rancho Viejo Road, San Juan Capistrano, CA 92675.
- <sup>23</sup>L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982).
- <sup>24</sup>W. M. Kays and A. L. London, *Compact Heat Exchangers* (McGraw-Hill, New York, 1964).
- <sup>25</sup>P. Kittel, "Ideal orifice pulse tube refrigerator performance," *Cryogenics* **32**, 843–844 (1992).
- <sup>26</sup>G. W. Swift and W. C. Ward, "Simple harmonic analysis of regenerators," *J. Thermophys. Heat Transfer* **10**, 652–662 (1996).
- <sup>27</sup>W. C. Ward and G. W. Swift, "Design environment for low amplitude thermoacoustic engines (DeltaE)," *J. Acoust. Soc. Am.* **95**, 3671–3672 (1994). Fully tested software and users guide available from Energy Science and Technology Software Center, US Department of Energy, Oak Ridge, Tennessee. To review DeltaE's capabilities, visit the Los Alamos thermoacoustics web site at <http://www.lanl.gov/thermoacoustics/>. For a beta-test version, contact [ww@lanl.gov](mailto:ww@lanl.gov) (Bill Ward) via Internet.
- <sup>28</sup>To obtain the DeltaE input file used to model the engine, visit <http://www.lanl.gov/projects/thermoacoustics/Tashe.out>.
- <sup>29</sup>R. W. Fox and A. T. McDonald, *Introduction to Fluid Mechanics* (Wiley, New York, 1985).
- <sup>30</sup>V. L. Streeter, *Handbook of Fluid Dynamics* (McGraw-Hill, New York, 1961).
- <sup>31</sup>M. Ohmi and M. Iguchi, "Critical Reynolds number in an oscillating pipe flow," *Bull. JSME* **25**, 165 (1982).
- <sup>32</sup>Unifrax Corp., 2351 Whirlpool St., P.O. Box 156, Niagara Falls, NY 14302.
- <sup>33</sup>M. A. Lewis, T. Kuriyama, F. Kuriyama, and R. Radebaugh, "Measurement of heat conduction through stacked screens," *Adv. Cryog. Eng.* **43**, 1611–1618 (1998).
- <sup>34</sup>C. F. Hickey. 304 Stainless steel. *Aerospace Structural Metals Handbook*, Code 1303, March 1973, Metals and Ceramics Information Center (1997).
- <sup>35</sup>G. W. Swift, *Thermoacoustics: A unifying perspective for some engines and refrigerators* (1999 draft available at <http://www.lanl.gov/thermoacoustics/Book/index.html>).

# Dynamics of gas bubbles in viscoelastic fluids. I. Linear viscoelasticity

John S. Allen

*Department of Biomedical Engineering, One Shields Avenue, University of California, Davis, Davis, California 95616*

Ronald A. Roy

*Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215*

(Received 17 May 1999; accepted for publication 13 March 2000)

The nonlinear oscillations of spherical gas bubbles in linear viscoelastic fluids are studied. A novel approach is implemented to derive a governing system of nonlinear ordinary differential equations. The linear Maxwell and Jeffreys models are chosen as the fluid constitutive equations. An advantage of this new formulation is that, when compared with previous approaches, it facilitates perturbation methods and numerical investigations. Analytical solutions are obtained using a multiple scale perturbation method and compared with the Newtonian results for various Deborah numbers. Numerical analysis of the full equations supports the perturbation analysis, and further reveals significant differences between the viscoelastic and Newtonian cases. Differences in the oscillation phase and harmonic structure characterize some of the viscoelastic effects. Subharmonic excitations at particular fluid parameters lead to a discrete group modulation of the radial excursions; this appears to be a unique, previously undiscovered phenomenon. Implications for medical ultrasound applications are discussed in light of these current findings. © 2000 Acoustical Society of America. [S0001-4966(00)04606-3]

PACS numbers: 43.35.Wa, 43.25.Yw, 43.80.Gx [HEB]

## LIST OF SYMBOLS

$\lambda_1$	relaxation time	$\bar{a} = p_a/p_0$	ratio of pressure amplitudes
$\lambda_2$	retardation time	$De = \lambda_1/R_0\sqrt{\rho/p_{g0}}$	Deborah number (perturbation analysis)
$\dot{\epsilon}$	rate of strain or deformation tensor	$\lambda = \lambda_2/\lambda_1$	ratio of time fluid time constants
$\eta_0$	zero shear-rate viscosity	$x$	relative amplitude in perturbation analysis
$\kappa$	polytropic exponent	$a$	real part of amplitude in perturbation analysis
$S(R, t) = \int_R^\infty (\tau_{rr}(r, t) dr/r)$	spatial integral over radial stress	$\omega_0^2 = 3\kappa - w$	detuning parameter (perturbation analysis)
$\bar{r} = r/R_0$	nondimensional radial coordinate	$\sigma$	nondimensional time (numerical analysis)
$\bar{R} = R/R_0$	nondimensional bubble radius	$\bar{t} = \omega t$	Deborah number (numerical analysis)
$\bar{t} = t/R_0\sqrt{\rho/p_{g0}}$	nondimensional time (perturbation analysis)	$De = \lambda_1\omega$	Deborah number (numerical analysis)
$\bar{\tau} = \tau/\eta_0 R_0\sqrt{\rho/p_{g0}}$	nondimensional radial stress	$Re = \rho\omega R_0^2/\eta_0$	Reynolds number (numerical analysis)
$w = 2\hat{\sigma}/p_{g0}R_0$	nondimensional Laplace pressure		
$\Omega = R_0\sqrt{\rho/p_{g0}}\omega$	ratio of forcing frequencies		
$Re = \eta_0/(p_{g0}\rho)^{1/2}R_0$	Reynolds number (perturbation analysis)		

## INTRODUCTION

The nonlinear oscillations of an isolated, acoustically forced gas bubble have been the subject of extensive research.<sup>1,2</sup> Most of these studies have concentrated on the behavior of gas bubbles in Newtonian fluids. However, there is a growing interest in the nature of gas bubble oscillations in non-Newtonian fluids. In part, this interest has been motivated by the increasingly important role of cavitation

and associated bubble dynamics in medical ultrasound applications,<sup>3,4</sup> where biological fluids and tissues often exhibit non-Newtonian behavior.<sup>5,6</sup> Unfortunately, a complete and comprehensive study of the various rheological properties of tissue is still far from complete. Nevertheless, bubbles are proving to be valuable tools for medical diagnosis as evidenced by the emergence of ultrasound contrast agents, which usually take the form of microbubbles introduced into

a patient's bloodstream to increase ultrasound scattering for enhanced image contrast.<sup>7</sup> Another biomedical area of growing importance is the monitoring and assessment of possible cavitation bioeffects from medical ultrasound;<sup>8</sup> however, there remains some debate as to which modalities and parameter ranges of medical ultrasound might produce cavitation bioeffects. Within this context, one of the limitations of current research efforts is the lack of a model for bubble dynamics that properly accounts for tissue rheology. Previous theoretical efforts at determining cavitation bioeffect criteria have been based on an understanding of Newtonian bubble dynamics,<sup>8</sup> a fact that has complicated the detailed, quantitative comparison of theoretical predictions and *in vivo* studies. Thus, there exists a need to account for the impact of non-Newtonian rheology on the dynamics of bubbles in biomedical systems.

The scope of this problem, however, extends well beyond the aforementioned biomedical applications. Research topics as diverse as the acoustic scattering from oceanic microbubble-cloud targets, cavitation erosion suppression studies, cavitation dynamics in food processing, decompression sickness and sonochemistry all involve bubbles in non-Newtonian fluids to one extent or another. The study of bubbles in non-Newtonian polymeric solutions is particularly germane to polymer processing and other industrial applications. Bubble behavior in polymeric solutions has been modeled in many previous studies using empirical shear-thinning constitutive relationships.<sup>9</sup> This approach has been employed for acoustically forced bubble oscillations. In these studies, the only non-Newtonian effect considered is shear thinning (viscosity modeled as a decreasing function of the shear stress). A less understood, and perhaps more profound non-Newtonian effect is viscoelasticity, a feature that is particularly relevant to many biomaterials.<sup>5,6</sup> From the physical point of view, dilation accompanying spherical bubble oscillations "stretches" the fluid, a scenario in which viscoelastic effects are important. Hence viscoelasticity, and its effect on spherical bubble dynamics, is the focus of this paper.

In 1970, Fogler and Goddard published a groundbreaking study of the free oscillations of a spherical bubble in a viscoelastic fluid.<sup>10</sup> For simplicity, they modeled the bubble as an empty void and used the linear Maxwell model as the constitutive equation for the surrounding fluid. They mathematically combined this constitutive equation with a Rayleigh-Plesset bubble dynamics formulation<sup>10</sup> in deriving an integro-differential equation for the radial dynamics. This equation was examined in certain asymptotic limits and solved numerically for a few cases. Their results suggested that for certain parameter ranges, the elasticity of the fluid retards the bubble collapse. However, numerical difficulties limited their integration to only a few cycles of bubble oscillation. Moreover, their choice of a void with no internal gas pressure constrains the physical interpretation of their results. Following the work of Fogler and Goddard, Tanasawa and Yang, and later Ting, studied a freely oscillating bubble using a three-constant Oldroyd constitutive equation.<sup>11,12</sup> Numerical difficulties in solving the resulting integro-differential equation also limited the scope of these

investigations. More recently, Kim undertook a study of the free oscillations of a non-Newtonian bubble with an objective form of the Maxwell model, solving the  $r$  and  $\theta$  components of the continuity, momentum and constitutive equations simultaneously using a finite-element method.<sup>13</sup> Kim's effort proved to be too computationally intensive to admit a comprehensive study of bubble behavior for a wide range of parameter values.

The investigation of acoustically forced bubble oscillations has received less attention. Shima and co-workers examined an acoustically forced bubble with the three-constant Oldroyd model<sup>14</sup> for the fluid. Their study provides some intriguing results; however, it was a pioneering study and thus left many fundamental questions unanswered.

One advantage of using the "objective" constitutive equations as opposed to a linear viscoelastic formulation is that the former is intrinsically valid for large deformations of the fluid. The term "objective" refers to the property of being invariant under changes in reference frame; observers subject to relative motion between each other see the same stress on a body. Furthermore, it has been demonstrated that some of these equations originate from molecular level considerations as opposed to empirical, macroscopic assumptions.<sup>15</sup> Conversely, linear viscoelastic equations originate mostly from empirical arguments. Nevertheless, even for linear viscoelastic fluids, bubble dynamics is poorly understood and, in particular, the behavior of acoustically forced bubbles remains relatively unexplored. This is due, in part, to the numerical difficulties associated with solving Fogler and Goddard's radial viscoelastic bubble equation. (More recently, researchers have attempted to compare their results from a more complicated constitutive equation with this original work.<sup>13</sup>) A more comprehensive study of bubble dynamics in linear viscoelastic fluids is justified, particularly as a prelude to tackling the complications associated with the adoption of objective constitutive relationships.

Thus this study will demonstrate that the linear viscoelastic bubble dynamics equations can be obtained using a novel approach offering several advantages over previous formulations. The governing equations are a system of nonlinear, ordinary differential equations. This allows for a robust numerical study and direct implementation of a standard perturbation analysis. These two methods of solution complement each other, yielding new physical insight into the bubble dynamics. The perturbation analysis facilitates a qualitative comparison with some of the previous results on resonance forcing behavior. The numerical investigations reveal intriguing, previously unreported phenomena.

Moreover, it is important to keep in mind the physical assumptions underpinning the linear viscoelastic constitutive equations. There are associated constraints on the amount of physically realizable radial deformation (i.e., growth and rebound) the bubble can assume, and hence on the overall magnitude of acoustic forcing. The perturbation method employed is similar to that previously used for the Newtonian case. This method is implemented for small-amplitude oscillations about the equilibrium radius; thus, the linear viscoelastic excursion limitations are satisfied. Nonlinear viscoelastic damping terms come in at a higher order and are

not considered in the analyses. Also, in the linear viscoelastic case the stress tensor is traceless, which simplifies the analysis since we need only consider the radial component of the stress tensor.

## I. GOVERNING EQUATIONS

The Rayleigh-Plesset equation for a gas bubble can be written in a general form, given in Eq. (1), which includes the stress tensor of the liquid.<sup>10</sup> This equation describes the motion of a spherically symmetric bubble of radius  $R(t)$  in an incompressible fluid of density  $\rho$  subject to a static surface tension  $\hat{\sigma}$ , an internal gas pressure  $p_i$ , and an ambient pressure  $p_\infty$ . For the linear viscoelastic (traceless) case, we need only consider the radial stress component,

$$R\ddot{R} + \frac{3}{2}\dot{R}^2 = \frac{1}{\rho_0} \left[ p_i - p_\infty - \frac{2\hat{\sigma}}{R} - 3 \int_R^\infty \frac{\tau_{rr} dr}{r} \right]. \quad (1)$$

This equation is coupled to a constitutive equation for the fluid which in turn relates the radial stress tensor,  $\tau_{rr}$ , to the spatial deformations in the fluid.

The linear Jeffreys model is chosen as the constitutive equation for this study.<sup>15</sup> It has been used as a starting point for more complicated models and reduces to the linear Maxwell model used by Fogler and Goddard in the case that the retardation time as defined below goes to zero.<sup>15</sup> The Jeffreys model can be expressed as

$$\tau_{rr} + \lambda_1 \frac{\partial \tau_{rr}}{\partial t} = -2\eta_0 \left( \dot{\epsilon} + \lambda_2 \frac{\partial \dot{\epsilon}}{\partial t} \right), \quad (2a)$$

$$\dot{\epsilon} = \frac{\partial u}{\partial r} = -\frac{2\dot{R}R^2}{r^3}; \quad u = \frac{\dot{R}R^2}{r^2}, \quad (2b)$$

where  $\lambda_1$  is the relaxation time,  $\lambda_2$  is the retardation time,  $\dot{\epsilon}$  is the rate of strain or deformation tensor, and  $\eta_0$  is the zero shear-rate viscosity. The relaxation time refers to the time it takes the stress to fall to  $1/e$  of its original equilibrium value and the retardation time is the time it takes the strain to fall to  $1/e$  of its equilibrium value after the removal of stress. The radial component of the deformation tensor corresponding to Eq. (2a) may be expressed as  $\partial u/\partial r$ , where  $u$  is radial velocity in the liquid for the bubble.<sup>9-14</sup> The radial stress, which results from the solution to (2), is coupled to Eq. (1) through a spatial integral over the surrounding fluid. Previous approaches such as that taken by Fogler and Goddard first solved the constitutive equation for the stress tensor Eq. (2) in terms of time and then substituted the resulting expression into the integral in Eq. (1).<sup>10-12</sup> Integration was subsequently performed in terms of the spatial variable. In this new approach, the stress tensor is first solved in terms of the spatial variables. We substitute from Eq. (2b) the deformation tensor in terms of the radial fluid velocity. Equation (2) is first divided by  $r$  and then integrated from the bubble wall to infinity,

$$\int_R^\infty \frac{\tau_{rr}(r,t) dr}{r} + \lambda_1 \int_R^\infty \frac{\partial}{\partial t} \left( \frac{\tau_{rr}(r,t)}{r} \right) dr = 4\eta_0 \left[ \int_R^\infty \frac{\dot{R}R^2 dr}{r^4} + \lambda_2 \int_R^\infty \frac{2R\dot{R}^2 + R^2\ddot{R} dr}{r^4} \right]. \quad (3)$$

The Leibnitz Rule can be used to simplify the second term on the left-hand side of Eq. (3),

$$\int_R^\infty \frac{\partial}{\partial t} \left( \frac{\tau_{rr}(r,t)}{r} \right) dr = \frac{d}{dt} \int_R^\infty \frac{\tau_{rr}(r,t) dr}{r} + \frac{dR}{dt} \frac{\tau_{rr}(R(t))}{R}. \quad (4)$$

Equation (4) is substituted into Eq. (3), and the two terms on the right-hand side of Eq. (3) are integrated out analytically, yielding the following equation:

$$\int_R^\infty \frac{\tau_{rr}(r,t) dr}{r} + \lambda_1 \left( \frac{d}{dt} \int_R^\infty \frac{\tau_{rr}(r,t) dr}{r} + \dot{R} \frac{\tau_{rr}(R(t))}{R} \right) = \frac{4}{3} \eta_0 \left( \frac{\dot{R}}{R} + \lambda_2 \left( \frac{2\dot{R}^2 + R\ddot{R}}{R^2} \right) \right). \quad (5)$$

This expression can be considered an ordinary differential equation for the quantity in the spatial integral, which we define as  $S$ . Use of this definition results in

$$S + \lambda_1 \left( \frac{dS}{dt} + \dot{R} \frac{\tau_{rr}(R)}{R} \right) = \frac{4}{3} \eta_0 \left( \frac{\dot{R}}{R} + \lambda_2 \left( \frac{2\dot{R}^2 + R\ddot{R}}{R^2} \right) \right), \quad (6)$$

such that

$$S(R,t) = \int_R^\infty \frac{\tau_{rr}(r,t) dr}{r}. \quad (7)$$

Equation (6) is thus coupled to the radial equation of the bubble wall motion by an equation for the stress tensor evaluated at the bubble wall. Evaluating Eq. (2) at  $r=R(t)$  yields

$$\tau_{rr}(R) + \lambda_1 \frac{\partial \tau_{rr}(R)}{\partial t} = 4\eta_0 \left( \frac{\dot{R}}{R} + \lambda_2 \left( \frac{2\dot{R}^2 + R\ddot{R}}{R^2} \right) \right). \quad (8)$$

This expresses the problem as a set of three coupled nonlinear differential Eqs. (1), (6), and (8). Assuming a polytropic relationship for the internal gas pressure closes the problem. Though this assumption has been shown to have some physical limitations especially with respect to driven oscillations and thermal damping,<sup>1</sup> these will not be of primary concern to us in this paper given the constraints in amplitude excursion imposed by the linear viscoelastic limit.<sup>5,6</sup> However, it is important to keep in mind the limitations of this assumption in mind when interpreting some of the new results.

Starting with this formulation of the problem, we implement perturbation techniques and obtain numerical solutions. To achieve this, we first chose a nondimensionalization appropriate to a multiple scale perturbation technique.

The following nondimensional quantities are introduced to make the problem more appropriate to a perturbation analysis. Some of the choices reflect those taken from previous perturbation studies of Newtonian bubbles. In the Newtonian limit such that the non-Newtonian terms vanish, the results can be checked against those previously obtained,

$$\begin{aligned}\bar{r} &= r/R_0; & \bar{R} &= R/R_0; & \bar{t} &= t / R_0 \sqrt{\frac{\rho}{p_{go}}}; \\ \bar{\tau} &= \frac{\tau}{\eta_0} R_0 \sqrt{\frac{\rho}{p_{go}}}.\end{aligned}\quad (9)$$

Time is nondimensionalized by essentially the linear bubble resonance frequency. In this perturbation method, we will focus on bubble behavior near its natural, linear resonance, which turns out to be one of the most important time scales. The initial gas pressure in the bubble is denoted by  $p_{go}$ . The choice of nondimensionalization for the last term warrants a brief explanation. This choice originates from a Newtonian fluid, where  $\tau \sim \eta_0 \partial u / \partial r$ . It is reasonable for our work, which concentrates on elastic deviations from the Newtonian case.

The radial bubble dynamics equation [Eq. (1)] is given by

$$\begin{aligned}\bar{R}\ddot{\bar{R}} + \frac{3}{2}\dot{\bar{R}}^2 &= (1/\bar{R}^{3\kappa}) - w(1/\bar{R}) - (1-w)(1 - \bar{a} \cos(\Omega \bar{t})) \\ &\quad - \frac{3}{\text{Re}} \int_R^\infty \frac{\bar{\tau}_{rr}(\bar{r}, \bar{t})}{\bar{r}},\end{aligned}\quad (10)$$

where we use a polytropic relationship for the exponent. The following nondimensional quantities are defined as

$$\begin{aligned}w &= \frac{2\hat{\sigma}}{p_{go}R_0}; & \Omega &= R_0 \sqrt{\frac{\rho}{p_{go}}} \omega; \\ \text{Re} &= \eta_0 / (p_{go}\rho)^{1/2} R_0; & \bar{a} &= \frac{p_a}{p_0},\end{aligned}\quad (11)$$

where the Reynolds number is denoted as Re. The quantity  $w$  represents a nondimensional Laplace pressure and  $\Omega$  the ratio of acoustic forcing frequency to the natural resonance frequency of the bubble. The equations for the stresses are

$$\bar{S} + \text{De} \left( \frac{d\bar{S}}{d\bar{t}} + \bar{R} \frac{\bar{\tau}_{rr}(\bar{R})}{\bar{R}} \right) = \frac{4}{3} \left( \frac{\dot{\bar{R}}}{\bar{R}} + \lambda \text{De} \frac{2\bar{R}^2 + \bar{R}\ddot{\bar{R}}}{\bar{R}^2} \right) \quad (12)$$

and

$$\bar{\tau}_{rr}(\bar{R}) + \text{De} \left( \frac{d\bar{\tau}_{rr}(\bar{R})}{d\bar{t}} \right) = 4 \left( \frac{\dot{\bar{R}}}{\bar{R}} + \lambda \text{De} \frac{2\bar{R}^2 + \bar{R}\ddot{\bar{R}}}{\bar{R}^2} \right), \quad (13)$$

and the nondimensional quantities such as the Deborah number (De) are expressed as follows:

$$\text{De} = \lambda_1 / R_0 \sqrt{\frac{\rho}{p_{go}}}; \quad \lambda = \frac{\lambda_2}{\lambda_1}. \quad (14)$$

The Deborah (De) number, a nondimensional number in non-Newtonian fluid dynamics, expresses the ratio of two time scales: the fluid to that of the flow.<sup>15</sup> This gives an estimate of the relative importance of elastic to viscous forces. It is usually taken as the ratio of a characteristic time of the fluid to a characteristic time for deformation. In this case, the Deborah number is defined as ratio of the relaxation time to the natural frequency of the bubble. The ratio of the retardation time to the relaxation time denoted by  $\lambda$  is expected to be a small quantity for relevant biological or polymeric fluids.<sup>5,15</sup> It should be noted again that in the limit the Deborah number goes zero, the Newtonian expressions for the viscous terms are recovered in Eqs. (10), (12) and (13).

## II. PERTURBATION ANALYSIS

Perturbation analyses have been used previously to describe the weakly nonlinear response of an acoustically forced gas bubble in a Newtonian fluid. Only methods that yield a uniform expansion for the nonlinear oscillations should be used.<sup>16</sup> Nayfeh and Saric were among the first to use the method of multiple scales to analyze a bubble in an inviscid, slightly compressible fluid governed by the Herring equation.<sup>17</sup> In a series of papers, Prosperetti used the method of averaging to examine the transient and steady-state response of a bubble governed by the Rayleigh-Plesset equation.<sup>18,19</sup> The Rayleigh-Plesset equation has been subsequently reexamined in a series of papers using the multiple scales analysis.<sup>20,21</sup> All of these studies were based on small-amplitude acoustic forcing and hence small oscillations about the equilibrium radius. We seek to undertake a similar analysis for the non-Newtonian, viscoelastic fluid. For notational convenience, we will drop the bars (denoting the nondimensional quantities) introduced in the previous section.

Following the previous approaches,<sup>18,19</sup> we seek a perturbation solution for small forcing, so we first expand the equilibrium radius in terms of relative amplitude  $x$ ,

$$R = 1 + x, \quad (15)$$

keeping terms to second order in  $x$ . We substitute Eq. (15) into Eqs. (10), (12), and (13) and obtain the following equations:

$$\begin{aligned}\ddot{x} + \omega_0^2 x &= F \cos(\Omega t) + \left( -\frac{3}{2}\dot{x}^2 + \alpha_1 x^2 - xF \cos(\Omega t) \right) \\ &\quad + \left( -\alpha_2 x^3 + x^2 F \cos(\Omega t) + \frac{3}{2}\dot{x}^2 x \right) + \alpha_3 S,\end{aligned}\quad (16)$$

$$\begin{aligned}S + \text{De} \left( \frac{dS}{dt} + \tau_{rr}(\dot{x} - x\dot{x}) \right) &= \frac{4}{3} ((\dot{x} - x\dot{x}) \\ &\quad + \lambda \text{De}(\ddot{x} + (2\dot{x}^2 - x\ddot{x}))),\end{aligned}\quad (17)$$

$$\tau_{rr} + \text{De} \frac{\partial \tau_{rr}}{\partial t} = 4((\dot{x} - x\dot{x}) + \lambda \text{De}((2\dot{x}^2 - \ddot{x}x) + \ddot{x})), \quad (18)$$

where we have introduced the following quantities:

$$\alpha_1 = \frac{9}{2}(\kappa + 1) - 2w; \quad \alpha_2 = \frac{\kappa}{2}(9\kappa^2 + 18\kappa + 11) - 3w; \quad (19)$$

$$\alpha_3 = \frac{-3}{\text{Re}}; \quad F = (1 - w)\bar{a}.$$

First, we employ the method of multiple scales to obtain an analytical solution near the primary resonance such that the forcing is near the natural linear frequency for free oscillations. In nonlinear oscillations of the bubble, the amplitude and phase of the oscillations depend upon each other. A standard perturbation scheme will fail to capture the variations of the amplitude with the phase and results in a nonuniform expansion. However, these problems may be overcome by allowing the solution to be a function of independent time scales. The amplitude and phase interaction tends to occur at slower scales than natural oscillations. We approximate this frequency as  $\omega_0^2 = 3\kappa - w$ . We need to order the excitation, the nonlinearity, and the viscoelastic damping effects so that their contributions appear at the same time in the perturbation scheme.<sup>16,20,21</sup> In this case, we define with the small parameter  $\varepsilon$ ,

$$F = \varepsilon^3 f; \quad x = \varepsilon u. \quad (20)$$

The viscoelastic damping effects from the stress Eqs. (12) and (13) need to appear at the same order. Our choice will be  $1/\text{Re} \sim O(\varepsilon)$  to match with the expansions of Eqs. (17) and (18). Using Eq. (20), the radial equation (16) can be expressed as

$$\ddot{u} + \omega_0^2 u = \varepsilon^2 f \cos(\Omega t) + \varepsilon \left( -\frac{3}{2} \dot{u}^2 + \alpha_1 u^2 + \alpha_3 S \right) + \varepsilon^2 \left( -\alpha_2 u^3 + \frac{3}{2} \dot{u}^2 u \right). \quad (21)$$

We can implement the method of multiple scales to obtain expansions for the dependent variables

$$u(t; \varepsilon) = u_0(T_0, T_1, T_2) + \varepsilon u_1(T_0, T_1, T_2) + \varepsilon^2 u_2(T_0, T_1, T_2) + \dots, \quad (22)$$

$$\tau_{rr}(t; \varepsilon) = \tau_0(T_0, T_1, T_2) + \varepsilon \tau_1(T_0, T_1, T_2) + \varepsilon^2 \tau_2(T_0, T_1, T_2) + \dots, \quad (23)$$

$$S(t; \varepsilon) = S_0(T_0, T_1, T_2) + \varepsilon S_1(T_0, T_1, T_2) + \varepsilon^2 S_2(T_0, T_1, T_2) + \dots, \quad (24)$$

where

$$T_n = \varepsilon^n t; \quad n = 0, 1, \dots, \quad (25)$$

and the time derivatives can be expressed in terms of the multiple time scales as

$$d/dt = \sum_{n=0} \varepsilon^n D_n, \quad D_n = \partial/\partial T_n, \quad (26a)$$

$$d/dt = D_0 + \varepsilon D_1 + \varepsilon^2 D_2 + \dots, \quad (26b)$$

$$d^2/dt^2 = D_0^2 + 2\varepsilon D_0 D_1 + \varepsilon^2 (D_1^2 + 2D_0 D_2) + \dots$$

Substituting Eqs. (22)–(24), and (26) into Eqs. (16)–(18), and equating powers of epsilon, we obtain

$$D_0^2 u_0 + \omega_0^2 u_0 = 0 \quad O(\varepsilon^0), \quad (27)$$

$$S_0 + \text{De}(D_0 S_0) = 0, \quad (28)$$

$$\tau_0 + \text{De}(D_0 \tau_0) = 0. \quad (29)$$

We see from Eqs. (28) and (29) that, to first order, the equations for the stresses decouple such that in this case we need only consider the  $S$  equations, the integral of the stress, in order to include the damping effects. We need to have the damping effects come in at the same time as the forcing, so we set these variables to zero in Eqs. (28) and (29),

$$S_0 = 0; \quad \tau_0 = 0. \quad (30)$$

The next order equations can be expressed as

$$D_0^2 u_1 + \omega_0^2 u_1 = -2D_0 D_1 u_0 - \frac{3}{2}(D_0 u_0)^2 + \alpha_1 u_0^2 O(\varepsilon^1), \quad (31)$$

$$S_1 + \text{De}(D_0 S_1) = \frac{4}{3}(D_0 u_0 + \lambda \text{De}(D_0^2 u_0)), \quad (32)$$

and performing the expansion out to second order in epsilon yields

$$D_0^2 u_2 + \omega_0^2 u_2 = -D_1^2 u_0 - 2D_0 D_1 u_1 - 2D_0 D_2 u_0 - 3(D_0 u_0)(D_0 u_1) + 2\alpha_1 u_0 u_1 - \alpha_2 u_0^3 + \frac{3}{2} u_0 (D_0 u_0)^2 - 3(D_0 u_0)(D_1 u_0) + \alpha_3 S_1 + f \cos[(\omega_0 T_0 + \sigma T_2)] O(\varepsilon^2) \quad (33)$$

$$\Omega = \omega_0 + \varepsilon^2 \sigma,$$

where  $\sigma$  represents the detuning parameter (frequency of the excitation).<sup>16</sup> The solution to Eq. (33) can be expressed as

$$u_0 = A(T_1, T_2) \exp(i\omega_0 T_0) + cc. \quad (34)$$

$cc.$ —complex conjugate

for the unknown complex function  $A$ . Substituting this into Eq. (31) yields

$$D_0^2 u_1 + \omega_0^2 u_1 = -2i\omega D_1 A e^{i\omega_0 T_0} - \frac{3}{2}\omega_0^2 (A\bar{A} - A^2 e^{i2\omega_0 T_0}) + \alpha_1 (A^2 e^{i2\omega_0 T_0} + A\bar{A}) + cc., \quad (35)$$

where the “ $\sim$ ” symbol denotes the complex conjugate. Using the freedom of multiple scales to eliminate the secular terms (those not uniformly valid as time increases), we let

$$A = A(T_2), \quad (36)$$

so that the first term on the left-hand side goes to zero and we can solve Eq. (35),

$$u_1 = c_1 A^2 \exp(2i\omega_0 T_0) + c_2 A\bar{A} + cc., \quad (37)$$

$$c_1 = -\frac{\alpha_1 + \frac{3}{2}\omega_0^2}{3\omega_0^2}; \quad c_2 = \frac{\alpha_1 - \frac{3}{2}\omega_0^2}{\omega_0^2}.$$

Likewise, we substitute Eq. (37) into the equation for  $S$ , Eq. (32). We assume a solution of the form

$$S_1 = C \exp(i\omega_0 T_0) + ED + cc. \quad (38)$$

The term  $ED$  depends on the initial conditions and is exponentially decaying. One objective of this analysis is to examine steady-state responses where this term does not contrib-

ute. In this case, we set it equal to zero and solve for the constant

$$C = \frac{4}{3} \frac{i\omega_0 - \text{De} \lambda \omega_0^2}{1 + i\text{De} \omega_0}, \quad (39)$$

which can be separated into its real and imaginary parts,

$$C = c_r + ic_i, \quad (40)$$

$$c_r = -\frac{4}{3} \frac{\text{De} \omega_0^2 (\lambda - 1)}{(1 + \text{De}^2 \omega_0^2)}, \quad (41)$$

$$c_i = \frac{4}{3} \frac{\omega_0 (1 + \text{De}^2 \lambda \omega_0^2)}{(1 + \text{De}^2 \omega_0^2)}. \quad (42)$$

It is advantageous to do this separation because the amplitude and phase equations will be separated into real and imaginary parts in the method of multiple scales, and these coefficients [Eqs. (41), (42)] have some physical significance. The imaginary part Eq. (42) corresponds to the storage or elastic component in the fluid and the real part Eq. (41) is proportional to the viscous damping or loss due to dissipation. A Newtonian fluid will only have a viscous loss term proportional to its natural frequency, as seen by taking the limit as the Deborah number goes to zero. As the ratio of the retardation to relaxation time ( $\lambda$ ) approaches one, the magnitude of the storage coefficient decreases for increasing loss coefficient.

We can substitute the results of Eqs. (34), (37), and (38) into Eq. (33), resulting in

$$\begin{aligned} D_0^2 u_2 + \omega_0^2 u_2 = & (-2i\omega_0 A' + c_4 A^2 \bar{A} + \frac{1}{2} f \exp(i\sigma T_0) \\ & + \alpha_3 (c_r + ic_i) A) e^{i\omega_0 T_0} + \text{nst} + cc., \end{aligned} \quad (43)$$

$$c_4 = \frac{20\alpha_1^2 - 30\alpha_1 \omega_0^2 - 18\alpha_2 \omega_0^2 + 27\omega_0^2}{6\omega_0^2},$$

where the prime denotes a derivative with respect to  $T_2$  and ‘nst’ refers to nonsecular terms. We can set the secular term in Eq. (43) to zero and separate into real and imaginary parts by introducing

$$A = \frac{1}{2} a(T_2) \exp(i\beta(T_2)). \quad (44)$$

The real and imaginary parts are

$$a\beta' + \frac{c_4}{8} a^3 + \frac{f}{2} \cos \gamma + \frac{a}{2} \alpha_3 c_r = 0, \quad (45)$$

$$-a' \omega_0 + \frac{f}{2} \sin \gamma + \frac{a}{2} \alpha_3 c_i = 0, \quad (46)$$

respectively. Eliminating  $\beta$  yields the following equations:

$$a\sigma\omega_0 - a\gamma' \omega_0 + \frac{c_4}{8} a^3 + \frac{f}{2} \cos \gamma + \frac{a}{2} \alpha_3 c_r = 0, \quad (47)$$

$$-a' \omega_0 + \frac{f}{2} \sin \gamma + \frac{a}{2} \alpha_3 c_i = 0, \quad (48)$$

$$\gamma = \sigma T_2 - \beta. \quad (49)$$

The steady-state response, corresponding to the stationary solutions to Eqs. (47) and (48), is

$$a\sigma = -\frac{c_4}{8\omega_0} a^3 - \frac{f \cos \gamma}{2\omega_0} - \frac{a}{2\omega_0} \alpha_3 c_r, \quad (50)$$

$$\frac{f}{2} \sin \gamma = -\frac{a}{2} \alpha_3 c_i. \quad (51)$$

These equations can be combined into a single equation that describes analytically the frequency-dependent amplitude response of the bubble,

$$f^2 = a^2 (\alpha_3 c_i)^2 + a^2 \left( 2\sigma\omega_0 + \frac{c_4}{4} a^2 + \alpha_3 c_r \right)^2. \quad (52)$$

This equation can be solved numerically to obtain a frequency response curve (amplitude,  $a$ , versus the frequency detuning,  $\sigma$ ). The response for different Deborah numbers indicates the effect of elasticity on the primary resonance. Some of these results are shown in Figs. 1 and 2. The results shown in Fig. 1. are the same case except at Deborah numbers of 0.0, 1.0, and 2.0 for fixed value of retardation to relaxation time. The  $\text{De}=0.0$  case corresponds to the Newtonian response. We notice for increasing Deborah number, and hence increasing elastic properties in the fluid, the amplitude of response becomes greater. The bubble does not exhibit nonlinear softening effects in the Newtonian ( $\text{De}=0.0$ ) case. However, these effects become apparent for Deborah numbers of 1.0 and 2.0. In these cases, the amplitude jumps to a lower value for decreasing frequency and a higher one for increasing frequency. Bubbles in Newtonian fluids also exhibit this softening phenomenon,<sup>17,18,20</sup> but this effect appears at lower forcing amplitude with the inclusion of viscoelasticity. A possible physical explanation for this is that the addition of elasticity seems to cause the bubble to stay at its point of maximum expansion longer. It appears it takes longer with elasticity for the stresses in the fluid to begin initiating the bubble collapse process.

Figure 2 shows the frequency response curve for  $\text{De}=1.0$  and for values of the ratio of the retardation to relaxation time of 0.0, 0.1, and 0.2. The linear Maxwell model corresponds to the  $\text{De}=0.0$  case. The softening effect diminishes with increasing retardation time. A possible explanation for this result might be that extra dissipation associated with the Jeffreys model damps out the nonlinearity even at these low amplitudes of forcing. The retardation term for this approximation can be seen as a loss that is proportional to the acceleration of the bubble wall.

The stability of these response curves for steady-state motion can be investigated by linearizing the amplitude and phase equations about the singular points. Information is gained about the local stability by examining small disturbances or perturbations about these points,

$$a = a_0 + a_1, \quad \gamma = \gamma_0 + \gamma_1. \quad (53)$$

The zero subscript denotes the singular points and the one subscript the disturbances about them. The eigenvalues of the coefficient matrix of the linearized equation determine the stability characteristic.<sup>16</sup> The system is stable if the characteristic roots have negative real parts.<sup>16</sup> This method was employed to investigate the stability of Eqs. (50) and (51). The structural characteristics that were examined appear

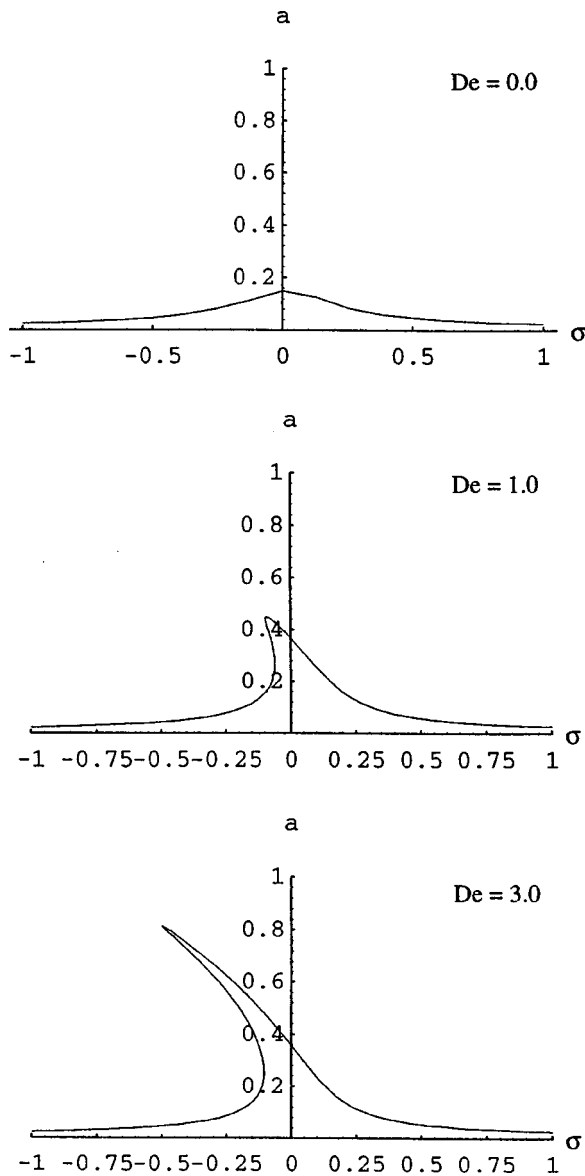


FIG. 1. Shown (from top to bottom) are the nondimensional frequency response curves for  $De=0.0$ ,  $1.0$ , and  $3.0$ . The other parameter values are  $\kappa=1.4$ ,  $w=0.145$ ,  $1/Re=0.1$  and  $\lambda=0.1$ . Note that the nonlinear softening response is visible for the viscoelastic cases but not the Newtonian case ( $De=0.0$ ). Moreover, this nonlinearity increases with increasing Deborah number.

quite similar to the Newtonian results. The stability requirements for steady-state motion are shifted by an overall constant proportional to the real and imaginary damping coefficients. We might expect more interesting results if we retained higher order terms from the constitutive equation expansions.

Numerical and experimental studies have shown the occurrence of secondary harmonics, namely superharmonics and subharmonics from acoustically driven Newtonian bubbles.<sup>1,2</sup> The method of multiple scales has been used to analytically examine these responses. Nayfeh and Saric's initial study determined the superharmonic responses and pressure amplitude threshold for the first subharmonic resonance from a form of the Herring equation.<sup>16</sup> Nayfeh and Mook used this same ordering of the damping, nonlinearity, and excitation terms in the perturbation analysis for a study of a

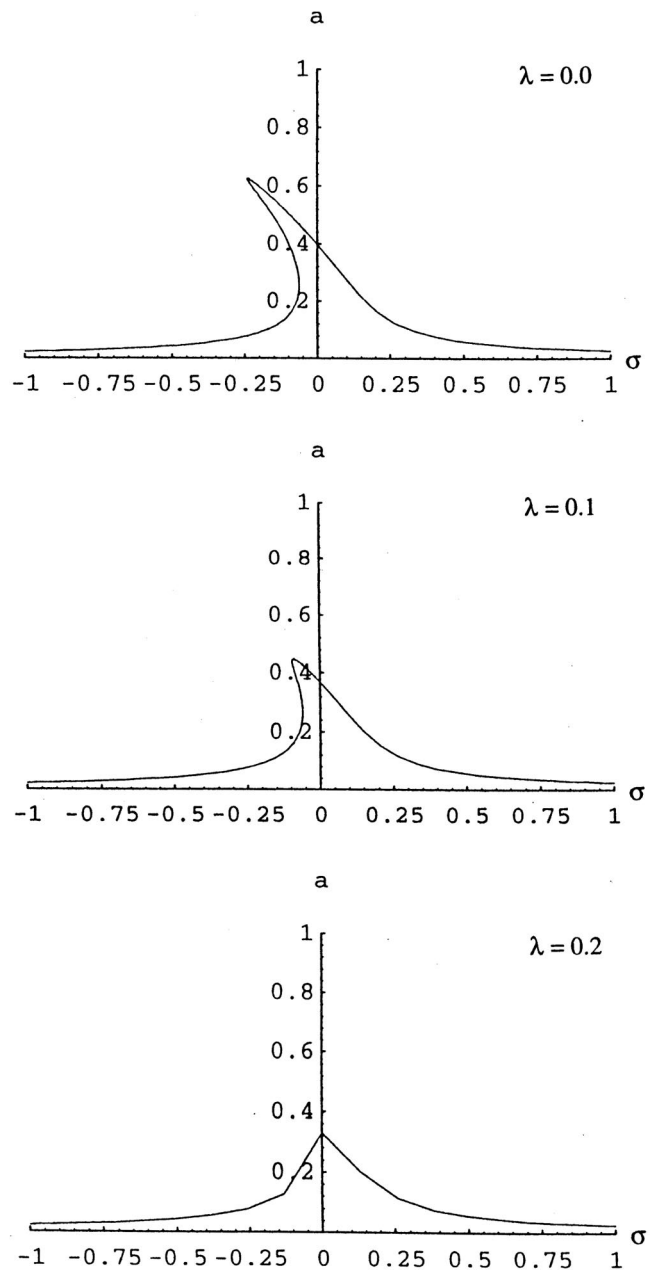


FIG. 2. Increases in the retardation time of the fluid damp of the frequency response curves, shown here for  $De=1.0$  (other parameters are the same as in Fig. 1) and  $\lambda=0.0$ ,  $0.1$ , and  $0.2$  (shown top to bottom). This indicates that a bubble response in a linear Jeffreys fluid will be diminished compared to that in a linear Maxwell fluid.

related nonlinear oscillation problem.<sup>15</sup> Francescutto and Nabergoj followed this ordering approach in their studies of the first subharmonic in the Rayleigh-Plesset equation.<sup>16,17</sup> However, in a subsequent paper Nayfeh addressed the limitations and validity of this ordering and offered a uniformly valid alternative.<sup>22</sup> In particular, he demonstrated that the previously used ordering is not valid for large times due to the fact that resonant terms are produced by the quadratic terms in the equation before a self-interaction by the nonlinearity. Using the alternative ordering, Nayfeh found a region where the subharmonics may be excited (depending on the initial conditions) which was not revealed in the previous studies.<sup>22</sup> In the improved first subharmonic ordering, Nay-



feh examines both the cases of small damping for closely tuned oscillations and larger damping for a case that is not as closely tuned.<sup>22</sup> The closely tuned subharmonic case refers to a nearness of  $\Omega$  to  $2\omega_0$  and the detuning  $\sigma$  can be defined as

$$\Omega = 2\omega_0 + \varepsilon^2\sigma. \quad (54)$$

The additional viscoelastic equations can be easily extended to Nayfeh's closely tuned case using the same expansions (23) and (34) and results in Eq. (30) for the primary harmonic where damping is scaled so it appears in the second order equations of the expansions. To first order, we expect the effects similar to those on the primary resonance from the real and imaginary constants. (We will leave a comprehensive study of the secondary harmonics, in particular the subharmonic case, for numerical studies that follow.) The reason for this being that a numerical solution of the system preserves more of the physics of the constitutive equations.

### III. NUMERICAL ANALYSIS

In this section, we obtain numerical solutions to Eqs. (1), (6) and (8). For these studies, we adopt a change in the choice of nondimensionalization of several quantities. The choice of Deborah number used in Eq. (14) was appropriate for investigating forcing near the natural linear frequency of the bubble. In that case, the results may easily be checked against previous perturbation solutions in Newtonian limit. In these numerical investigations of the complete set of equations, it is more desirable that the nondimensional numbers reflect the way the inertial and elastic forces depend on the forcing frequency. In this section, the Deborah and Reynolds numbers are rescaled to reflect the acoustic forcing frequency for off-resonance cases. Time is re-dimensionalized such that

$$\bar{t} = \omega t, \quad (55)$$

then the Deborah number is rewritten as

$$De = \lambda_1 \omega, \quad (56)$$

and also the Reynolds number is rewritten,

$$Re = \rho \omega R_0^2 / \eta_0. \quad (57)$$

This definition of the Deborah number approximates Eq. (14) as the acoustic forcing approaches the linear resonance frequency. Also with this choice, a Deborah number of order one or greater represents physically reasonable quantities for studies of polymeric and biological fluids. A nondimensional rescaling of the stress tensor was examined; however, it was found that it was still physically and numerically reasonable to essentially retain the previous nondimensionalization of the stress tensor with the substitution of the ambient pressure for the equilibrium bubble pressure.

First, we consider a case of acoustic forcing away from the bubble's natural frequency. We examine a 0.5-micron equilibrium radius bubble at 1.0 MHz forcing for various values of Reynolds and Deborah numbers. In light of potential biomedical applications related to this work, the values are chosen to roughly approximate biological fluids or tissue.<sup>5</sup> Polymer stabilized (xanthan gum) bubble clouds have been manufactured for underwater acoustics applications<sup>25</sup> with Deborah and Reynolds numbers compa-

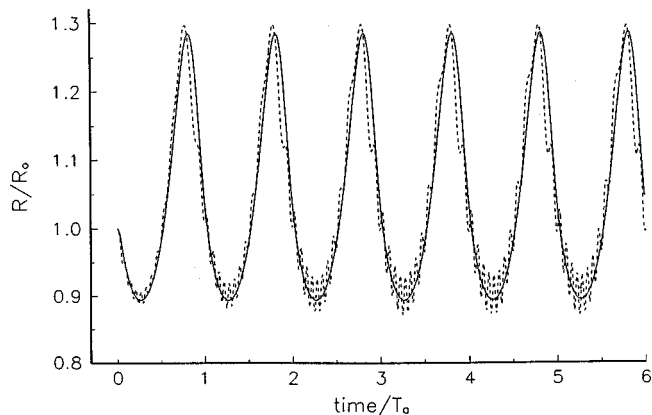


FIG. 3. The oscillatory behavior of a 0.5-micron radius bubble (0.2 MPa, 1.0 MHz,  $Re=0.2$ ) demonstrates the differences in bubble response in a viscoelastic linear Maxwell fluid at  $De=1.0$  (dashed line) and a Newtonian fluid (solid line). A slight phase shift exists between the two cases. Small oscillations occur near the cycle minima for the viscoelastic case.  $T_a$  is the acoustic period.

ble to those used in generating Figs. 9–12. The pressure amplitudes are kept low (usually below 0.2 MPa) in order to stay within the restrictions of linear viscoelasticity. The surface tension is kept fixed so the effects of varying  $Re$  and  $De$  number can be discerned. For all the numerical cases, the interior gas is assumed to be air, such that  $\kappa=1.4$ , and that the fluid has surface tension ( $72.5 \text{ dyn/cm}^2$ ) and density ( $0.999 \text{ cm/gm}^3$ ) of that of water. Actual values for biological fluids and tissue are probably different; however, such values are not well established and the aim of this initial paper is to examine the effects of viscoelasticity. In the limit that the Deborah number, elasticity of the fluid, goes to zero a viscous Newtonian fluid is recovered ( $\lambda_1=0$ , Newtonian fluid). Figure 3 shows bubble oscillations where the  $De=1.0$  and  $Re=0.2$  compared to a Newtonian fluid of the same Reynolds number.

Several novel features distinguish the viscoelastic case from the corresponding Newtonian case. First, the overall phase of the oscillation is slightly shifted from the Newtonian case. This may be due to the presence of elasticity, which results in the storage of energy in the fluid. Furthermore, the viscoelastic bubble reaches a maximum slightly before the Newtonian bubble. Near the cycle minimum there are many small-scale oscillations, essentially manifestations of relatively small fluctuations in radius. One initial concern might be that these oscillations are numerical artifacts and not physical in origin. However, numerical tests varying the accuracy parameters in the integration method were done to check for these problems; robustness was verified. The oscillations were also found to be a function of the physical parameters in the study, thus casting further doubt on the notion that these oscillations are due to numerical effects. Finally, it should be noted that researchers who have examined viscoelastic bubble oscillations previously and related viscoelastic flow problems have encountered similar oscillatory phenomena, which they concluded to be physical in nature.<sup>23</sup> Our conclusion is that these oscillations are probably due to weak coupling of the elastic fluid effects with the bubble's motion. They occur near the oscillation cycle mini-

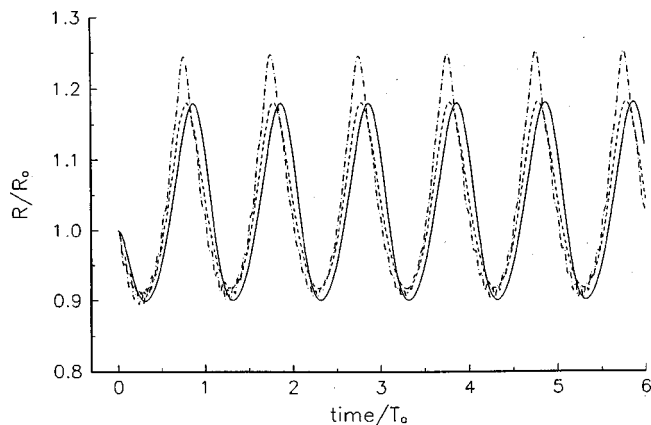


FIG. 4. The 0.2-micron bubble from Fig. 3 is shown for  $Re=0.05$  in Newtonian (solid line),  $De=1.0$  (dashed line) and  $De=3.0$  (dot-dash line) fluids. The lower Reynolds number leads to a damping of the small oscillations, though the phase shift effect remains visible. The amplitude of radial excursion increases with increasing Deborah number, a result that is consistent with the perturbation analysis.

num as the bubble wall velocity begins to slow down, consistent with our previous notions that time is required for the stress effects to catch up to bubble's wall motion.

We further demonstrate that these small-scale oscillations are a function of the physical parameter, as indicated in Fig. 4; the Reynolds and Deborah are set to  $Re=0.05$  and  $De=1.0, 3.0$ ; the Newtonian case ( $De=0.0$ ). The smaller Reynolds number corresponds to an increased viscosity damping. The small oscillations near the cycle minima have been diminished and the overall phase shift from the Newtonian case seen previously remains visible. In this case, an increase from  $De=0.0$  to  $De=1.0$  results in a visible phase shift but little change in amplitude. Increasing the Deborah number from 1.0 to 3.0 results in clear growth in the amplitude of the response. For these parameters it appears a sufficient Deborah number must be reached for the amplitude effect to develop.

A plot of the radial stress tensor at the bubble wall for the  $De=1.0$  and  $De=3.0$  cases from Fig. 4 is given in Fig. 5. For  $De=3.0$ , the plot shows some stress oscillations corresponding with small oscillations in radius. The excursions in stress are less in the  $De=1.0$  case compared to the  $De=3.0$  condition. A plot of the bubble wall velocity further illus-

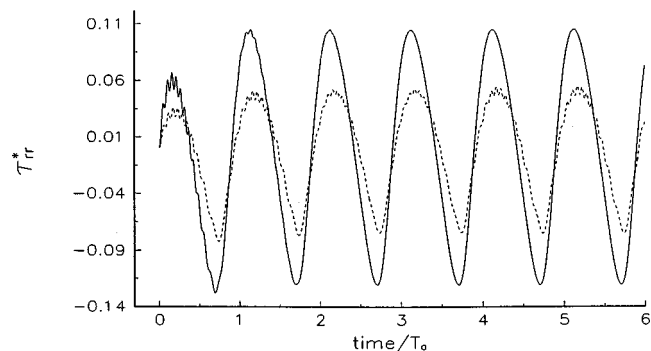


FIG. 5. The radial stress tensor at the bubble wall for the same parameters given in Fig. 4 is shown for  $De=1.0$  (solid line) and  $De=3.0$  (dashed line) cases. The elasticity seems to influence the magnitude of the stress produced at the bubble wall.

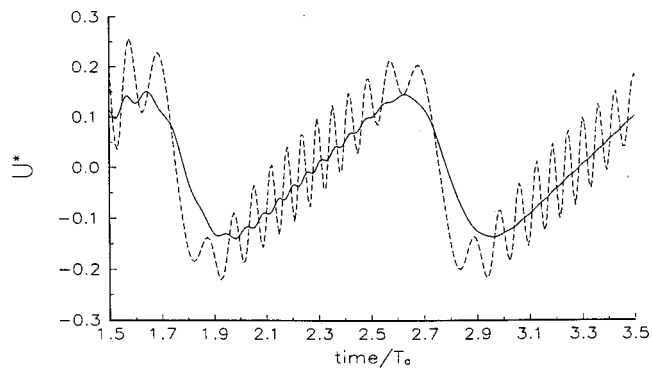


FIG. 6. The bubble wall velocity for the  $De=3.0$  (dashed line) and  $De=1.0$  (solid line) cases from Fig. 5. A steep drop in velocity followed by fluctuations corresponds with a less intense collapse.

trates this point (see Fig. 6). During the initial moments of collapse, the greater Deborah number results in a faster collapse. As the velocity begins to bottom out, the elastic stress starts inducing oscillations in the radius and the radial velocity. These effects serve to lessen the magnitude of the overall collapse strength. This phenomenon is a direct consequence of stress relaxation effects incorporated in the linear Maxwell model. Furthermore, these results hold important implications in the assessment of cavitation damage in media with increasing elasticity. Kim, using a more complicated objective model, made some similar observations for free bubble oscillations.<sup>13</sup> Some of the available experimental evidence indicates that polymeric additives may suppress damage effects resulting from bubble collapse, at least in the hydrodynamic case.<sup>24</sup> Our results offer physical insight into the possible governing mechanisms behind this important effect.

The previously depicted results are probably due to a weak coupling between the elastic properties of the fluid and the bubble oscillations. More profound effects are exhibited for near-resonance forcing when compared with the Newtonian case. This is shown in Fig. 7. The secondary resonance shown in the dashed curve is not present in the corresponding Newtonian case. Hence, an additional resonance effect associated with the elasticity of the fluid is manifest. While

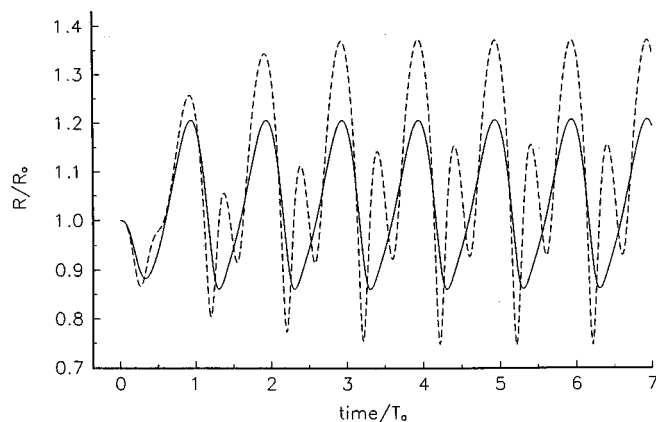


FIG. 7. A bubble's harmonic response appears to be linked to the fluid's viscoelastic properties under certain conditions. A 0.5-micron radius bubble forced near resonance (6.4 MHz, 0.2 MPa,  $Re=1.93$ ) in a  $De=1.0$  (dashed line) fluid reveals a secondary harmonic. The corresponding Newtonian case has the expected single harmonic component.

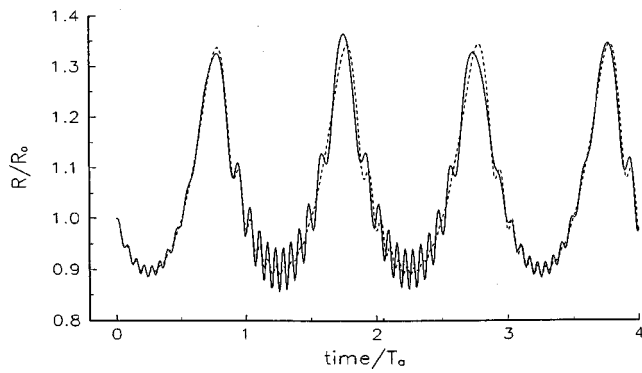


FIG. 8. The linear Jeffreys model (dashed line) for  $\lambda=0.1$  for a 0.5-micron radius bubble ( $Re=0.3$ ,  $De=2.0$ ) exhibits damped small oscillations when compared with the corresponding Maxwell fluid (solid line) case for the same parameters. These results agree qualitatively with the perturbation results.

only heightened amplitude effects were seen in the perturbation analysis, the inclusion of the complete stress equations serves to uncover an additional secondary resonance. The previously considered off-resonance case shows subtle minor variation from the Newtonian case, but here we see a more dramatic difference. In particular, this result has important implications for contrast agent harmonic imaging applications. It indicates that the occurrence of secondary resonances may be a function of specific tissue properties. Modeling tissue as viscous Newtonian fluid may not be an appropriate assumption for the prediction of secondary harmonics in sufficiently elastic tissue. Likewise, at relatively lower amplitudes of acoustic forcing, tissue elasticity may enhance secondary harmonic growth despite the highly viscous nature of the medium. Also, it should be noted a greater maximum radius is reached with the addition of elasticity. This is an important parameter for accessing potential inertia cavitation effects. The role maximum radius value plays is key in the development of the ultrasound cavitation safety criteria known as the Mechanical Index (MI).<sup>4,8</sup> This implies that tissue elasticity may need to be incorporated in future formulations of the MI.

In the perturbation scheme, the inclusion of the retardation time with the Jeffreys model results in an increasing damping of the amplitude response. This result agrees qualitatively with Shima and Tomita's study of viscoelastic bubble oscillations with the three-constant Oldroyd model.<sup>14</sup> In Fig. 8, we show that the Jeffreys model damps out oscillations that are present in the linear Maxwell model for the same set of parameters. Additional numerical studies indicate that increasing the ratio of retardation to relaxation times further damps the oscillations. We leave a more comprehensive examination of the retardation term to future studies.

As mentioned above, subharmonic responses have been observed theoretically and experimentally for bubbles in Newtonian fluids. This effect, however, has not been thoroughly investigated for viscoelastic bubbles. We will examine the 2/1 subharmonic for a 0.1-mm radius bubble driven at roughly twice its natural frequency, 64.74 kHz for a 0.04 MPa pressure amplitude using the linear Maxwell model. These parameters are chosen because subharmonic behavior

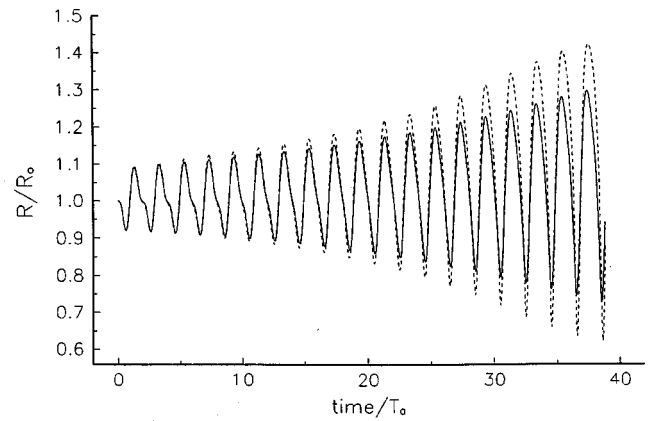


FIG. 9. The subharmonic response (0.1 mm radius, 64.74 kHz, 0.04 MPa,  $Re=203$ ) at  $De=1.0$  (dashed line) shows a more rapid growth when compared with the corresponding Newtonian ( $De=0.0$ ) (solid line) case. Notice the heightened amplitude grows for a longer forcing period.

in kHz bubbles is attainable at much lower pressure amplitudes than for micron-size bubbles at megahertz frequencies. Figure 9 shows the case where  $Re=203$  (20 cP) for a viscous Newtonian fluid, and for a viscoelastic fluid where  $De=1.0$ . The higher Reynolds number in this case in part reflects the larger equilibrium bubble radius. For Reynolds numbers smaller than this value it was found that the subharmonic structure was obscured due to viscous damping. We notice a faster growth in the radial response for the  $De=1.0$  case as compared to the corresponding Newtonian case. Indeed, the chosen parameter values allow for sufficient dissipation to permit observation of elastic effects and also restrict the amplitude excursions to be within the linear viscoelastic deformation limits. For the same Reynolds number, the Newtonian case is plotted against the  $De=3.0$  fluid in Fig. 10. This elevation in elasticity yields an even greater increase in the amplitude of the radial motion; however, the oscillation amplitude reaches a maximum and appears to fall off after about 40 cycles of acoustic forcing. We believe that the growth and decay is due at least in part to the elastic component of the fluid. However, our results must be viewed in light of our omission of thermal damping, which becomes more important in the parameter range for this case. Thermal damping

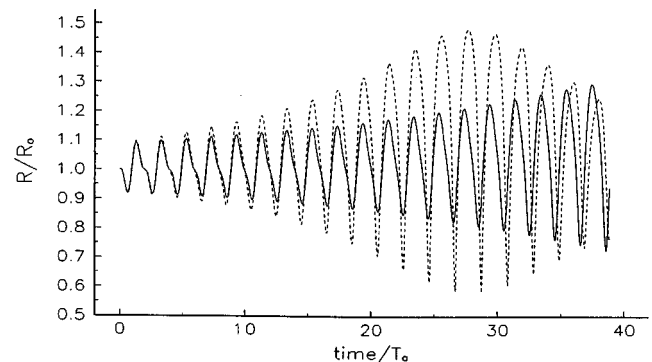


FIG. 10. A rapid growth of bubble oscillations occurs for  $De=3.0$  fluid (dashed line) when compared with the Newtonian case ( $De=0.0$ ). The subharmonic for  $De=3.0$  reaches a maximum during the first 40 cycles of forcing (compare with the  $De=1.0$  case in Fig. 9).

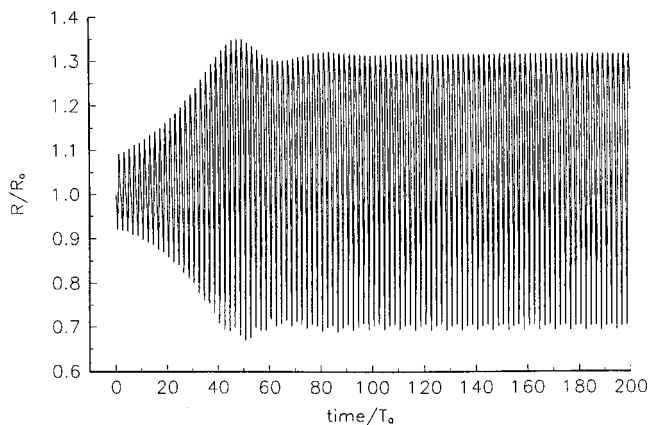


FIG. 11. A bubble in a Newtonian fluid ( $De=0.0$ ) forced by the given subharmonic excitation (see the caption for Fig. 9) reaches a steady-state excursion amplitude after about 55 cycles of forcing.

may in fact reduce the pronounced differences seen here between the viscous and viscoelastic cases.

The Newtonian case was examined over upward of 100 cycles of forcing, as shown in Fig. 11. After about 45 cycles, the subharmonic reaches a maximum and it subsequently decreases over the next 10 cycles to a steady-state value. The non-Newtonian cases for various Deborah numbers were also examined over long periods of forcing. Particularly dramatic effects appeared for  $De=10.0$ , as shown in Fig. 12. The subharmonic not only reaches maximum much more rapidly than in the Newtonian case, but its amplitude undergoes a beat-type modulation. For lower Deborah numbers, we see the gradual onset of separate amplitude packets, but eventually these effects would give way to the steady-state subharmonics observed in the Newtonian case. However, at  $De=10$  we see a full modulation effect over many cycles of forcing. The motion is itself a periodic subharmonic but appears to be modulated in a periodic fashion.

It is important to note the two simultaneous physical interactions occurring. One is the growth of the steady-state subharmonic governed by a coupling of the free oscillations of the bubble with the acoustic forcing frequency. This is immediately present in the Newtonian case and such motion

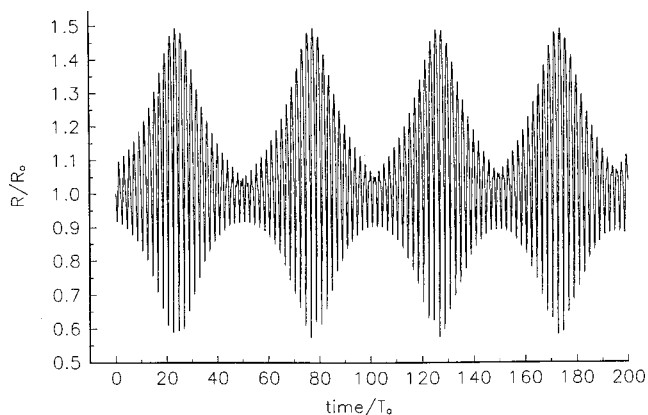


FIG. 12. For a bubble in a highly elastic fluid ( $De=10.0$ ) forced by the given subharmonic excitation (Figs. 9–11), the amplitude undergoes a novel beat modulation. Discrete groups of oscillations occur in contrast to the Newtonian case shown in Fig. 11. The width of each of these modules is related to the characteristic relaxation time of the fluid.

eventually dominates for fluids with Deborah numbers approximately less than ten. The other physical process is the addition of elasticity to the fluid, which seems to increase the oscillatory nature of the bubble response. The elasticity effects are apparently sufficient at about  $De=10.0$  that an additional, previously unknown, coupling occurs. The relaxation time of the fluid couples with the subharmonic such that it influences the magnitude of its amplitude. The characteristic time scale for amplitude modulation is on the order of the relaxation time of the fluid. The fact that the modulation time scales with the relaxation time indicates that this phenomenon is directly linked to the viscoelastic nature of the fluid. As mentioned earlier, thermal damping may modify the extent of this modulation and is the subject of further study. Nevertheless, this result is a bubble dynamics phenomenon that is unique to viscoelastic fluids and is a model prediction that lends itself to direct experimental verification.

#### IV. CONCLUSIONS

This study reveals several new and interesting features about gas bubble oscillations in viscoelastic fluids. A system of coupled nonlinear ordinary differential equations has been derived which describes the radial bubble dynamics. This formulation greatly facilitates numerical and analytical analysis of the problem. A perturbation approach using multiple scales predicts increasing damping with increasing retardation time and heightened amplitude response for increasing Deborah number. The additional elasticity adds a further softening of the nonlinear bubble response.

While the perturbation approach includes the viscoelastic effects to first order, the numerical studies incorporate the entire constitutive equation. A slight phase shift and additional oscillations distinguish the linear Maxwell case from the Newtonian case. The addition of a retardation term within the Jeffreys model damps the oscillations, a result also predicted by the perturbation theory. Finally, the addition of elasticity to the fluid produces some novel effects in the study of subharmonic excitations for kHz forcing frequencies. While the corresponding Newtonian case reaches steady amplitude subharmonic response, the amplitude of the highly viscoelastic case may be modulated in a manner seemingly governed by the relaxation time of the fluid. The impact of thermal damping on these results will be examined elsewhere.

This work has implications for many scientific and engineering applications. It was demonstrated that a bubble's response to acoustic forcing in a viscoelastic fluid might be dramatically different than that in Newtonian fluid. The amplitude restrictions imposed by the linear viscoelastic assumptions do not allow for investigations at the high-pressure amplitudes linked to potential ultrasound bioeffects.<sup>8</sup> However, our results indicate that viscoelasticity is an important effect that warrants further investigation in the development of future medical ultrasound safety criteria. For certain conditions, an enhanced maximum radius and expansion ratio (maximum to minimum radius) is observed. The fact that this parameter depends directly on elasticity suggests the Mechanical Index criteria for medical ultra-

sound bioeffects assessment should eventually be reformulated to account for different tissue elasticity values and the contribution of the ultrasound pulse duration. Finally, for ultrasound contrast agent applications involving harmonic imaging, a Newtonian fluid model may not always be the best predictor of secondary harmonic responses.

The present formulation could be readily extended for application studies where the role of large bubble deformation is not particularly critical to the analysis. Such an application might be the scattering of acoustic waves off of polymeric-gel bubble clouds. These polymeric gels, seeded with gas bubbles, have been used as experimental models of oceanic bubble clouds in geophysical acoustic studies.<sup>25</sup> This gel model eliminates some of the experimental difficulties in investigating oceanic bubble clouds in the field. However, there is some question as to the accuracy of conventional scattering theory when applied to the resonance response of bubbles in such a viscoelastic material. A future paper by the authors will address this area of application in more detail. Another area of potential application is the study of free oscillations of bubbles in polymers formed due to pressure variations in the processing methods.<sup>26</sup>

However, many biomedical and industrial applications involve bubble deformations beyond the linear viscoelastic limit. Future work should consider the use of objective constitutive equations for the fluid.<sup>15</sup> Indeed, the present work serves as a foundation for such a study. The present formulation offers a straightforward, robust benchmark for direct comparison against such future models in the linear viscoelastic limit. Finally, it should be mentioned that future efforts should also include experimental verification of these unique viscoelastic results. Currently, little experimental work currently exists on forced or even free oscillations of bubbles in viscoelastic fluids. There is a paucity of published experimental work devoted to the forced (or even free) oscillations of bubbles in viscoelastic fluids. The contributions in this work should allow for a rigorous and direct comparison of theoretical predictions with future experimental results for linear viscoelasticity.

## ACKNOWLEDGMENTS

The authors wish to acknowledge the generous financial support of ONR, NIH and DARPA. Helpful comments and suggestions by James J. Riley, Lewis Wedgewood, and Larry Crum were greatly appreciated.

- <sup>1</sup>T.J. Leighton, *The Acoustic Bubble* (Academic, San Diego, 1994).
- <sup>2</sup>E.A. Neppiris, "Acoustic cavitation," *Phys. Rep.* **61**, 159–251 (1980).
- <sup>3</sup>W.L. Nyborg and D.L. Miller, "Biophysical implications of bubble mechanics" *Appl. Sci. Res.* **38**, 17–25 (1982).
- <sup>4</sup>R.E. Apfel and C.K. Holland, "Gauging the likelihood of cavitation from short-pulse-low-duty cycle diagnostic ultrasound," *Ultrasound Med. Biol.* **17**, 179–81 (1991).
- <sup>5</sup>Y.C. Fung, *Biomechanics: Mechanical Properties of Living Tissues* (Springer-Verlag, New York, 1981).
- <sup>6</sup>J.F. Vincent, *Structural Biomaterials*, 2nd ed. (Princeton University Press, Princeton, NJ, 1990).
- <sup>7</sup>J. Ophir and K.J. Parker, "Contrast agents in diagnostic ultrasound," *Ultrasound Med. Biol.* **15**, 319–33 (1989).
- <sup>8</sup>R.E. Apfel, "Possibility of microcavitation from diagnostic ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **33**, 139–142 (1982).
- <sup>9</sup>A. Shima, T. Tsujino, H. Nanjo, and N. Miura, "Study of nonlinear oscillations of bubbles in Powell-Eyring fluids," *J. Acoust. Soc. Am.* **77**, 1702–1709 (1985).
- <sup>10</sup>H.S. Fogler and J.D. Goddard, "Collapse of spherical cavities in viscoelastic fluids," *Phys. Fluids* **13**, 1135–1141 (1970).
- <sup>11</sup>I. Tanasawa and W.J. Yang, "Dynamics behavior of a gas bubble in viscoelastic liquids," *J. Appl. Phys.* **41**, 4526–4531 (1970).
- <sup>12</sup>R.Y. Ting, "Viscoelastic effect of polymers on single bubble dynamics," *AIChE J.* **21**, 810–813 (1975).
- <sup>13</sup>C. Kim, "Collapse of spherical bubbles in Maxwell fluids," *J. Non-Newtonian Fluid Mech.* **55**, 37–58 (1994).
- <sup>14</sup>A. Shima, T. Tsujino, and H. Nanjo, "Nonlinear oscillations of gas bubbles in viscoelastic fluids," *Ultrasonics* **24**, 142–147 (1986).
- <sup>15</sup>R.B. Bird, R.C. Armstrong, and O. Hassager, *Dynamics of Polymeric Liquids*, 2nd ed. (Wiley, New York, 1987), Vol. 1.
- <sup>16</sup>A.H. Nayfeh and D.K. Mook, *Nonlinear Oscillations* (Wiley, New York, 1979).
- <sup>17</sup>A.H. Nayfeh and W.S. Saric, "Nonlinear acoustic response of a spherical bubble," *J. Sound Vib.* **30**, 445–453 (1973).
- <sup>18</sup>A. Prosperetti, "Nonlinear oscillations of gas bubbles in liquids: Steady state solutions," *J. Acoust. Soc. Am.* **56**, 878–885 (1974).
- <sup>19</sup>A. Prosperetti, "Nonlinear oscillations of gas bubbles in liquids: Transient solutions and the connection between subharmonic signal and cavitation," *J. Acoust. Soc. Am.* **83**, 502–514 (1988).
- <sup>20</sup>A. Francescutto and R. Nabergoj, "Steady-state oscillations of gas bubbles in liquids: Explicit formulas for frequency-response curves," *J. Acoust. Soc. Am.* **73**, 457–460 (1983).
- <sup>21</sup>A. Francescutto and R. Nabergoj, "A multiscale analysis of gas bubble oscillations: Transient and steady-state-solutions," *Acustica* **56**, 12–22 (1984).
- <sup>22</sup>A.H. Nayfeh, "The response of single degree of freedom systems with quadratic and cubic nonlinearities to a subharmonic excitation," *J. Sound Vib.* **89**, 457–470 (1983).
- <sup>23</sup>R.E. Khayat and A. Garcia-Rejo, "Uniaxial and biaxial unsteady inflations of viscoelastic material," *J. Non-Newtonian Fluid Mech.* **43**, 31–59 (1992).
- <sup>24</sup>C. Brennen, "Some cavitation experiments in dilute polymer solutions," *J. Fluid Mech.* **44**, 51–63 (1970).
- <sup>25</sup>P.A. Hwang, R.A. Roy, and L.A. Crum, "Artificial bubble cloud targets for under-water acoustic remote sensing," *J. Atmos. Ocean. Technol.* **12**, 1289–1302 (1995).
- <sup>26</sup>H.J. Yoo and C.D. Ha, "Oscillatory behavior of a gas bubble growing or collapsing in viscoelastic liquids," *AIChE J.* **28**, 1002–1009 (1982).

# Diffraction acoustic elements for laser ultrasonics

M. Clark, S. D. Sharples, and Michael G. Somekh

*School of Electrical and Electronic Engineering, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom*

(Received 13 August 1999; revised 10 March 2000; accepted 15 March 2000)

Laser ultrasonics is an effective means of generating surface acoustic waves (SAWs). We have shown in previous publications how computer-generated holograms (CGHs) can be used to project optical distributions onto the sample surface. These can be used to control both the frequency content and the spatial distribution of the resulting ultrasound field. In this paper the concept is extended further to produce distributions which themselves act as diffractive acoustic elements (DAEs) for SAWs. It is demonstrated how frequency suppression, multiple foci, and frequency selective focusing of Rayleigh waves may be achieved with these elements. Agreement between the distributions predicted from the designs and those actually measured is excellent. © 2000 Acoustical Society of America. [S0001-4966(00)05006-2]

PACS numbers: 43.38.Zp, 43.38.Rh, 43.35.Sx [SLE]

## INTRODUCTION

Laser ultrasonics<sup>1</sup> enables samples to be tested in situations where access can be difficult and where coupling can be hazardous or degrade the measurement accuracy. The major disadvantage of laser ultrasonics over contacting technologies is that the sensitivity of optical detectors is poor compared to piezoelectric detection. This means that, in many cases, high laser powers have to be used to generate measurable ultrasonic amplitudes; this may lead to ablation of the surface. It is thus helpful if the total optical power used to generate the ultrasound is spread over the sample. Further improvements in the peak displacement can be achieved if the generated ultrasound is concentrated (focused) into a small region.

In order to increase the amplitude of the surface acoustic waves (SAWs) generated on the sample while minimizing the sample damage through ablation, two approaches have been used by previous authors. One method to increase the SAW amplitude involves passing the light from the generating laser through an axicon,<sup>2</sup> so that it focuses onto an annulus on the sample. This results in the generation of a converging (and diverging) surface acoustic wave, so that the amplitude of the surface acoustic wave at the center of the annulus is considerably increased. The second approach involves spreading the power of the generating beam using a moving grating formed by interfering two optical beams with slightly different frequencies.<sup>3</sup> Matching the wavelength of the grating to that of the surface wave at the difference frequency allows a strong ultrasonic signal to be generated with relatively low power density on the surface. Another related approach involves scanning a point line source at the phase velocity of the required surface wave.<sup>4</sup>

In this paper we discuss the design of a range of different computer-generated holograms (CGHs) which focus the SAW distributions as well as performing various filtering tasks on the ultrasonic distribution. The method is robust, allows extremely stable operation and, as we will show, is very flexible compared to other methods. We have reported previously on CGHs, which project an arc onto the surface of

the sample, providing a diffraction-limited SAW (Rayleigh wave) focus.<sup>5,6</sup> These CGHs are analogous to using a refracting optical lens because they produce a continuous variation of the acoustic phase to effect the focusing. It is also possible to produce an analog of an optical diffractive CGH where control of the resulting beam is affected by discretely varying the phase of the ultrasonic source distribution. We describe continuous, discrete, and hybrid elements, which may be thought of as the Rayleigh wave analog of hybrid refractive/diffractive optical elements. We introduced the concept of diffractive acoustic elements (DAEs) and hybrid elements in Ref. 7; however, the present paper describes a wider range of elements performing many different tasks and we also describe special considerations that need to be borne in mind when designing these elements.

Section I describes the experimental setup and briefly discusses the method used to model the SAW distributions. In view of the distinction between continuous and discrete SAW elements, the rest of the paper is divided up as follows: Sec. II describes continuous distributions, used for focusing and frequency selection; Sec. III describes the application of discrete and hybrid distributions; Sec. IV discusses some of the design issues related to the discrete phase elements and considers the prospects for the elements discussed in this paper.

## I. EXPERIMENTAL SETUP AND MODELING SAW DISTRIBUTIONS

The experimental setup used in this paper has been described elsewhere,<sup>8</sup> so a brief description will suffice. The system is shown schematically in Fig. 1. A mode-locked *Q*-switched laser was used to excite the ultrasonic beam. The laser produced a burst of approximately 30 short pulses each separated by 12 ns. Each burst was repeated every millisecond. The light from the source was then passed through different CGHs, which focused the beam onto the sample surface to form the profiles described throughout the paper. The resulting SAW distributions were detected with a specialized knife edge detector<sup>6</sup> which was mechanically scanned rela-

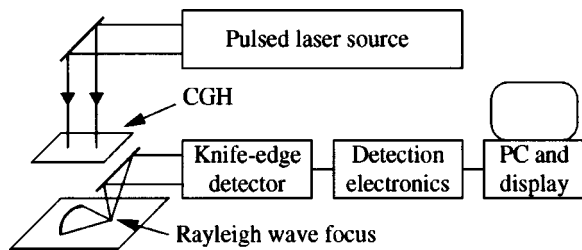


FIG. 1. Schematic diagram of laser ultrasound system. (CGH=computer-generated hologram.)

tive to the sample, while the illumination optics remained fixed. The peak envelope of the ultrasonic tone bursts were normally acquired using the envelope detection electronics described in Ref. 9.

The CGHs were designed using a direct search algorithm described elsewhere;<sup>10</sup> these produced the desired optical distributions on the sample surface. All elements were binary phase and etched in quartz. The CGHs profiled and shaped the laser beam so that no additional optical elements were required. The CGH thus replaces a focusing cylindrical lens in a conventional laser ultrasonic system.

In all the experiments the system operated in the thermoelastic regime and no surface treatment was performed on the samples. Inspection of the samples after the experiments revealed no detectable damage.

In order to predict the surface wave distributions expected from the samples, the predicted light distribution was calculated from the CGH design. This light distribution was then used to calculate the ultrasonic distribution close to the generation point. Since we are confident of always operating in the thermoelastic regime, the model of Krylov and Pavlov<sup>11</sup> was used, which predicts that the surface displacement due to Rayleigh wave motion,  $a_r$ , is given by the following relationship:

$$a_r \propto \frac{1}{\omega_r} F_s(k_r) F_t(\omega_r), \quad (1)$$

where  $F_s(k_r)$  is the Fourier transform of the spatial distribution of the source intensity at the SAW wave number  $k_r$  and  $F_t(\omega_r)$  is the corresponding temporal frequency component of the source at SAW frequency  $\omega_r$ . This SAW distribution is then propagated along the sample using angular spectrum propagation as described by Goodman.<sup>12</sup> Here the scalar field of the SAWs at a line is decomposed into plane waves (the angular spectrum) by taking the Fourier transform of the complex amplitude of the SAW at the line. These plane waves are then propagated to the reconstruction line using an appropriate propagation factor,  $\exp jk_r \cos \theta z$ , where  $z$  is the propagation distance between the source and reconstruction line and  $\theta$  is the angle between the plane wave and the axis of propagation (see Goodman<sup>12</sup>). The simulated SAW amplitude is then reconstructed from the resultant plane waves (by taking the inverse Fourier transform).

## II. CONTINUOUS DISTRIBUTIONS FOR FOCUSING AND FREQUENCY CONTROL

### A. SAW focusing

Focusing the Rayleigh wave distribution increases the displacement on the sample, thus improving the detectability of the surface wave distribution. In previous publications we have shown that this allows one to detect the ultrasonic distribution without averaging.<sup>6,9</sup> We have also shown how the increase in surface wave amplitude allows fast analog electronics to be used rather than a digital storage oscilloscope; this greatly improves the rate of image acquisition obtained in a noncontacting surface wave acoustic microscope.<sup>9</sup> The light distributions on the sample were obtained using a CGH, which produced a single arc. This provided focusing of the laser-generated surface waves, but offered little control of the frequency content of the resulting surface wave distribution. In this section the use of several concentric arcs is described; this increases the maximum power that the sample can withstand without ablation, thus providing an even greater enhancement of the SAW amplitude at the focus. The spacing of the arcs determined the frequency selection.

### B. Frequency control: High pass filtering

Enhancing the harmonic signal and suppressing the fundamental can be important when trying to image at high frequencies where the signal at the fundamental is often much larger and may tend to swamp the desired harmonic. Figure 2 shows the SAW focus obtained at the focus of a four-arc distribution, which is shown schematically in Fig. 2(a). The separation between adjacent arcs is half a wavelength at 82 MHz and thus a whole wavelength at 164 MHz. This means that the excitation from each 82-MHz line cancels out, whereas the 164-MHz signals add in phase. The distributions shown in Fig. 2 clearly demonstrate this, since no 82-MHz signal can be detected above the noise in Fig. 2(b) and (d), whereas a diffraction-limited focus for the 164-MHz signal was observed in Fig. 2(c) and (d).

### C. Frequency control: Low-pass filtering

Harmonic imaging involves excitation at a fundamental frequency and detection of the harmonic generated by nonlinearities in the sample. Harmonic imaging has been used in the fluid coupled scanning acoustic microscope but the large nonlinearity in the fluid means that more of the generation occurs in the couplant rather than the sample. In conventional fluid coupled acoustic microscopy, information about the couplant liquids and biological samples<sup>13</sup> may be obtained, but little information about the nonlinear acoustic properties of hard solid materials is accessible. Since laser ultrasonics does not use a couplant it is a promising technique for SAW harmonic generation in solid materials, particularly at cracks and discontinuities. The DAE described below is well-suited for such an application. To successfully perform harmonic imaging it is essential that any harmonic content in the input signal be suppressed so that it is not confused with the generated harmonic. To ensure efficient suppression of the harmonic a two-arc distribution was used

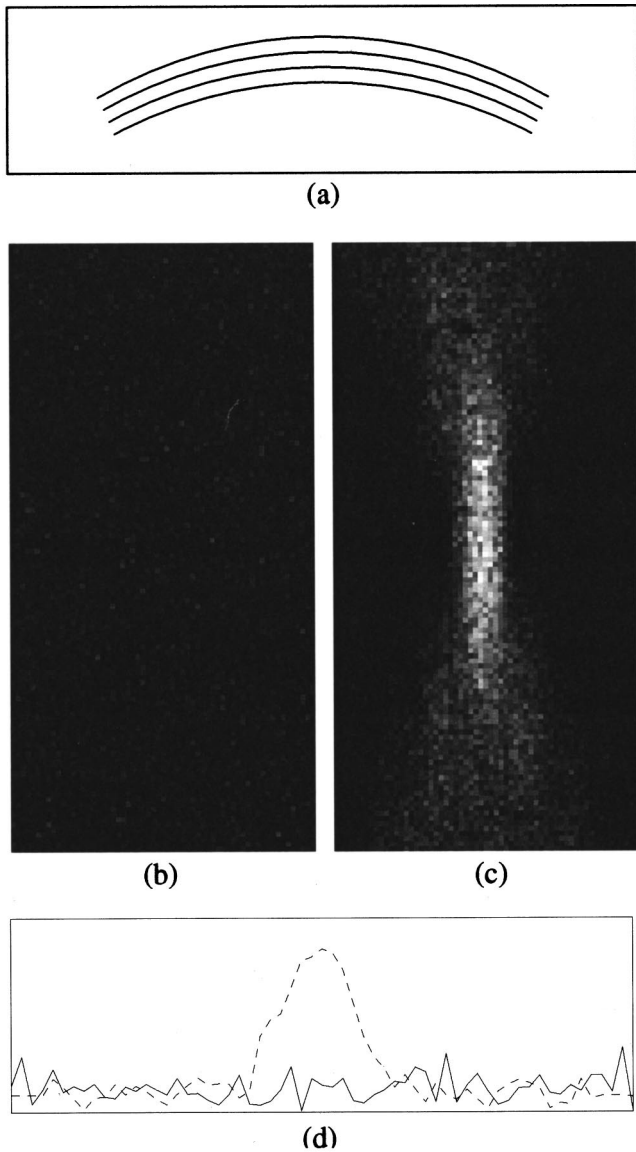


FIG. 2. Focus of a four-arc distribution, where each arc is separated by half a wavelength at 82 MHz. (a) Schematic of four-arc optical distribution; (b) no observable signal at 82 MHz; (c) strong diffraction-limited focus at 164 MHz; (d) Line plot through the focus, solid line 82 MHz, dashed 164 MHz. Image size 600 by 300  $\mu\text{m}$ . Ultrasonic source 1 cm above center of image.

[Fig. 3(a)]. In this case the arcs were separated by three half-wavelengths at 164 MHz, thus ensuring the cancellation of this signal. The separation at the fundamental frequency is therefore three quarter-wavelengths, so that the waves generated at 82 MHz from each line add in quadrature. This means that the amplitude of the resulting surface wave at the fundamental frequency has a large amplitude [see Fig. 3(b)] although it is reduced by a factor of  $1/\sqrt{2}$  compared to the optimum case, in which all the contributing arcs add in phase rather than quadrature. On the other hand, the harmonic signal is suppressed, so that it is not visible above the noise [see Fig. 3(c) and (d)].

### III. DISCRETE DIFFRACTIVE ACOUSTIC ELEMENTS

In Sec. II we showed how interference between SAWs could be used to suppress different frequency components; in

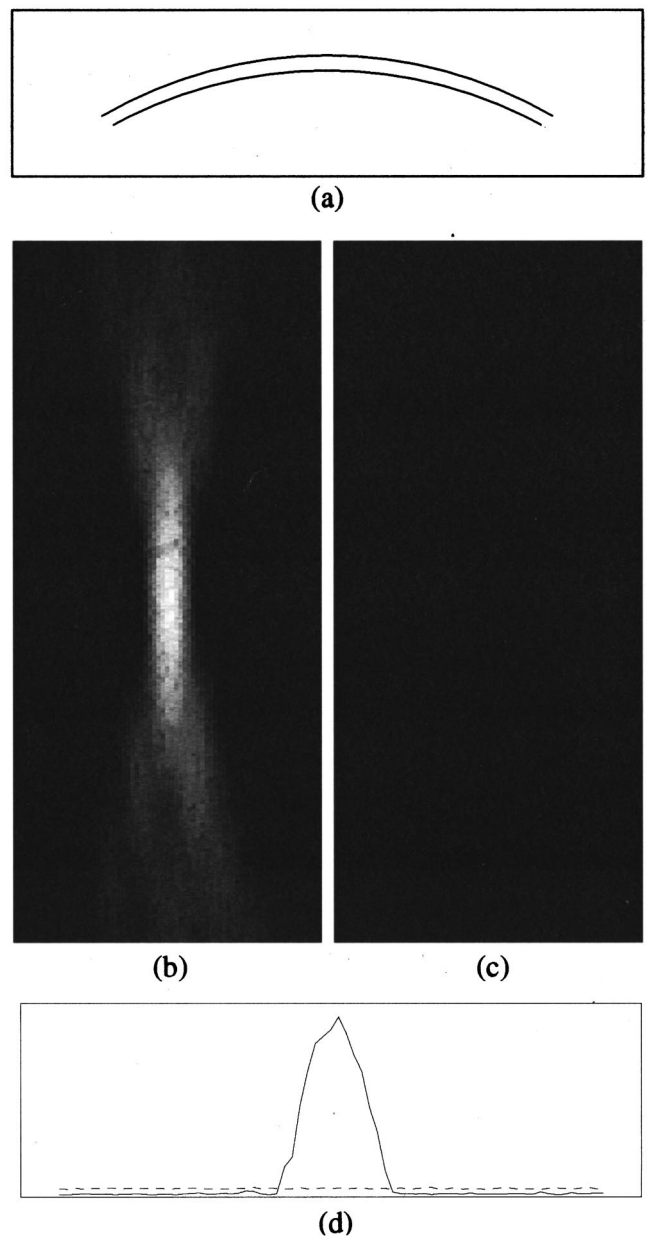


FIG. 3. Focus of two-arc hologram, where each arc is separated by three quarters of a wavelength at 82 MHz. (a) Schematic of two-arc optical distribution; (b) strong diffraction-limited focus at 82 MHz; (c) no observable signal at 164 MHz; (d) line plot through the focus, solid line 82 MHz, dashed 164 MHz. Image size 600 by 300  $\mu\text{m}$ . Ultrasonic source 1 cm above center of image.

this section we will show how CGHs may be used to produce discrete distributions, which themselves act as DAEs. DAEs, like optical diffractive elements, are very dispersive; this dispersion can be inconvenient or it can be useful to produce, for instance, frequency-dependent surface wave focusing.

#### A. Discrete acoustic element for focusing

In optics, one of the simplest diffractive elements is the binary phase plate, this introduces 180-deg phase shifts between adjacent zones, so that it acts as a “discrete” lens. For a binary (two-level) phase plate a parallel beam incident on the element will produce both a converging and a diverging beam from the +1 and -1 diffracted orders. Projecting an



array of short lines offset from each other by half the Rayleigh wave wavelength forms a DAE which acts as a surface wave analog of a zone plate [see Fig. 4(a)]. Ideally, this should produce a focus of SAWs. Figure 4(b) shows the predicted diffraction pattern from the ultrasonic phase plate of Fig. 4(a). The experimental focal plane distribution is shown in Fig. 4(c) and we note that the agreement is not particularly good. The reason for this is that there is a large path difference between extreme ray paths for this element. This path difference is approximately 1.4 mm, which corresponds to a relative time delay of 450 ns or 37 cycles of 82-MHz signal (the number of zones between the center and the edge of the phase plate). The large time delay means that all the surface waves do not interfere at the focus, due to the short duration of the  $Q$ -switched envelope (approximately 360 ns). This can be seen from the waveforms shown in Fig. 4(d). The top trace shows the compact waveform from a focusing arc; the middle trace shows the envelope at the focus obtained from the surface wave zone plate, where one may note the spreading; the bottom trace was obtained at the position of the expected first minimum of the focal distribution. This shows considerable spreading of the waveform as well as a phase dislocation, indicating that for a single frequency wave there would be phase cancellation. To perform effective interference it is necessary for the length of the individual pulse trains to be considerably longer than the maximum time delay between different paths. This can be achieved by passing the electronic signal through a very narrow bandpass filter. This was performed digitally by detecting the output ultrasonic waveform from the optical detector in the digital storage oscilloscope. Recall, that for the remaining measurements in this paper only the envelope of the output signal was detected. The waveforms detected at each scan point were then Fourier transformed and the component at 82 MHz was recovered. This was then used to form the corrected focal distribution, which is shown in Fig. 4(f), next to an expanded version of the theoretical distribution, Fig. 4(e). We can now see that the focal distribution from the phase plate agrees quite well with the expected distribution for a single frequency pulse.

## B. Hybrid elements for frequency-dependent focusing

The simple phase plate is clearly unsatisfactory if we wish to use an envelope detector. The solution to this problem is to use a hybrid continuous/discrete DAE. In this case a weak phase plate was imposed on an arc as shown in Fig. 5(a). The arc forms the more powerful focusing element and the phase plate imposes a relatively weak perturbation; since the diffractive element imposed on the arc has only six zones between the center and the edge, effective interference can take place from our 30-cycle source. As mentioned earlier, the zone plate gives a  $+1$  and a  $-1$  diffracted order so that two foci are formed—one inside the geometrical focus of the arc and one outside. If the focal length of the arc is 10 mm and the focal length of the zone plate is 76 mm, we expect the two foci to be located at 8.8 and 11.5 mm, respectively. Figure 5(b) and (c) show excellent agreement between the simulated and measured focal distributions.

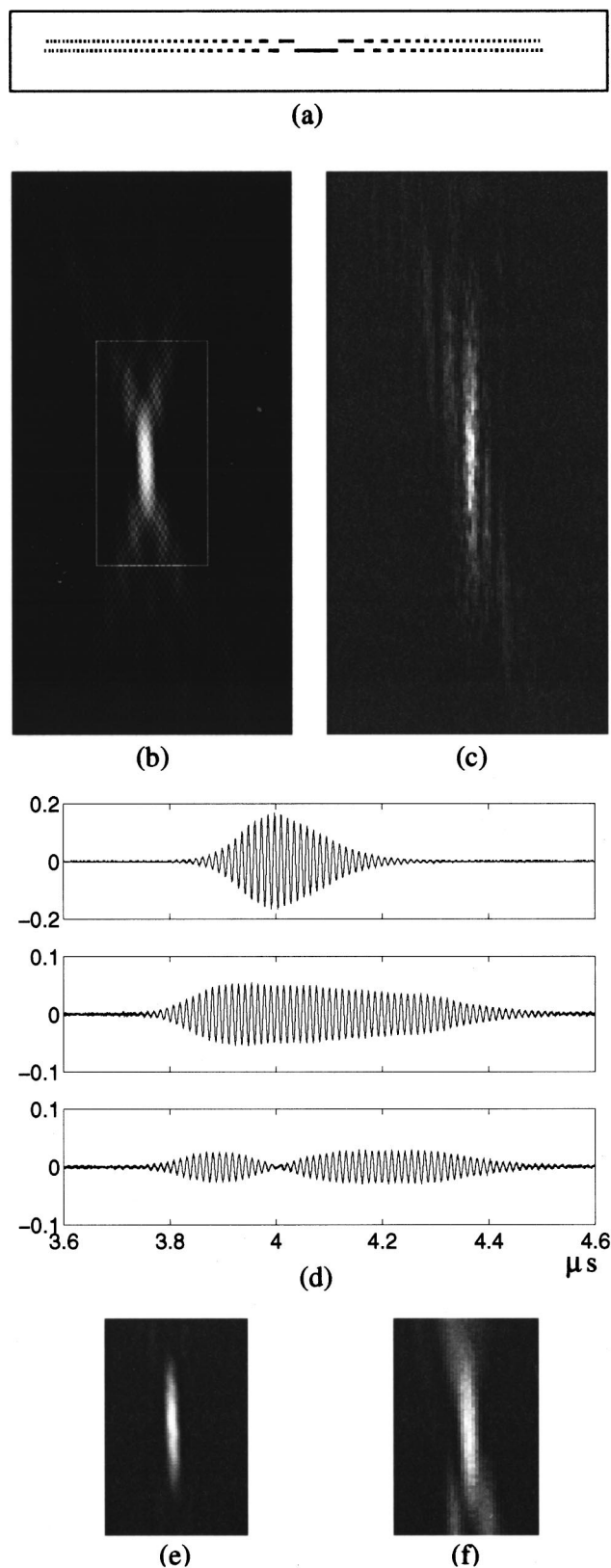


FIG. 4. Results obtained with acoustic zone plate. (a) Schematic of optical distribution; (b) predicted point spread function at focus. Image dimensions 2 by 1 mm; (c) measured point spread function, image dimensions as (b); (d) waveforms showing amplitude versus propagation time from generation source (microseconds). Top trace waveform at focus from an arc with same focal length, middle trace waveform at focus, bottom trace waveform at position of first zero (notice phase dislocation in waveform); (e) zoom of predicted point spread function (b), image size 600 by 400  $\mu\text{m}$ ; (f) measured point spread function after filtering. Image dimensions 600 by 400  $\mu\text{m}$ .

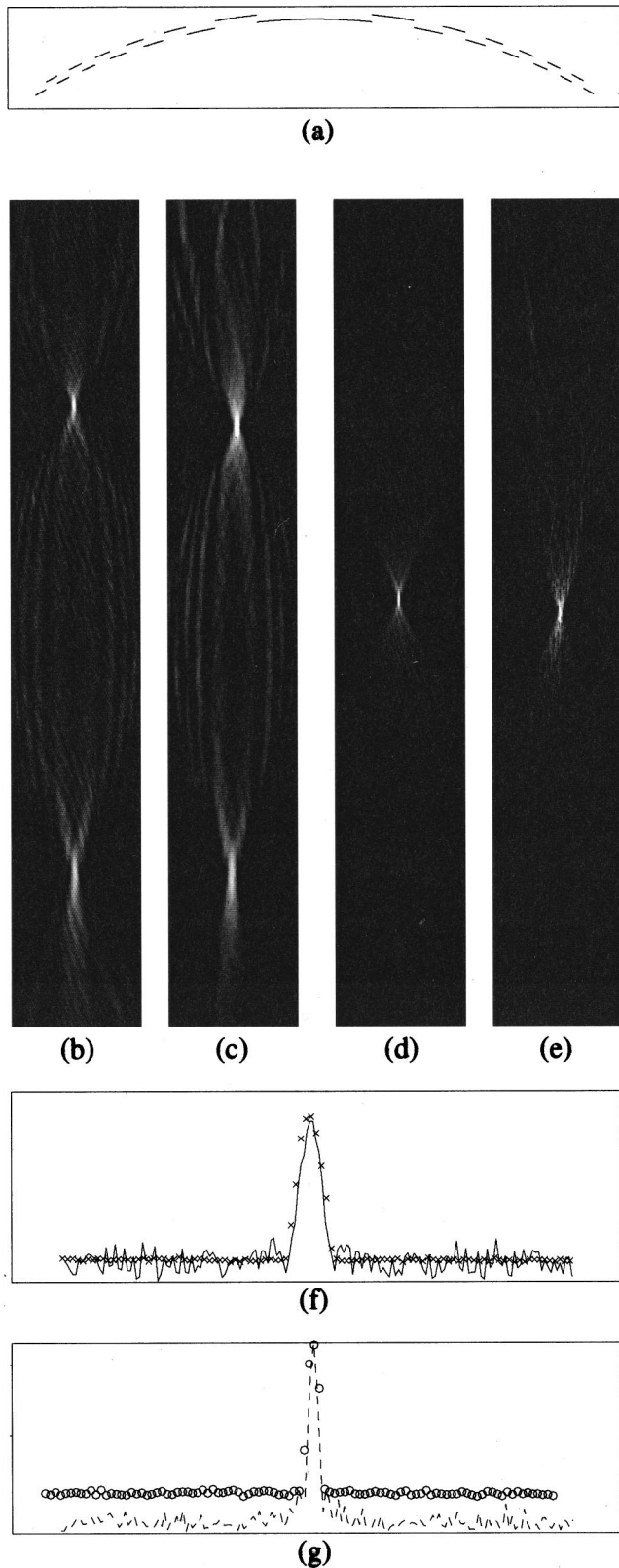


FIG. 5. Hybrid diffractive element with superimposed continuous wave front, with two focal lengths. (a) Schematic of optical distribution; (b) predicted focal distribution at 82 MHz; (c) measured distribution at 82 MHz; (d) predicted focal distribution at 164 MHz; (e) measured distribution at 164 MHz; (f) line plot through the first 82 MHz focus, solid line 82-MHz simulation, crosses measured SAW amplitude at 82 MHz. (g) line plot through the 164-MHz focus, dashed line 164-MHz simulation, circles measured SAW amplitude at 164 MHz. Image size 6.4 mm by 0.8 mm.

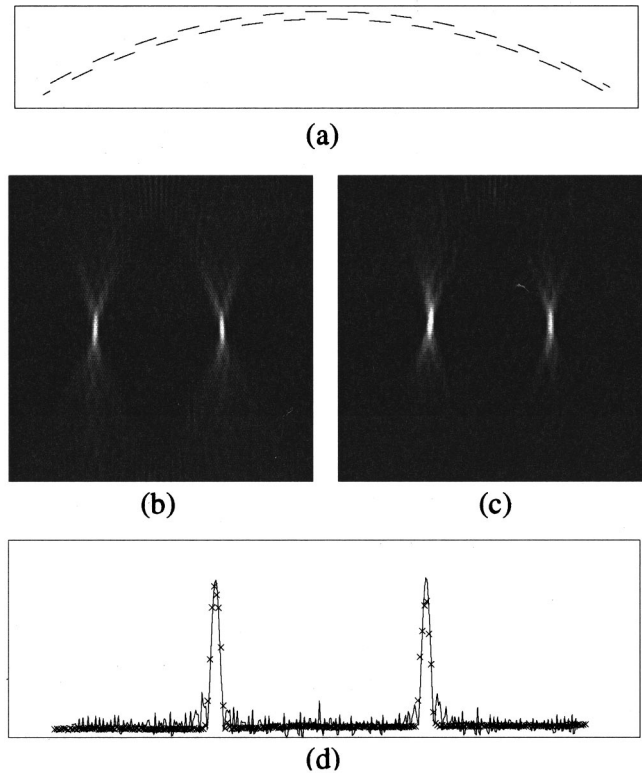


FIG. 6. Hybrid diffractive element with superimposed continuous wave front, with laterally offset foci. (a) Schematic of optical distribution; (b) predicted distribution from DAE showing laterally displaced foci at 82 MHz; (c) measured distribution from DAE showing laterally displaced foci at 82 MHz; (d) line plot through foci, solid line 82-MHz simulation, crosses measured SAW amplitude at 82 MHz. Image size 2.4 by 2.4 mm.

The half-wavelength phase delay at 82 MHz imposed by the displaced elements gives a whole wavelength relative delay at 164 MHz, so that the phase plate is “invisible” to the fundamental frequency. The focus for 164 MHz thus occurs at the geometrical focus of the arc. This is verified in Fig. 5(d) (predicted) and (e) (measured). The quantitative agreement between measured and predicted distributions may be determined from Fig. 5(f) and (g), which show the measured and predicted plots at 82 and 164 MHz, respectively.

### C. Hybrid elements for displaced focusing

In addition to controlling the focal position along the axial direction, a combined continuous discrete element can also act as a Rayleigh wave analog of a Nomarski objective in optics, which forms two adjacent foci. This is achieved using an arc on which a grating is imposed [Fig. 6(a)], the +1 diffracted order introduces a phase tilt on the beam, which displaces the focus upwards and, similarly, the -1 order displaces the focus downwards. The simulations and the measured distributions are shown in Fig. 6(b) and (c), respectively, and a lineplot comparing experiment and theory is shown in 6(d); once again the agreement between experiment and theory is striking.

In order to demonstrate the flexibility of our approach we once again used a continuous arc with a superimposed discrete phase plate, but rather than maintain the symmetry

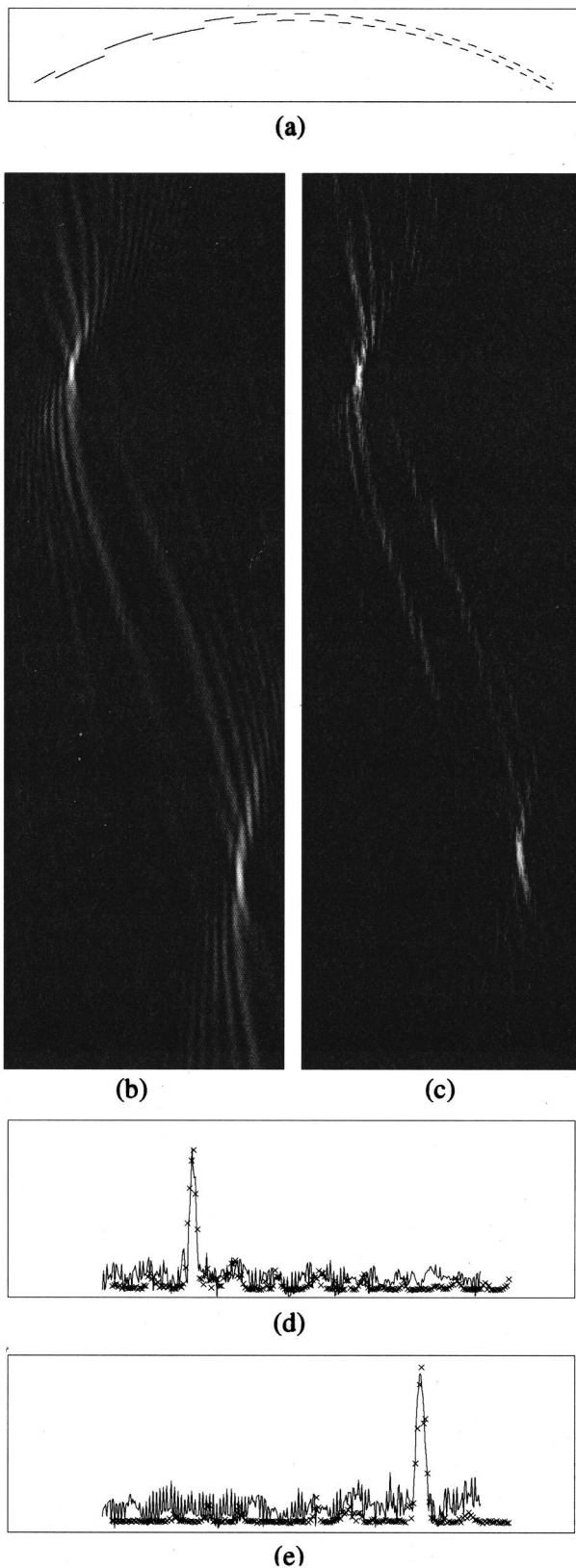


FIG. 7. Hybrid diffractive element with superimposed continuous wave front, with focuses offset both laterally and axially. (a) Schematic of optical distribution; (b) predicted distribution from DAE showing laterally and axially displaced foci; (c) measured distribution from DAE showing laterally and axially displaced foci; (d) and (e) line plots through first and second foci, respectively, solid lines 82-MHz simulation, crosses measured SAW amplitude at 82 MHz. Image size 6.4 by 2 mm.

of this phase plate about the axis the plate was displaced [Fig. 7(a)]. The +1 and -1 order focuses now differ not only in their axial position, but also in their height relative to the axis. Figure 7(b)–(e) show the excellent agreement between the simulated and measured distributions, even showing extremely close agreement in the fine detail of the wave distribution between the two foci.

#### IV. DISCUSSION

We have shown how from a specific optical design it is possible to generate an arbitrary optical distribution in the sample surface, which, in turn, acts as a Rayleigh wave diffractive element. The surface wave distribution is predicted entirely from the design of the optical element and gives excellent agreement. Combining curved sources with discrete phase plates gives control of the focal distributions and achieves multiple foci as well as frequency selective focusing.

We note that the DAEs described here differ from their optical counterparts in two important respects. First, the number of zones is by necessity rather small, to ensure good interference over a finite duration acoustic pulse; second, these DAEs are essentially one-dimensional—as opposed to two-dimensional—elements, since they are concerned with the focusing of SAWs. When phase plates are used in optics, the only criterion necessary to ensure suppression of the zero order is that there is 180-deg phase shift between adjacent zones. The large number of zones and the approximately equal areas in each zone—arising from the two-dimensional nature of the element—ensures that the areas corresponding to each phase shift are approximately equal. To ensure good zero-order suppression with DAEs, a little more care is required in the design, to ensure that the total line lengths (and hence generated SAW power) corresponding to each phase shift are approximately equal. In our designs the 0-order of diffraction is undetectable with a signal-to-noise ratio of 100:1 (at the peak of a focus).

In this paper we have concentrated our discussion on SAWs; however, concepts described here can also be applied to bulk waves. Projection of a distribution, which acts as a DAE for bulk waves, may be readily produced by projecting successive light and dark patches onto the sample. This element will act as an amplitude zone plate rather than a phase plate, but will produce a well-defined focus in the bulk of the sample. The methodology we have described has been applied to flat surfaces only, but the method should find application to curved surfaces, as the CGHs may be readily designed to produce a desired optical distribution over any arbitrary surface.

In future work we intend to use hybrid continuous/discrete elements to enhance the operation of a noncontacting SAW microscope and, in particular, examine the use of the frequency selectivity of these elements to facilitate harmonic imaging of defects.

#### ACKNOWLEDGMENTS

We are grateful to the Engineering and Physical Science Research Council and Rolls Royce plc for supporting this

work. The CGHs used in this study were designed in Nottingham and fabricated at Glasgow University.

- <sup>1</sup>C. B. Scruby and L. E. Drain, *Laser Ultrasonics, Techniques and Applications* (Adam Hilger, Bristol, UK, 1990).
- <sup>2</sup>X. Maldague, P. Cielo, and C. K. Jen, "NDT applications of laser generated focused acoustic waves," *Mater. Eval.* **44**(9), 1120–1124 (1986).
- <sup>3</sup>H. Nishino, Y. Tsukahara, Y. Nagata, T. Koda, and K. Yamanaka, "Excitation of high-frequency surface acoustic-waves by phase-velocity scanning of a laser interference fringe," *Appl. Phys. Lett.* **62**(17), 2036–2038 (1993).
- <sup>4</sup>K. Yamanaka, O. V. Kolosov, Y. Nagata, T. Koda, H. Nishino, and Y. Tsukahara, "Analysis of excitation and coherent amplitude enhancement of surface acoustic-waves by the phase-velocity scanning method," *J. Appl. Phys.* **74**(11), 6511–6522 (1993).
- <sup>5</sup>M. Liu, H. P. Ho, M. G. Somekh, and J. M. R. Weaver, "Noncontacting optical-generation of focused surface acoustic-waves using a customized zoneplate," *Electron. Lett.* **31**(4), 264–265 (1995).
- <sup>6</sup>F. Linnane, M. Clark, M. G. Somekh, and D. Zhang, "Surface acoustic

- wave generation with customized optical beams," 1996 IEEE Ultrasonics Symposium, pp. 479–483.
- <sup>7</sup>M. Clark, S. D. Sharples, and M. G. Somekh, "Noncontact continuous wavefront/diffractive acoustic wave elements for Rayleigh wave control," *Appl. Phys. Lett.* **74**(24), 3604–3606 (1999).
- <sup>8</sup>M. Clark, F. Linnane, S. D. Sharples, and M. G. Somekh, "Frequency control in laser ultrasound with computer generated holography," *Appl. Phys. Lett.* **72**(16), 1963–1965 (1998).
- <sup>9</sup>M. Clark, S. D. Sharples, M. G. Somekh, and A. S. Leitch, "Non-contacting surface acoustic wave microscope," *Electron. Lett.* **35**(4), 346–347 (1999).
- <sup>10</sup>M. Clark and R. Smith, "A direct search method for the computer design of holograms," *Opt. Commun.* **124**, 150–164 (1996).
- <sup>11</sup>V. V. Krylov and V. I. Pavlov, "Thermo-optical generation of surface acoustic waves in a solid," *Sov. Phys. Acoust.* **28**(6), 493–494 (1982).
- <sup>12</sup>J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, New York, 1996).
- <sup>13</sup>L. Germain and J. D. N. Cheeke, "Generation and detection of high-order harmonics in liquids using a scanning acoustic microscope," *J. Acoust. Soc. Am.* **83**(3), 942–949 (1988).

# Power flow in coupled bending and longitudinal waves in beams

Richard P. Szwerc

*Carderock Division, Naval Surface Warfare Center, Bethesda, Maryland 20817-5700*

Courtney B. Burroughs, Stephen A. Hambric, and Timothy E. McDevitt

*Applied Research Laboratory, The Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804*

(Received 14 October 1999; accepted for publication 10 March 2000)

In complex structures, curvature and impedance discontinuities (e.g., junctions) couple bending and longitudinal waves. Propagation losses for longitudinal waves are often much less than losses for bending waves, and damping treatments often less effective on longitudinal waves. When the dissipation in longitudinal waves is less than that on bending waves, longitudinal waves can provide an efficient means of power flow between bending waves generated at one location and bending waves that are a source of acoustic radiation at another location. In order to design and locate effective treatments, knowledge of the power flow in longitudinal as well as bending waves is required. The measurement of power in both bending and longitudinal waves when both waves are present is demonstrated. Measurements conducted on a straight beam and a T-beam are compared to predictions obtained using finite element methods. The effect of coupling between waves at the junction in a T-beam is illustrated using results from measurements of power flow. © 2000 Acoustical Society of America. [S0001-4966(00)03906-0]

PACS numbers: 43.40.Cw [PJR]

## INTRODUCTION

Curvature and impedance discontinuities (e.g., junctions) in structural systems couple bending (flexural) and longitudinal (in-plane) waves. Because transverse displacements are dominant in bending waves, they are usually a greater source of acoustic radiation and therefore have received more attention in the development of structural intensity measurement methods than have longitudinal waves. Bending waves have shorter wavelengths than longitudinal waves and are more subject to dissipation by added damping treatments where shear stresses are more readily transmitted into the treatment by bending waves than by longitudinal waves. Because of a higher dissipation rate, bending waves typically transport energy over shorter distances than longitudinal waves. Therefore, longitudinal waves, generated by the incidence of bending waves at structural junctions, can be an efficient means of transporting energy in structures from one location to another, where coupling back to bending waves may occur to produce acoustic radiation. It is therefore important to track energy transmission in both bending and longitudinal waves so that effective treatments can be designed and located. Treating only bending waves may result in flanking paths for longitudinal waves that reduce the effectiveness of the treatment.

The first objective of this paper is to validate a method for the measurement of the spatial distribution of power in bending and longitudinal waves when both waves are present with unknown amplitudes. Measurements will be validated by comparison to predictions generated by finite element methods. The second objective is to employ both measurements and predictions to investigate the effects of structural junctions on the coupling of bending and longitudinal waves. To meet these objectives, measurements and predictions of power are made on a T-beam (see Fig. 1). This is one of the

simplest structures in which bending and longitudinal waves are coupled.

## I. DESCRIPTION OF T-BEAM

Two beam structures are used, (1) a straight beam and (2) a T-beam, both constructed from the same  $2\frac{1}{4}$ -in.-thick sheet of thermoplastic. This material was chosen over metal primarily for the high material damping. The damping reduced standing waves and provided a decrease in structural intensity with distance along the beam from the source. The straight beam had the same dimensions as the base of the T-beam. The measured Young's modulus for the beam was  $2.62 \times 10^9$  N/m<sup>2</sup> and the density 1240 kg/m<sup>3</sup>. In this work, the beams were supported by soft springs at the ends to simulate free ends. Vibration was excited with a mechanical shaker attached at one end. The shaker was oriented to either excite bending or longitudinal waves.

## II. INTENSITY MEASUREMENTS

### A. Experimental method

Noiseaux<sup>1</sup> developed an expression for far-field structural power flow in terms of transverse and rotational velocities so that a measurement of the transverse and rotational velocity at only one point was required. Pavic<sup>2</sup> extended the work of Noiseaux by developing expressions for the power flow of bending vibration in the near and far fields which required measurements of only the transverse velocity. Verheij<sup>3</sup> described how the frequency domain could be exploited using cross-spectral density measurements to measure the power flow in bending or longitudinal waves in beams and pipes. Linjama and Lahti<sup>4</sup> expanded the work of Verheij to include the power flow in near and far fields.

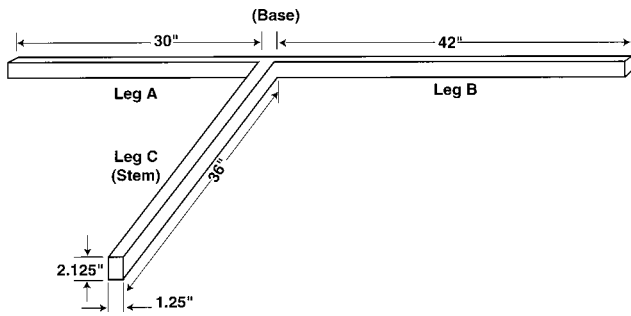


FIG. 1. Sketch of T-beam with dimensions.

Using principles given by Verheij, McDevitt *et al.* conducted measurements of power in bending<sup>5</sup> and longitudinal<sup>6</sup> waves in a beam using a two-channel laser. However, coupling of transverse and in-plane motions in bending and longitudinal waves limited these measurements to isolated bending or longitudinal waves in a straight beam. This paper validates the approach of Verheij<sup>3</sup> using an extension of the experimental techniques of Linjama and Lahti.<sup>4</sup> Verheij<sup>3</sup> presents a method and structure-borne power data as a function of frequency at a single location on a complex structure. Although his paper shows the relative powers of different wave types, the data are not rigorously substantiated. In this paper, the measured input and structure-borne powers are compared with numerical predictions, and the decay in the spatial power flow fields are measured. Thus the reader can observe that the structure-borne powers compare well to the input powers near the drive point, and decay to zero at free ends as expected and are in agreement with numerical predictions.

To determine the power in bending and longitudinal waves, when they coexist, it is necessary to determine the vibration amplitudes and phases of both waves. In this paper, works presented by Linjama and Lahti<sup>4</sup> and Verheij<sup>3</sup> are combined to develop equations and a measurement technique which depends only on two sensors (an accelerometer and a force gage) for the determination of power flow in both bending and longitudinal waves in the presence of both waves. The accelerometer is moved to several locations to complete the required measurements. For bending waves, the transverse velocities at four locations, two on opposite sides of the beam at one axial location and two others on opposite sides at another axial location within a small fraction of a bending wavelength of the first pair, are made. For longitudinal waves, velocity measurements are also made at locations in a similar configuration, except measurements are made with the axes of the accelerometers parallel to the axis of the beam, instead of perpendicular as was the case for power measurements of bending waves. Ming and Craik<sup>7</sup> examined errors in the measurement of structure-borne power in beams, and recommend against the use of four acceleration measurements to separate longitudinal and bending wave powers because of inevitable phase differences between the sensors.

Both the bending and longitudinal waves have displacements in the transverse and axial directions at the surface of the beam. As observed by accelerometers attached transversely to opposite sides of a beam, the transverse displace-

ments for bending waves on both surfaces are completely out of phase. Thus, the bending wave signal may be separated from the longitudinal by

$$H_b = \frac{H_{tt} - H_{tb}}{2}, \quad (1)$$

where  $H_b$  is the bending wave acceleration frequency response and  $H_{tt}$  ( $H_{tb}$ ) is the transverse accelerometer response on the top (bottom) surface of the beam. As observed by accelerometers attached axially to opposite surfaces of the beam, the axial displacements for longitudinal waves on both surfaces are in phase and the axial displacements for bending waves are completely out of phase. Thus, longitudinal waves may be separated from bending waves by

$$H_l = \frac{H_{at} + H_{ab}}{2}, \quad (2)$$

where  $H_{at}$  ( $H_{ab}$ ) is the axial accelerometer response on the top (bottom) surface of the beam.

The use of a single roaming accelerometer removes phase errors introduced by phase differences in the response of multiple accelerometers. The cross-spectral density between a wave amplitude at two adjacent measurement points is estimated from<sup>8</sup>

$$G_{12}^b = H_{b1}^* H_{b2} G_{ff}, \quad (3)$$

where  $G_{ff}$  is the auto spectral density of the reference signal,  $H_{b1}$  and  $H_{b2}$  are the frequency responses of the signal from the accelerometer at positions 1 and 2, respectively, divided by the reference signal, and  $G_{12}^b$  is the estimated cross-spectral density between the bending wave acceleration signals at positions 1 and 2. The \* indicates a complex conjugate.

Using Eq. (4), the far-field structure-borne power in the bending wave,  $P_B$ , was then estimated from the measurements made with the scanned transverse accelerometer, using Eq. (1) to remove the contribution to the transverse displacement by longitudinal waves:

$$P_B = 2 \frac{(Dm)^{1/2}}{\Delta \omega^2} \text{Im} \{G_{12}^b\}. \quad (4)$$

The signal from the transverse accelerometer is referenced to the impedance head force gage. In Eq. (4),  $D$  is the bending rigidity,  $m$  is the mass per unit length of the beam,  $\Delta$  is the distance between measurement locations,  $\omega$  is the radian frequency, and  $\text{Im}$  indicates the imaginary part. Power can be measured in the near field of the bending wave using a modified form of Eq. (4) and four measurement locations, rather than two locations.<sup>9</sup> Using Eq. (5), the structural power in the longitudinal waves,  $P_L$ , was estimated from the measurements made with the scanning axial accelerometer using Eq. (2) to remove the contribution to the axial acceleration by bending waves:

$$P_L = \frac{SE}{\Delta \omega^3} \text{Im} \{G_{12}^l\}, \quad (5)$$

The superscripts "b" and "l" in Eqs. (4) and (5) refer to the bending and longitudinal waves, respectively. In Eq. (5),  $S$  is

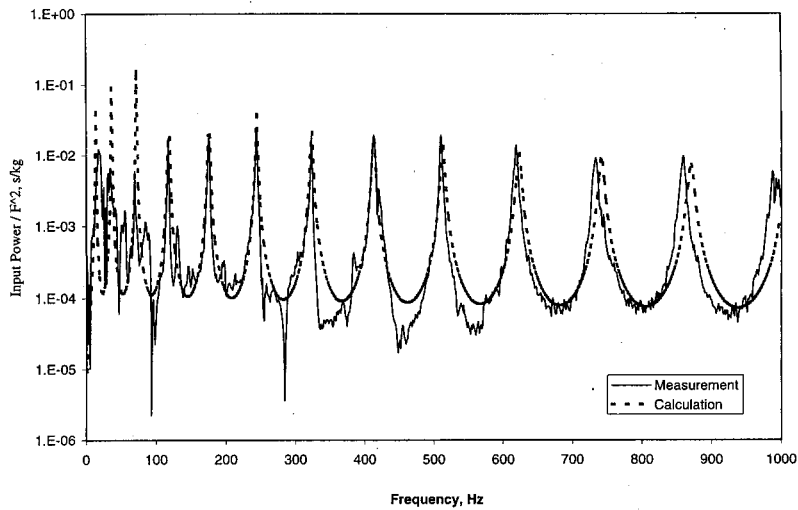


FIG. 2. Predicted and measured bending wave power input to transversely driven straight beam.

the cross-sectional area and  $E$  is Young's modulus for the beam. Equations (4) and (5) apply only to structures with constant thickness and when the measured quantity is acceleration. Also, note that the power as obtained via Eqs. (4) and (5) must be adjusted by a finite difference correction as necessitated by the discrete transducer spacing,  $\Delta$ . Specifically, the result must be multiplied by the term  $k\Delta/\sin k\Delta$ , where  $k$  is the appropriate (bending or longitudinal) wave number. This adjustment has been made for all of the measurements reported.

In addition to the powers in bending and longitudinal waves along the beams, the power inputs to the beams were also measured. The input power was determined from

$$P_i = \frac{\text{Im}\{G_{fa}\}}{\omega}, \quad (6)$$

where  $G_{fa}$  is the cross-spectral density between the force and acceleration outputs of an impedance head, which is located between a stinger at the output of the shaker, and the beam. The end of leg A (see Fig. 1) was driven by the shaker in either of two orientations; one transverse to the beam on the  $\frac{1}{4}$ -in. face and the other axial on the end. The first drive was

designed to generate bending waves and the second longitudinal waves in leg A. Measurements were made over a frequency range from 0 to 1 kHz. Similar measurements were made on the straight beam, which represented legs A and B (i.e., the base), without leg C (i.e., the stem), of the T-beam.

### B. Sources of experimental error

Three sources of experimental error are estimated, one random and two bias. Assuming that individual estimates of the autospectra have a random Gaussian distribution, then the random error in the autospectra computed from an average of ensemble estimates is one over the square root of the number of ensembles.<sup>10</sup> Therefore, the expected error for the 300 ensemble averages used in the measurements reported here is within 7%.

Two sources of bias error are considered. First, the error associated with phase error at different measurement locations, and second, the contamination of the measurement by transverse sensitivity of the accelerometer are considered.

Since a single accelerometer was used in all of the measurements, the only potential source of phase error is differ-

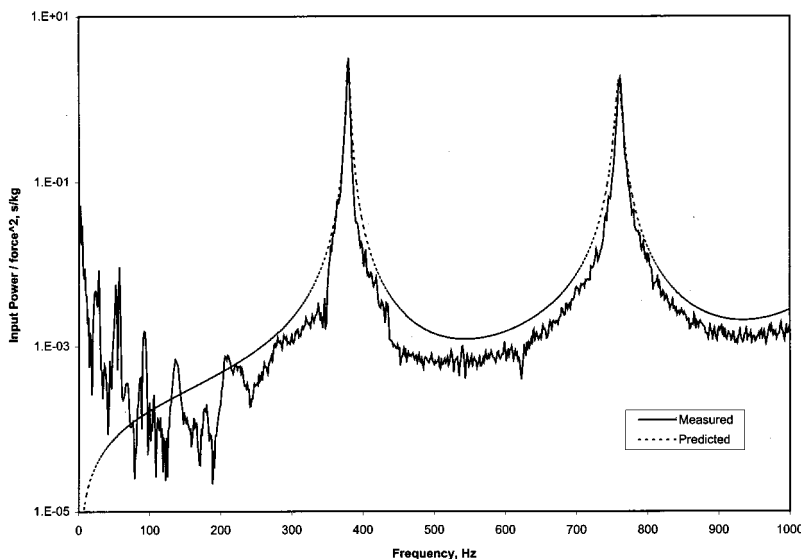


FIG. 3. Predicted and measured longitudinal wave power input to axially driven straight beam.

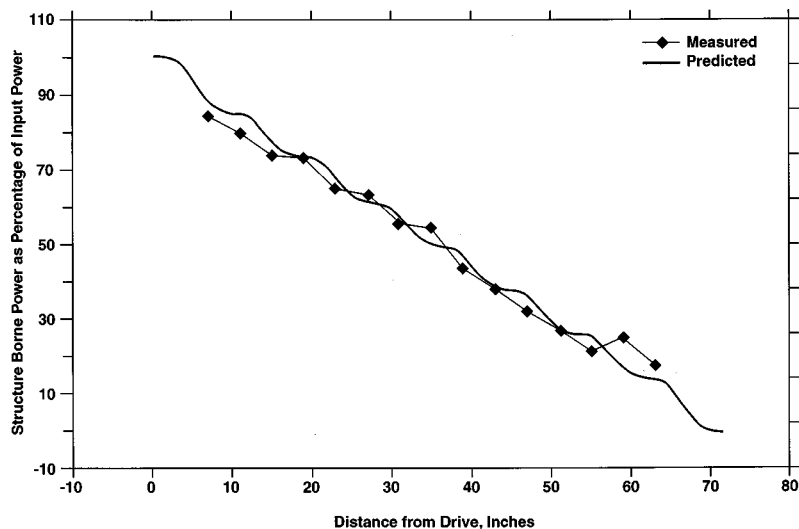


FIG. 4. Measured and predicted bending wave power in straight beam at 420 Hz.

ences in phase introduced by the relocation of the accelerometer. To assess the effect of phase changes due to relocation, the accelerometer was repeatedly installed on a mass attached to a shaker and the variation in the phase between signals from the accelerometer and the force gage located between the shaker and mass measured. Using 300 ensemble averages in each measurement resulted in variations in the phase of approximately 0.25 degrees. Including this phase error into Eqs. (4) and (5) for the power resulted in estimated bias errors of less than 1%.

For longitudinal waves, the out-of-plane displacement is related to the in-plane displacement by the factor  $-jk_l\mu h/2$ , where  $\mu$  is Poisson's ratio and  $k_l$  is the longitudinal wave number.<sup>8</sup> For bending waves, the in-plane displacement is related to the out-of-plane displacement by the factor  $-jk_b h/2$ , where  $k_b$  is the bending wave number.<sup>8</sup> Combining these with an accelerometer transverse sensitivity,  $\alpha$ , the bias errors for longitudinal and bending waves, associated with contamination by out-of-plane displacement or in-plane displacement, respectively, and transverse sensitivity of the accelerometer, are  $[\alpha\mu k_l h/2]^2$  for longitudinal waves and  $[\alpha k_b h/2]^2$  for bending waves. With a transverse sensitivity

of less than 5%, this results in bias errors of less than 1% for the measurements reported here.

### III. PREDICTIONS

Power flow in the beams was predicted using McPOW (Mechanical POWER), a postprocessor for the finite element program NASTRAN, developed by Hambric<sup>11</sup> and validated by Hambric and Taylor.<sup>12</sup> The beams were modeled using 1-in. elements which gives a mesh density sufficient to resolve longitudinal and bending waves for frequencies up to 1 kHz. The masses of the shaker, impedance head, and stinger were included at the driven end of the model. Since the excitations and responses of the cross sections of the beams are planar for all cases considered in this paper, powers are computed only for bending and longitudinal waves

$$P = \text{Re}\{F_x v_x^* + V_y v_y^* + M_z \dot{\theta}_z^*\}, \quad (7)$$

where the cross section of the beams are in the  $yz$  rectangular coordinate plane. In Eq. (7),  $P$  is power,  $F_x$  is the axial force,  $v_x$  is the axial velocity,  $V_y$  is the transverse shear force,  $v_y$  is the transverse velocity,  $M_x$  is the bending moment, and  $\dot{\theta}_z$  is

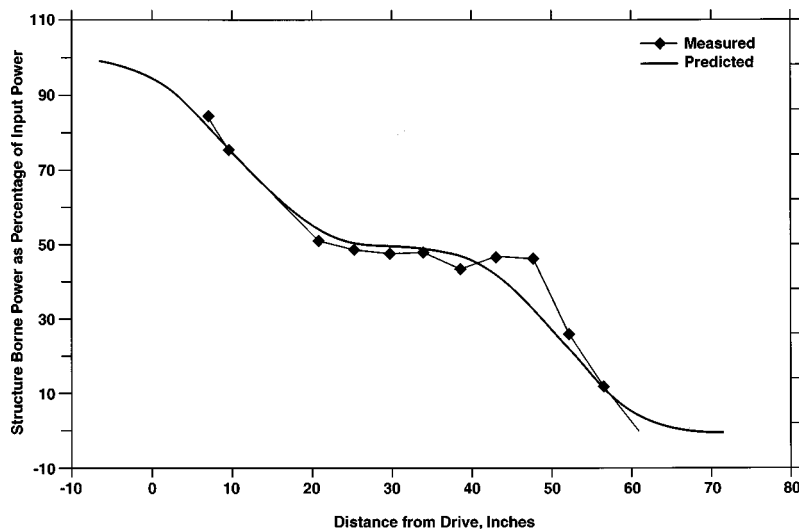


FIG. 5. Measured and predicted longitudinal wave power in straight beam at 780 Hz.



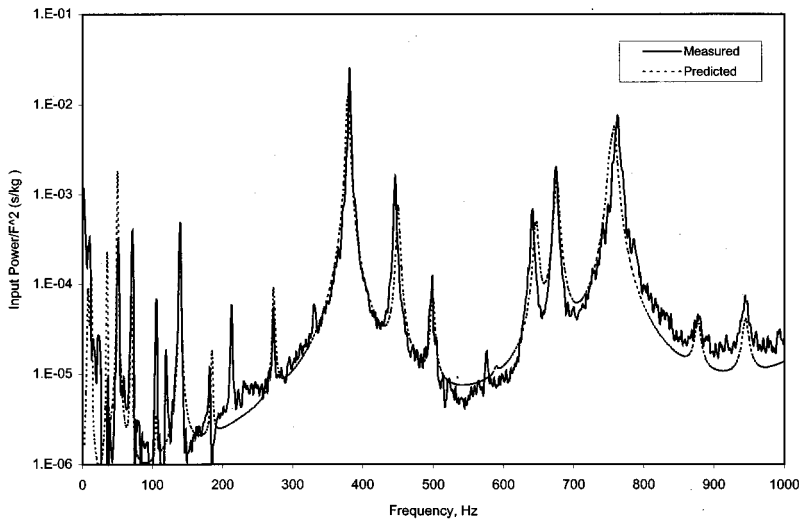


FIG. 6. Measured and predicted power input for axially driven T-beam.

the rotational velocity about the main axis of the beam. The term  $F_x v_x^*$  is the power in the longitudinal wave and  $V_y v_y^* + M_z \theta_z^*$  is the power in the bending wave.

Input powers were computed by McPOW as

$$P_i = \text{Re}\{F_i v_i^*\}, \quad (8)$$

where  $F_i$  is the drive force and  $v_i$  is the velocity in the direction of the applied force at the drive location.

#### IV. RESULTS

Results are presented for the straight beam first to demonstrate the validity of the measurements on a simple structure, where there is no coupling between bending and longitudinal waves, before presenting results for the T-beams, where coupling occurs at the junction.

##### A. Straight beam

Measured and predicted input powers to the straight beam are compared in Figs. 2 and 3, for transverse and axial drives, respectively. For both drives, the peak levels will

depend on the damping. Damping loss factors were measured to be approximately 0.01 across the frequency range using the half-power bandwidth method.

In Fig. 4, the measured and predicted axial distribution of structure-borne power at 420 Hz, a bending wave resonance, are plotted for the transverse drive. The powers are normalized to the input power to compare experimental and predicted results. Near the drive, the measured bending wave power is approximately 100% of the input power, and decreases to zero at the end of the beam. This agrees with expectations and provides a check on the validity of the structure-borne power measurements. The measured power in the longitudinal wave is 80 dB below the power in the bending wave near the drive, indicating that the bending wave dominates the response of the straight beam for the transverse drive at 420 Hz, as expected since there is no mechanism for coupling between waves. Agreement between measurements and predictions is good, which indicates the validity of both. However, the predictions show a modulation in power not seen in the measurements. Deflections in the standing wave in the beam are maximum at antinodes, where the power dissipation is greatest and minimum at

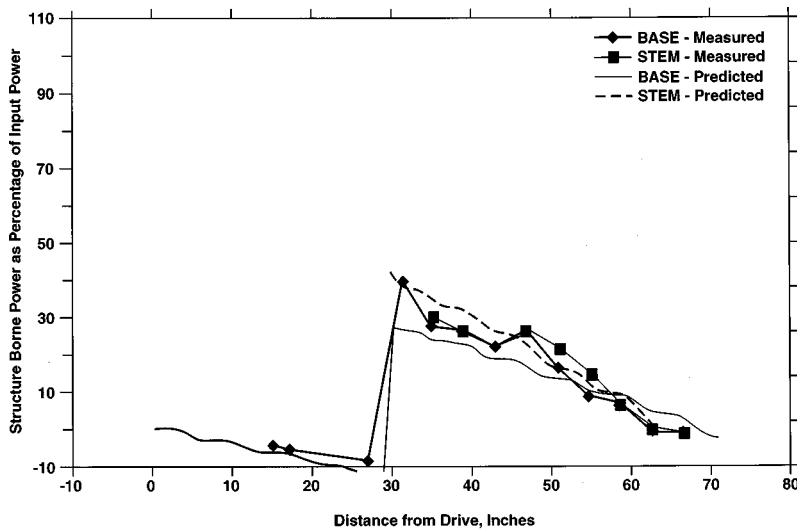


FIG. 7. Measured and predicted bending wave power in axially driven T-beam at 670 Hz.

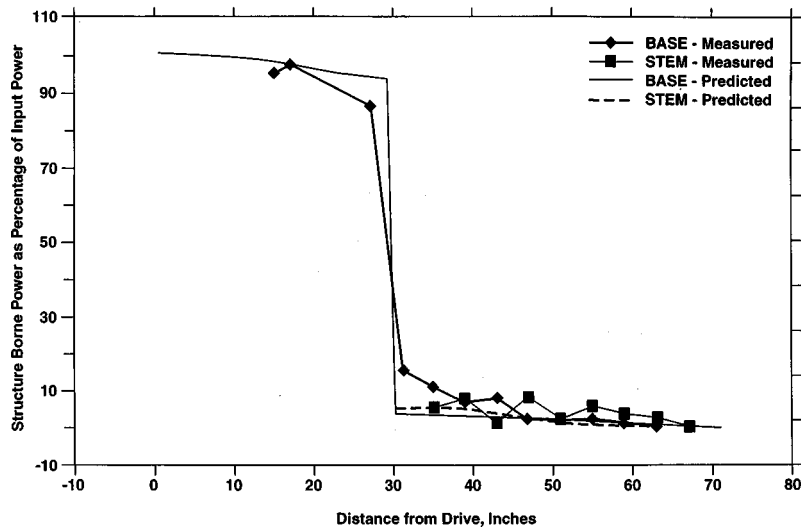


FIG. 8. Measured and predicted longitudinal wave power in axially driven T-beam at 670 Hz.

nodes where power dissipation is minimum. This results in the modulation shown in the predictions in Fig. 4, where the modulation has a period equal to one-half the bending wavelength. The amplitude of the modulation in power is too small and the measurement points too far apart for the measurements to follow the modulation in the predictions.

With an axial drive at the longitudinal resonance of 780 Hz, measured and predicted power flows are compared in Fig. 5. With the exception of the points around 50 in. from the drive, the comparison between the measured and predicted longitudinal power is good. It is likely that these two bad points are the result of a single bad measurement, since the power at one point requires measurements at two points. Consequently, each measurement point contributes to measurements at two neighboring points. The measured power in the bending wave is greater than 80 dB below the measured power in the longitudinal wave near the drive point, indicating that the longitudinal waves were dominant in the beam. The modulation in the predicted power is also seen in the measurements. Unlike for the bending waves at 420 Hz, the distance between measurements points is small enough, relative to the longitudinal wavelength, and the amplitude of the

modulation large enough for the measurements to follow the modulation.

### B. T-beam

The measured input power for the axial drive at the end of leg A of the T-beam is compared to predictions in Fig. 6. The addition of the cross leg in the T-beam introduces several resonances seen in both the measurements and predictions, which show good agreement.

Measurements and predictions are shown in Figs. 7–10 at two frequencies, 670 and 735 Hz. These two frequencies are selected because they show different power flow paths. The stem contains an axial resonance at 735 Hz but not at 670 Hz. Consequently, a greater percentage of longitudinal wave power flows into the stem at this frequency, leaving less power available to flow into leg B. At 670 Hz, because less longitudinal power flows into the stem, the total power levels in leg B are higher.

In Figs. 7 and 8, the measured and predicted powers in bending and longitudinal waves, respectively, are plotted for the base (legs A and B) and the stem (leg C) of the T-beam

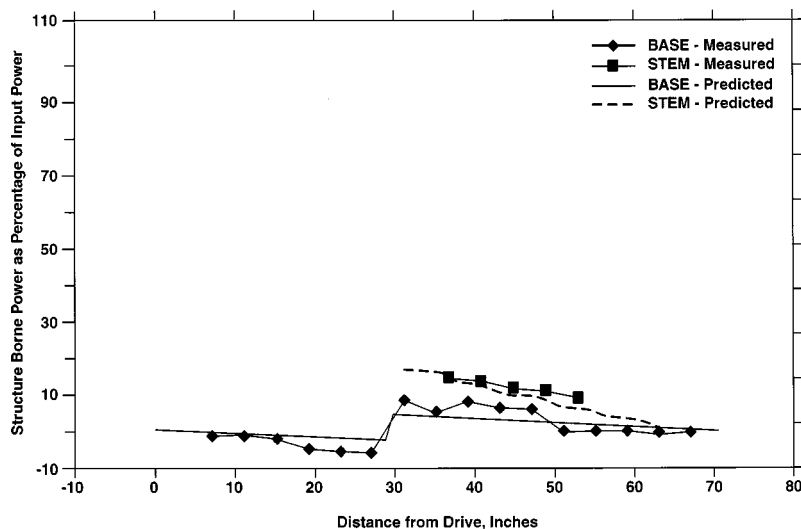


FIG. 9. Measured and predicted bending wave power in axially driven T-beam at 735 Hz.

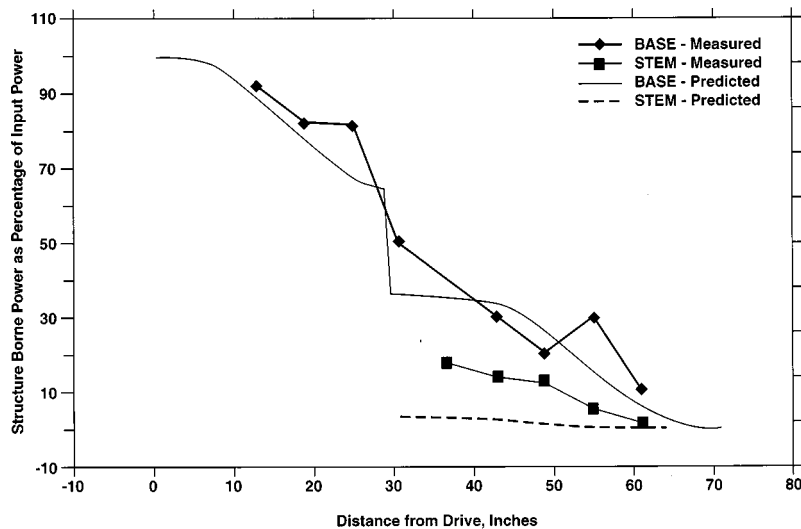


FIG. 10. Measured and predicted longitudinal wave power in axially driven T-beam at 735 Hz.

at 670 Hz, when the force is applied axially to the end of leg A. The agreement between the predictions and measurements is good for both bending and longitudinal waves. The bending wave powers shown in leg A are small, and slightly negative, indicating power flow toward the drive point along the base that originates at the intersection with the stem. However, the bending wave power in the stem is much higher. At the base of the stem, the bending wave power is 40% of the input power. Figure 8 shows that the measured longitudinal wave power in leg A is approximately 85% of the input power near the junction with the stem and then drops to 15% of the input power after passing the junction. This drop in power can be explained by the power required to generate the bending wave power in the stem. Therefore, it appears that the bending wave power in the stem is caused by longitudinal waves coupled to bending waves at the junction. The negative bending wave power in leg A is also produced by longitudinal waves at the junction coupling to bending waves which propagate back to the source.

Figures 9 and 10 show predicted and measured powers at 735 Hz for the axial drive on the T-beam in bending and longitudinal waves, respectively. Agreement between predictions and measurements is good for bending waves. However, agreement for longitudinal waves at 735 Hz is not as

good as for bending waves, or for longitudinal waves at 670 Hz. The longitudinal wave in leg A is the source of the bending waves at the junction, so that all of the longitudinal wave is excited in the stem at 735 Hz than at 670 Hz. Because of the longitudinal resonance at 735 Hz, bending waves generated by longitudinal waves incident on the junction are all much weaker at 735 Hz than observed at 670 Hz. At 735 Hz, the bending wave powers in legs B and C are less than 5% of the drive power.

Measured and predicted input powers for the transverse drive on the T-beam are compared in Fig. 11. Agreement between the predictions and measurements is good around the peaks, but the measurements are lower than predictions between peaks. This was also true for both types of drives on the straight beam and for the axial drive on the T-beam.

At 670 Hz, predictions and measurements of the powers in bending and longitudinal waves are presented in Figs. 12 and 13, respectively, for the transverse drive on the T-beam. Again, good agreement between predictions and measurements can be observed. Of the bending wave power incident on the junction, approximately 30% is transmitted through the junction to leg B and 30% to the stem. Longitudinal waves are generated at the junction and, as shown in Fig. 13,

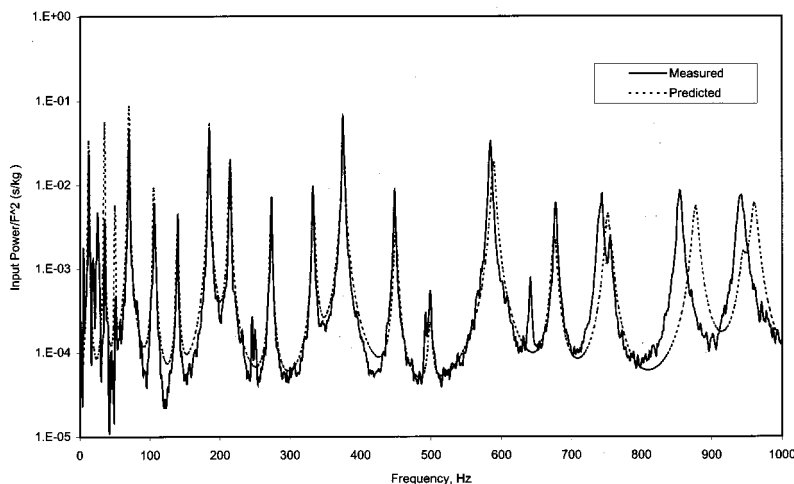


FIG. 11. Measured and predicted power input for transversely driven T-beam.

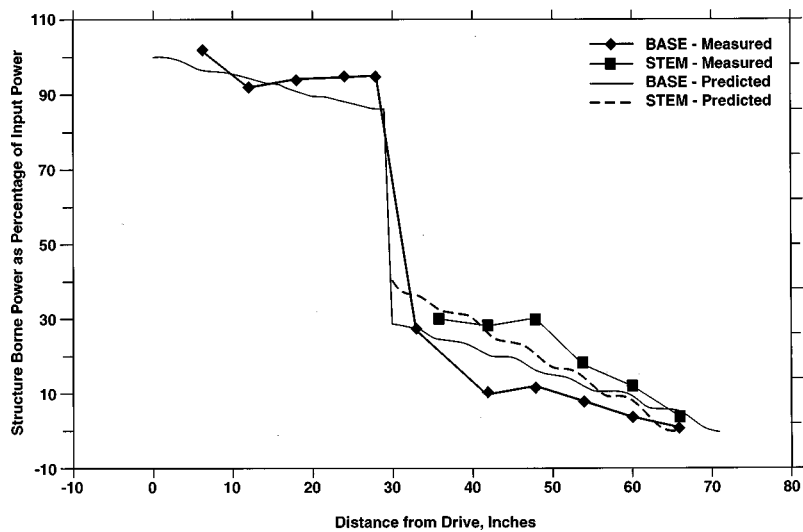


FIG. 12. Measured and predicted bending wave power in transversely driven T-beam at 670 Hz.

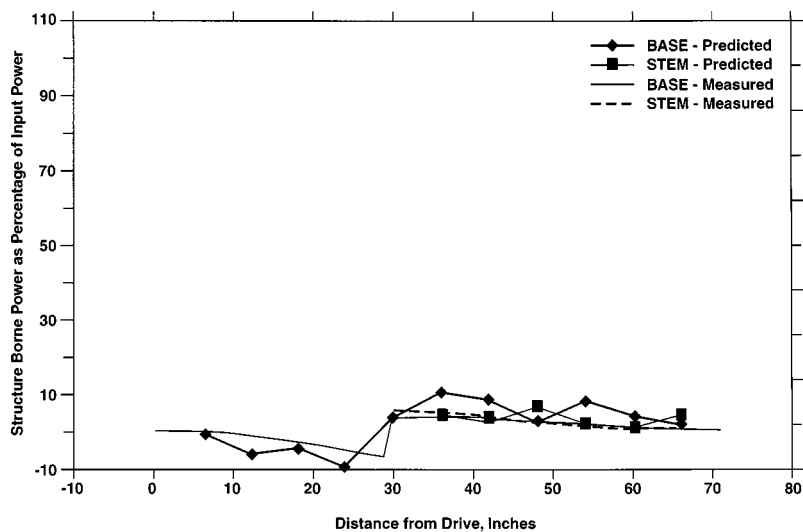


FIG. 13. Measured and predicted longitudinal wave power in transversely driven T-beam at 670 Hz.

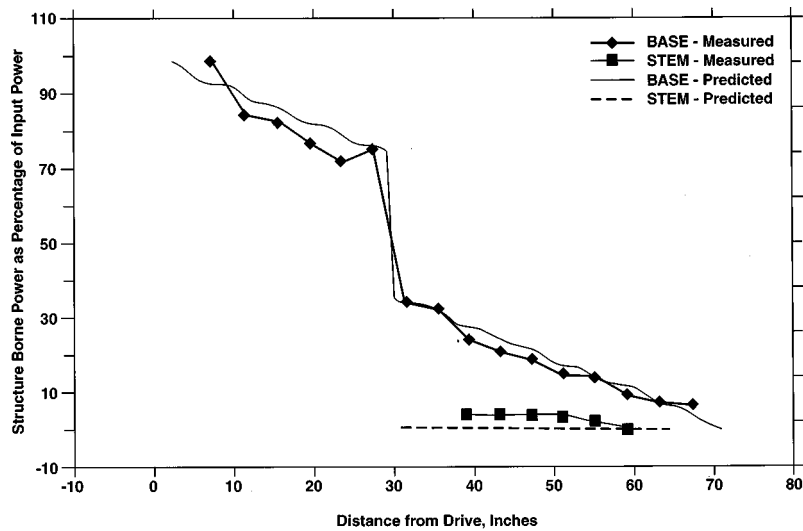


FIG. 14. Measured and predicted bending wave power in transversely driven T-beam at 735 Hz.

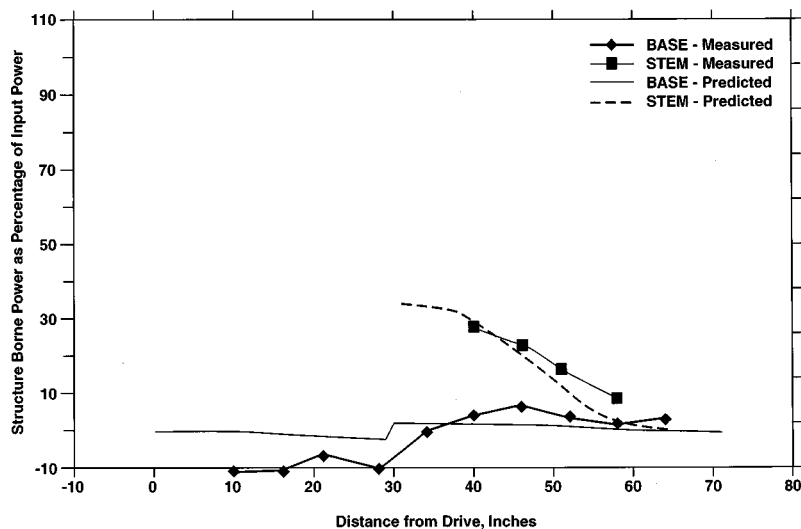


FIG. 15. Measured and predicted longitudinal wave power in transversely driven T-beam at 735 Hz.

power in these waves is transmitted away from the junction. Comparing Figs. 9 and 13, it can be seen that coupling of bending to longitudinal waves is weaker than the coupling from longitudinal to bending waves. Predicted and measured powers show good agreement in Figs. 14 and 15 for the transverse drive at 735 Hz. At 735 Hz, a strong longitudinal wave resonance exists in the stem with no significant bending resonances. As shown in Fig. 15, 35% of the bending wave power incident at the junction is transmitted into the stem as power in a longitudinal wave, with little power in the bending waves, consistent with the resonances. All of the power in the longitudinal waves propagates away from the junctions.

Finally, the bending wave and longitudinal wave powers are experimentally determined as a function of frequency at a position on the stem at 10 in. from the joint. These powers, compared with finite element predictions, are shown in Figs. 16 and 17, respectively.

## V. CONCLUSIONS

Agreement between the predictions and measurements shown in this paper is good, indicating that the measure-

ments are reasonably accurate and the experimental method for the measurement of power in bending and longitudinal waves is valid. Powers in bending and longitudinal waves can be separated when both waves are present.

Measurements and predictions both show coupling between bending and longitudinal waves that have been predicted in a number of studies.<sup>13-20</sup> Also, it was shown that the junction becomes a source of new waves that, without knowledge of power flow in both wave types, could be interpreted as the original source. Measurements that include the power in both wave types show power flow from the original source and redistribution of power by coupling of the waves at the junctions.

The method presented in this paper requires measurements on both sides of the structure. Both sides of a structure are usually not available for measurement. Also, measurements presented here were conducted with accelerometers, which make a high density of measurement points difficult. A laser vibrometer may be used for both transverse and in-plane measurements as demonstrated by McDevitt *et al.*;<sup>5,6</sup> however, it may not be possible to scan a laser beam on both sides of most structures. Therefore, work presented here is

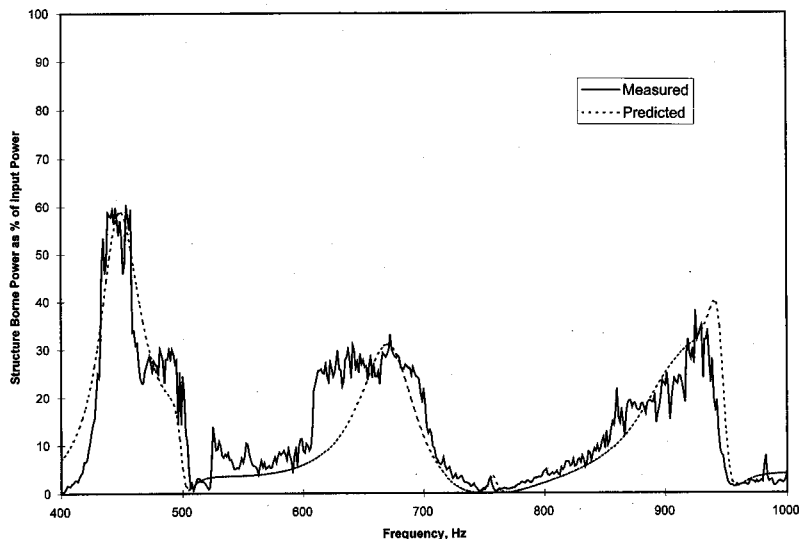


FIG. 16. Predicted and measured bending wave power on stem of T-beam.

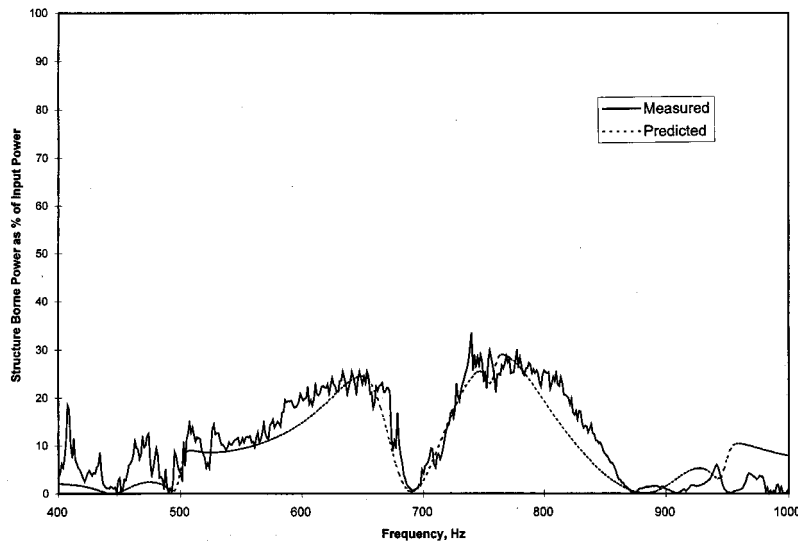


FIG. 17. Predicted and measured longitudinal wave power on stem of T-beam.

currently being extended to one-sided measurements made with a scanning laser vibrometer.

### ACKNOWLEDGMENTS

The support of the Naval Surface Warfare Center, and Dr. Peter Majumdar and Dr. Richard Vogelsong of the Office of Naval Research are recognized and appreciated.

<sup>1</sup>D. Noiseaux, "Measurement of power flow in uniform beams and plates," *J. Acoust. Soc. Am.* **47**, 238–247 (1970).  
<sup>2</sup>G. Pavic, "Measurement of Structure Borne Wave Intensity, Part I: Formulation of the Methods," *J. Sound Vib.* **49**(2), 221–230 (1976).  
<sup>3</sup>J. W. Verheij, "Measurements of Structure-Borne Wave Intensity of Lightly Damped Pipes," *Noise Control Eng. J.* **35**(2), 69–76 (1990).  
<sup>4</sup>J. Linjama and T. Lahti, "Estimation of Bending Wave Intensity in Beams Using the Frequency Response Technique," *J. Sound Vib.* **153**(1), 21–36 (1992).  
<sup>5</sup>T. E. McDevitt, G. H. Koopmann, and C. B. Burroughs, "Two Channel Laser Vibrometer Techniques for Vibrational Intensity Measurements, Part 1: Flexural Intensity," *J. Acoust. Vib.* **115**, 436–440 (1993).  
<sup>6</sup>T. E. McDevitt, G. H. Koopmann, and C. B. Burroughs, "Two Channel Laser Vibrometer Techniques for Vibrational Intensity Measurements, Part 2: Longitudinal Intensity," *J. Acoust. Vib.* **116**, 100–104 (1994).  
<sup>7</sup>R. S. Ming and R. J. Craik, "Errors in the Measurement of Structure-Borne Power Flow using Two-Accelerometer Techniques," *J. Sound Vib.* **204**(1), 59–71 (1997).  
<sup>8</sup>L. Cremer, M. Heckl, and E. Ungar, *Structure-Borne Sound* (Springer-Verlag, New York, 1973).  
<sup>9</sup>R. P. Szwerc and S. A. Hambric, "The Calculation and Measurement of

Flexural and Longitudinal Structural Power Flow on a Tee Shaped Beam," NSWCCD-SIG-96/014-7250, February 1996.  
<sup>10</sup>J. S. Bendat and A. G. Piersol, *Random Data; Analysis and Measurement Procedures* (Wiley, New York, 1986).  
<sup>11</sup>S. A. Hambric, "Power Flow and Mechanical Intensity Calculations in Structural Finite Element Analysis," *Trans. ASME, J. Vib. Acoust.* **112**, 542–549 (1990).  
<sup>12</sup>S. A. Hambric and P. D. Taylor, "Comparison of Experimental and Finite Element Structure-Borne Flexural Power Measurements for a Straight Beam," *J. Sound Vib.* **170**(5), 595–605 (1994).  
<sup>13</sup>S. V. Budrin and A. S. Nikiforov, "Wave Transmission Through Assorted Plate Joints," *Sov. Phys. Acoust.* **9**(4), 333–336 (1964).  
<sup>14</sup>T. Kihlman, "Sound Transmission in Building Structures of Concrete," *J. Sound Vib.* **11**(4), 435–445 (1970).  
<sup>15</sup>G. Rosenhouse, H. Ertel, and F. P. Mechel, "Theoretical and experimental investigations of structure-borne sound transmission through a 'T' joint in a finite system," *J. Acoust. Soc. Am.* **70**(2), 492–499 (1981).  
<sup>16</sup>M. D. McCollum and J. M. Cuschieri, "Bending and in-plane wave transmission in thick connected plates using statistical energy analysis," *J. Acoust. Soc. Am.* **88**(3), 1480–1485 (1990).  
<sup>17</sup>J. A. Moore, "Vibration transmission through frames of beam junctions," *J. Acoust. Soc. Am.* **88**(6), 2766–2766 (1990).  
<sup>18</sup>J. L. Horner and R. G. White, "Prediction of Vibrational Power Transmission Through Bends and Joints in Beam-Like Structures," *J. Sound Vib.* **147**(1), 87–103 (1991).  
<sup>19</sup>I-T. Lu, H. L. Bertoni, and H-Y. Chen, "Coupling of plate waves at joints," *J. Acoust. Soc. Am.* **92**(1), 510–526 (1992).  
<sup>20</sup>R. C. N. Leung and R. J. Pinnington, "Wave Propagation Through Right-Angled Joints with Compliance: Longitudinal Wave," *J. Sound Vib.* **153**(2), 223–237 (1992).

# Influence of circumferential partial coating on the acoustic radiation from a fluid-loaded shell

J. M. Cuschieri

*Center for Acoustics and Vibration, Department of Ocean Engineering, Florida Atlantic University, Boca Raton, Florida 33431*

D. Feit

*Carderock Division, NSWC, David Taylor Research Center, Bethesda, Maryland 20817-5000*

(Received 1 October 1998; revised 8 October 1999; accepted 9 February 2000)

Acoustic compliant coatings are a common approach to mitigate the radiation and scattering of sound from fluid-loaded submerged structures. An acoustic compliant coating is a coating that decouples an acoustic source from the surrounding acoustic medium; that is, it provides an acoustic impedance mismatch (different density and speed of sound product). Such a coating is distinct from an ordinary compliant coating in that it may not be resilient in the sense of low stiffness, but still provides an acoustic impedance mismatch. Ideally, the acoustic coating is applied uniformly over the entire surface of the fluid-loaded structure to minimize the acoustic radiation and scattering. However, in certain instances, because of appendages, it may not be practically possible to completely cover the surface of a fluid-loaded structure to decouple it from the adjacent acoustic medium. Furthermore, there may be some inherent advantages to optimizing the distribution of the coating around areas from which the acoustic radiation appears to be dominant. This would be analogous to the application of damping treatment to a vibrating structure in areas where the vibration levels are highest. In the case of the acoustic radiation the problem is more complex because of the coupling between the acoustic fluid and the structure. In this paper, the influence of a partial coating on the acoustic radiation from a fluid-loaded, cylindrical shell of infinite extent and excited by either a line force or an incident plane acoustic wave is examined. The solution to the response and scattered pressure is developed following the procedure used by the authors in previous work on the scattering from fluid-loaded plates and shells. The coating is assumed to be normally reacting providing a decoupling layer between the acoustic medium and the structure; that is, it does not add mass or stiffness to the base structure. The influence of added mass or stiffness of the coating can be included as an added inhomogeneity and treated separately in the solution. © 2000 Acoustical Society of America. [S0001-4966(00)02305-5]

PACS numbers: 43.40.Fz, 43.40.Rj [CBB]

## INTRODUCTION

The acoustic scattering and radiation from fluid-loaded submerged structures can be controlled through the use of compliant coatings. However, real submerged structures can be very complex and generally it may be practically impossible to apply a uniform acoustic compliant coating over the whole wet surface of the submerged structure. The structure may have appendages that would preclude the applications of a uniform coating. Alternatively, it may be of some advantage to optimize the distribution of the coating to either modify the scattering from a particular mode or group of modes, or to reduce the added weight from the coating on the submersible, which may translate into either increased speed or reduced required propulsion power and hence less noise. Since bulkheads and discontinuities in the fluid-loaded structure are the main source of acoustic scattering, the application of the coating near the bulkheads can be optimized to achieve maximum reduction in the acoustic scattering and radiation without exceeding a limitation imposed on the weight that can be added to the structure by the coating. Furthermore, it may be possible to strategically apply partial

acoustic coatings to achieve a desired acoustic scattering and radiation behavior from the submerged structure.

The optimization of the application of the coating would be analogous to the application of damping treatment in areas of high vibration amplitude. However, because of the coupling between the acoustic fluid and the structure, the acoustic scattering and radiation are more closely related to the whole structural surface motion. That is, the acoustic radiation and scattering are dependent on the whole of the vibrating surface, making the problem more complex. Because of this complex coupling between the whole structure surface and the acoustic radiation and scattering, it is of interest to model the acoustic radiation and scattering from a fluid-loaded shell structure with partial circumferential acoustic compliant coating. In this paper the problem of acoustic radiation and scattering from a fluid-loaded cylindrical shell with acoustic coating strips that do not extend all around the circumference is considered.

Partial coatings influence the acoustic scattering and radiation in at least two ways. First, the energy associated with supersonic waves contained within the exposed (uncoated) areas of the structure can find their way into acoustic radia-

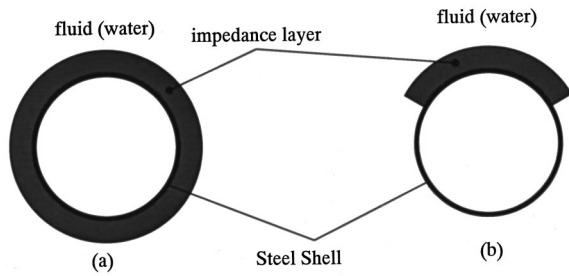


FIG. 1. Fully coated (a) and partially coated (b) shells. Angular spatial distribution of the coating will be considered through a circumferential mode decomposition.

tion. Second, the boundaries of the coating will create discontinuities in the submerged structure impedance, and these discontinuities can scatter some of the subsonic vibration energy contained in the structure into supersonic wave components, increasing the acoustic radiation and scattering. In the case of partial coatings, the edge “smoothness” of the coating and the thickness shaping may play a significant role in defining the acoustic target strength of the submersible. Observations based on the formulation developed and presented in later sections of this paper and earlier findings of the authors that were reported in Ref. 1 both show that this is expected. Thus an understanding of the physics controlling acoustic scattering and radiation from partially coated, fluid-loaded structures is important. A comparison of the changes in the target strength between the uncoated, fully coated, and partially coated submerged cylindrical shell can provide some of the required understanding.

Maidanik and Tucker<sup>2</sup> and Maidanik *et al.*<sup>3</sup> considered the problem of coated panels immersed in a fluid medium. The formulation presented in their paper is applicable to fully coated panels under different types of excitation. The coating is modeled as a damped spring with a surface impedance which is a fraction of the characteristic acoustic impedance of the surrounding fluid. The results obtained show the onset of resonances of the compliant layer, and that the radiation may actually increase in the vicinity of the first resonance formed by the plate mass moving against the resilience of the coating. In these papers, the coating is assumed to be uniformly distributed on the surface of the panel. Ferri *et al.*<sup>4</sup> considered the problem of scattering from partially coated cylinders. However, in their paper, the coated and uncoated sections of the cylinder are modeled as constant impedance surfaces; that is, the elasticity of the shell is not included in the formulation, and thereby there are no traveling waves within the combined shell coating structure. Laulagnet and Guyader<sup>5</sup> considered the radiation from partially coated finite shells and developed results for the total radiated acoustic power and the surface average velocity response. In the paper by Laulagnet and Guyader<sup>5</sup> the pressure and velocity distributions around the shell circumference are not presented. Furthermore, the results are for only the radiation case.

The objective of the present paper is to model and hence determine the influence of partial angular coating around the shell circumference, on the acoustic scattering, and radiation from a fluid-loaded cylindrical shell excited by either an in-

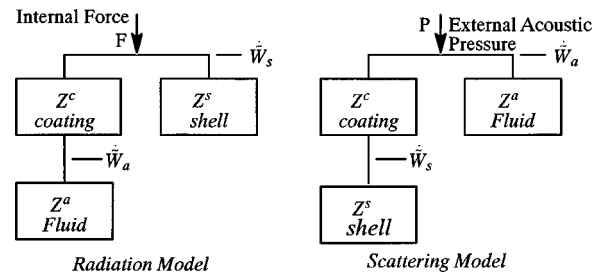


FIG. 2. Impedance models for the radiation and scattering problems.

cident plane acoustic wave (for acoustic scattering) or by a concentrated axial load (for acoustic radiation). The solution to this problem is based on the approach developed by the authors in their previous work on scattering from fluid-loaded cylindrical shells with different internal structures.<sup>6,7</sup> A partial circumferential coating would influence the different circumferential modes of the shell in different ways. The influence would be a function of the angular spread of the coating. Thus in determining the dependency of the acoustic scattering and radiation from the shell due to the angular spread and orientation of the coating, all the contributions from the significant circumferential modes have to be taken into account. The significant circumferential modes can be determined from a circumferential decomposition of the coating angular distribution.

## GENERAL FORMULATION

The model of the fluid-loaded cylindrical shell is based on the classical (Donnell–Mushtari)<sup>8</sup> cylindrical shell theory modified to include the impedance of the acoustic compliant coating. The coatings is introduced by inserting a mass/stiffness impedance in series/parallel to the surface of the shell. The impedance layer decouples the fluid from the structure. This modeling approach is similar to that used by Maidanik and Tucker<sup>2</sup> and Maidanik *et al.*<sup>3</sup> for a fluid-loaded coated plate. The angular distribution of the coating is included through a modal decomposition of the distribution function in the circumferential direction. The basic problem to be considered is shown in Fig. 1.

Two different impedance models are required to model the radiation and scattering problem; that is, the impedance model changes whether the shell is excited by an internal mechanical load (radiation problem) or externally excited by an incident acoustic wave (scattering problem). The two impedance models are shown in Fig. 2. In these impedance models, the impedance is the modal, wavenumber transform of the spatial impedance.

In Fig. 2,  $\dot{W}_s$  and  $\dot{W}_a$  are, respectively, the radial velocity at the shell and acoustic fluid interfaces, and  $Z^c$ ,  $Z^s$ , and  $Z^a$  are, the spatial transformed impedance of the coating, impedance of the shell, and impedance of the fluid, respectively. For the radiation condition,

$$\dot{W}_a = \frac{F}{\left[ Z^a + Z^s + \frac{Z^s Z^a}{Z^c} \right]} \quad \text{and} \quad \dot{W}_s = \frac{F}{\left[ Z^s + \frac{Z^c Z^a}{Z^c + Z^a} \right]} \quad (1)$$

For the scattering condition,



$$\dot{W}_a = \frac{P_i}{\left[ Z^a + \frac{Z^c Z^s}{Z^c + Z^s} \right]} \quad \text{and} \quad \dot{W}_s = \frac{P_i}{\left[ Z^s + Z^a + \frac{Z^a Z^s}{Z^c} \right]}, \quad (2)$$

where  $F$  is the internal excitation force, and  $P_i$  is the incident pressure.

A simplified form of the coating impedance can be taken to be simply that of a decoupling layer; that is the coating has only radial stiffness.<sup>3</sup> Using the approach of Refs. 2 and 3, the coating impedance  $Z^c$  is written in the form:

$$Z^c = -j\rho_0 c_0 \xi(1 + j\eta); \quad (3)$$

the coating impedance is expressed as a fraction ( $\xi$ ) of the impedance of the acoustic fluid  $\rho_0 c_0$ , where  $\rho_0$  and  $c_0$  are the density and acoustic wave speed of the external fluid medium, together with a damping term ( $\eta$ ).

### CYLINDRICAL SHELL FORMULATION

The Donnell–Mushtari<sup>7</sup> shell equations are used to model the behavior of the fluid-loaded shell. Thus starting from the equations of motion<sup>7</sup> given by

$$a^2 \frac{\partial^2 u(x, \phi)}{\partial x^2} + \left( \frac{1-v}{2} \right) \frac{\partial^2 u(x, \phi)}{\partial \phi^2} + \left( \frac{1+v}{2} \right) a \frac{\partial^2 v(x, \phi)}{\partial \phi \partial x} + v a \frac{\partial w^s(x, \phi)}{\partial x} + \Omega^2 u(x, \phi) = 0, \quad (4)$$

$$\left( \frac{1+v}{2} \right) a \frac{\partial^2 u(x, \phi)}{\partial \phi \partial x} + a^2 \left( \frac{1-v}{2} \right) \frac{\partial^2 v(x, \phi)}{\partial x^2} + (1+\beta^2) \frac{\partial^2 v(x, \phi)}{\partial \phi^2} + \left( \frac{\partial w^s(x, \phi)}{\partial \phi} - a^2 \beta^2 \frac{\partial^3 w^s(x, \phi)}{\partial x^2 \partial \phi} - \beta^2 \frac{\partial^3 w^s(x, \phi)}{\partial \phi^3} \right) + \Omega^2 v(x, \phi) = 0, \quad (5)$$

$$-v a \frac{\partial u(x, \phi)}{\partial x} - \left( \frac{\partial v(x, \phi)}{\partial \phi} - a^2 \beta^2 \frac{\partial^3 v(x, \phi)}{\partial x^2 \partial \phi} - \beta^2 \frac{\partial^3 v(x, \phi)}{\partial \phi^3} \right) - \left[ w^s(x, \phi) + \beta^2 \left( a^2 \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial \phi^2} \right) w^s(x, \phi) \right] + \Omega^2 w^s(x, \phi) = \gamma [F_w \delta(\phi) - p_s(x, \phi) - p_i(x, \phi)], \quad (6)$$

where  $a$  and  $h$  are, respectively, the radius and wall thickness of the shell,  $E$ ,  $\rho$ , and  $\nu$  are, respectively, the shell material modulus of elasticity, density and Poisson's ratio,  $\beta = h/(a\sqrt{12})$ ,  $\gamma = a^2(1-\nu^2)/(Eh)$ ,  $F_w$  is the amplitude of the radial internal axial line force,  $\Omega$  is the frequency parameter representing the frequency normalized with respect to the ring frequency,  $\Omega = \omega/\omega_r$ ,  $\omega_r = c_p/a$ ,  $c_p = \sqrt{E/[\rho(1-\nu^2)]}$ ,  $p_s$  is the scattered acoustic pressure, and  $p_i$  is the incident excitation pressure.

The cylindrical shell is excited by either a line force located at an angular location  $\phi=0$  degrees and extending along the axis of the shell, or by an incident plane acoustic wave from an incidence angle of  $\phi=180$  degrees (Fig. 3).

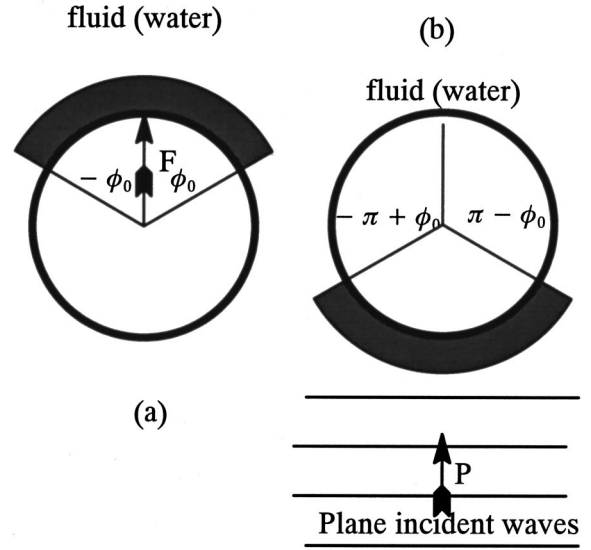


FIG. 3. Excitation and partial coating configurations.

The change in orientation of the coating between the internal line force driven and the acoustic scattering is solely due to the formulation of the expressions. From the formulation of the equations of motion in Ref. 8 radial forces are applied at  $\phi=0$  degrees, while scattering is considered from waves incident on the shell from  $\phi=180$  degrees.

With these forms of excitation the response of the infinitely long cylindrical shell has no axial dependency; that is, both forms of excitation are independent of the axial variable  $x$ . In this case the response of the shell is also independent of the variable  $x$ , thus the above equations of motion simplify to

$$(1 + \beta^2) \frac{\partial^2 v(\phi)}{\partial \phi^2} + \left[ \frac{\partial}{\partial \phi} - \beta^2 \frac{\partial^3}{\partial \phi^3} \right] w^s(\phi) + \Omega^2 v(\phi) = 0, \quad (7)$$

$$- \left( \frac{\partial v(\phi)}{\partial \phi} - \beta^2 \frac{\partial^3 v(\phi)}{\partial \phi^3} \right) - \left[ w^s(\phi) + \beta^2 \frac{\partial^4 w^s(\phi)}{\partial \phi^4} \right] + \Omega^2 w^s(\phi) = \gamma [F_w \delta(\phi) - p_s(\phi) - p_i(\phi)]. \quad (8)$$

The radial displacement  $w^s(\phi)$  in the above equations is the displacement response of the shell surface. The axial independency decouples the axial displacement  $u(\phi)$  from the circumferential  $v(\phi)$  and radial displacement  $w^s(\phi)$  and for this reason the expression for  $u(\phi)$  is not included. Using a modal decomposition in the circumferential direction,

$$v(\phi) = \sum_n V_n \sin(n\phi), \quad (9)$$

$$w^s(\phi) = \sum_n W_n^s \cos(n\phi)$$

and for the scattered and incident pressures,

$$\gamma p_s(\phi) = \Omega^2 \left( \frac{a}{h} \right) \left( \frac{\rho_0}{\rho} \right) \sum_n \frac{H_n(\chi)}{\chi H_n'(\chi)} W_n^a \cos(n\phi), \quad (10)$$

$$\gamma p_i(\phi) = 4 \gamma P_i \sum_n \frac{\epsilon_n}{2\pi} \frac{j^{n+1}}{\chi H_n'(\chi)} \cos(n\phi), \quad (11)$$

where  $n$  is the circumferential mode number,  $j = \sqrt{-1}$ ,  $H_n(\chi)$  is a Hankel function of the first kind of order  $n$ ,  $H'_n(\chi)$  is the first derivative of the Hankel function with respect to the argument  $\chi$ ,  $\chi = \Omega M$ ,  $M = c_p/c_0$ ,  $\epsilon_n = 1$  for  $n = 0$  and  $\epsilon_n = 2$  for  $n > 0$ , and  $W_n^a$  is the modal displacement of the outside surface of the shell or the coating.  $W_n^a$  is related to  $w^a(\phi)$  through a modal decomposition and expressions for  $w^a(\phi)$  and  $W_n^a$  are given later in the analysis. Substituting Eqs. (12)–(14) in Eqs. (10) and (11) [since  $w(\phi)$  and  $\nu(\phi)$  are decoupled from  $u(\phi)$ , the latter axial displacement component is not considered further in the analysis],

$$\begin{aligned} & [(1 + \beta^2)n^2 - \Omega^2] \sum_n V_n \sin(n\phi) \\ & + [n(1 + n^2\beta^2)] \sum_n W_n^s \sin(n\phi) = 0 \end{aligned} \quad (12)$$

and

$$\begin{aligned} & [n(1 + n^2\beta^2)] \sum_n V_n \cos(n\phi) \\ & + [1 + \beta^2n^4 - \Omega^2] \sum_n W_n^s \cos(n\phi) \\ & = \gamma F_w \delta(\phi) - \Omega^2 \left( \frac{a}{h} \right) \left( \frac{\rho_0}{\rho} \right) \sum_n \frac{H_n(\chi)}{\chi H'_n(\chi)} W_n^a \cos(n\phi) \\ & - 4\gamma P_i \sum_n \frac{\epsilon_n}{2\pi} \frac{j^{n+1}}{\chi H'_n(\chi)} \cos(n\phi). \end{aligned} \quad (13)$$

Multiplying both sides of Eq. (15) by  $(\epsilon_n/2\pi)\sin(m\phi)$  and Eq. (16) by  $(\epsilon_n/2\pi)\cos(m\phi)$ , and integrating between  $-\pi$  and  $\pi$ , and then substituting for  $V_n$ , the following expression is obtained:

$$\begin{aligned} W_n^s \left[ (1 + \beta^2n^4 - \Omega^2) - \frac{n^2(1 + n^2\beta^2)^2}{(1 + \beta^2)n^2 - \Omega^2} \right] + F_n^L W_n^a \\ = \frac{\epsilon_n}{2\pi} \gamma F_n^w - 4\gamma P_i \frac{\epsilon_n}{2\pi} \frac{j^{n+1}}{H'_n(\chi)}, \end{aligned} \quad (14)$$

where  $F_n^L$  is the fluid-loading (related to the radiation impedance of the external acoustic medium), and is given by

$$F_n^L = \Omega^2 \left( \frac{a}{h} \right) \left( \frac{\rho_0}{\rho} \right) \frac{H_n(\chi)}{\chi H'_n(\chi)}. \quad (15)$$

In the radiation case [Fig. 3(a)] the displacement of the outside surface of the shell or the coating  $w^a(\phi)$ , satisfies the condition<sup>5</sup>

$$j\omega[w^a(\phi) - w^s(\phi)] = \begin{cases} -\frac{P_s(\phi)}{Z^c} & |\phi| < \phi_0, \\ 0 & \phi_0 < |\phi| < \pi \end{cases}, \quad (16)$$

where it is assumed that the coating extends over the circumferential section from  $-\phi_0$  to  $\phi_0$ . Thus from the definition of modal expansion,

$$W_n^a = \frac{\epsilon_n}{2\pi} \int_{-\pi}^{\pi} w^a(\phi) \cos(n\phi) d\phi, \quad (17)$$

applied to Eq. (16),

$$\begin{aligned} & \int_{-\phi_0}^{\phi_0} w^a(\phi) \cos(n\phi) d\phi - \int_{-\phi_0}^{\phi_0} w^s(\phi) \cos(n\phi) d\phi \\ & = -\frac{1}{j\omega Z^c} \int_{-\phi_0}^{\phi_0} \left[ \sum_m^{\infty} P_m^s \cos(m\phi) \right] \cos(n\phi) d\phi, \quad (18) \\ & \int_{-\pi}^{-\phi_0} w^a(\phi) \cos(n\phi) d\phi + \int_{\phi_0}^{\pi} w^a(\phi) \cos(n\phi) d\phi \\ & - \int_{-\pi}^{-\phi_0} w^s(\phi) \cos(n\phi) d\phi \\ & - \int_{\phi_0}^{\pi} w^s(\phi) \cos(n\phi) d\phi = 0. \end{aligned} \quad (19)$$

Combining Eqs. (18) and (19) by summation, the result can be simplified to

$$\frac{2\pi}{\epsilon_n} (W_n^a - W_n^s) = -\frac{1}{j\omega Z^c} \sum_m P_m^s A_{mn}, \quad (20)$$

where

$$\begin{aligned} A_{mn} &= \int_{-\phi_0}^{\phi_0} \cos(m\phi) \cos(n\phi) d\phi \\ &= \begin{cases} \left[ \frac{\sin[(n+m)\phi_0]}{(n+m)} + \frac{\sin[(n-m)\phi_0]}{(n-m)} \right] & n \neq m \\ \left[ \phi_0 + \frac{\sin(2n\phi_0)}{2n} \right] & n = m \\ [2\phi_0] & n = m = 0 \end{cases}. \end{aligned} \quad (21)$$

Substituting for  $P_m^s$  using Eq. (10),

$$[W_n^a - W_n^s] = -\frac{\epsilon_n}{2\pi} \frac{1}{j\omega \gamma Z^c} \sum_m F_m^L W_m^a A_{mn}. \quad (22)$$

From Eq. (14), leaving out the incident pressure term,

$$S_n W_n^s + F_n^L W_n^s + F_n^L [W_n^a - W_n^s] = \frac{\epsilon_n}{2\pi} \gamma F_n^w, \quad (23)$$

where

$$S_n = (1 + \beta^2n^4 - \Omega^2) - \frac{n^2(1 + n^2\beta^2)^2}{(1 + \beta^2)n^2 - \Omega^2} \quad (24)$$

substituting  $[W_n^a - W_n^s]$  from Eq. (20);

$$S_n W_n^s + F_n^L W_n^s - \frac{F_n^L}{Z^c} \frac{\epsilon_n}{2\pi} \sum_m F_m^L W_m^a A_{mn} = \frac{\epsilon_n}{2\pi} \gamma F_n^w, \quad (25)$$

where a normalized coating impedance is defined using the shell normalizing parameters, by including the definition of  $Z^c$  from Eq. (3) with the  $j\omega\gamma$  terms; that is,

$$\bar{Z}^c = j\omega \gamma Z^c = \left( \frac{a}{h} \right) \left( \frac{\rho_0}{\rho} \right) \frac{\Omega \xi(1 + j\eta)}{M}. \quad (26)$$

From Eq. (14)

$$W_m^a = \frac{(\epsilon_m/2\pi) \gamma F_m^w - S_m W_m^s}{F_m^L}. \quad (27)$$

Substituting Eq. (27) into Eq. (25) and rearranging:

$$\begin{aligned} S_n W_n^s + F_n^L W_n^s + \frac{F_n^L}{\bar{Z}^c} \frac{\epsilon_n}{2\pi} \sum_m A_{mn} W_m^s S_m \\ = \frac{\epsilon_n}{2\pi} \gamma F_n^w + \frac{F_n^L}{\bar{Z}^c} \frac{\epsilon_n}{2\pi} \sum_m A_{mn} \frac{\epsilon_m}{2\pi} \gamma F_m^w. \end{aligned} \quad (28)$$

If the coating extends over the whole of the circumference, that is,  $|\phi_0| = \pi$ , then Eq. (28) simplifies to

$$\left[ S_n + \frac{\bar{Z}^c F_n^L}{\bar{Z}^c + F_n^L} \right] W_n^s = \frac{\epsilon_n}{2\pi} \gamma F_n^w, \quad (29)$$

which is similar to Eq. (1), with the exception that each of the impedance terms are defined in terms of displacements instead of velocities.

In the scattering case [Fig. 3(b)] the displacement of the outside surface of the shell or the coating  $w^a(\phi)$ , satisfies the condition,

$$j\omega[w^a(\phi) - w^s(\phi)] = \begin{cases} 0 & |\phi| < \phi_0 \\ -\frac{p_s(\phi)}{Z^c} - \frac{p_i(\phi)}{Z^c} & \phi_0 < |\phi| < \pi \end{cases}. \quad (30)$$

In this case the incident plane wave is from an angle of 180 degrees, hence the coating is assumed to be located between  $\pm(\pi - \phi_0)$ . Applying the modal decomposition,

$$\begin{aligned} \frac{2\pi}{\epsilon_n} (W_n^a - W_n^s) = - \sum_m \frac{F_m^L}{j\omega\gamma Z^c} A'_{mn} W_m^a \\ - 4\gamma P_i \sum_m \frac{1}{j\omega\gamma Z^c} A'_{mn} \frac{\epsilon_m}{2\pi} \frac{j^{m+1}}{\chi H'_m(\chi)}, \end{aligned} \quad (31)$$

where

$$\begin{aligned} A'_{mn} = 2 \int_{\phi_0}^{\pi} \cos(m\phi) \cos(n\phi) d\phi \\ = \begin{cases} \left[ -\frac{\sin[(n+m)\phi_0]}{(n+m)} - \frac{\sin[(n-m)\phi_0]}{(n-m)} \right] & n \neq m \\ \left[ (\pi - \phi_0) - \frac{\sin(2n\phi_0)}{2n} \right] & n = m \\ [2(\pi - \phi_0)] & n = m = 0 \end{cases}. \end{aligned} \quad (32)$$

From Eq. (14), this time retaining only the incident and scattering components,

$$S_n W_n^s + F_n^L W_n^s + F_n^L [W_n^a - W_n^s] = -4\gamma P_i \frac{\epsilon_n}{2\pi} \frac{j^{n+1}}{\chi H'_n(\chi)} \quad (33)$$

and substituting Eqs. (11) and (31) into Eq. (33);

$$\begin{aligned} S_n W_n^s + F_n^L W_n^s - F_n^L \frac{\epsilon_n}{2\pi} \sum_m \frac{F_m^L}{\bar{Z}^c} A'_{mn} W_m^a \\ - 4\gamma P_i \frac{\epsilon_m}{2\pi} F_n^L \sum_m \frac{1}{\bar{Z}^c} A'_{mn} \frac{\epsilon_m}{2\pi} \frac{j^{m+1}}{\chi H'_m(\chi)} \\ = -4\gamma P_i \frac{\epsilon_n}{2\pi} \frac{j^{n+1}}{\chi H'_n(\chi)}, \end{aligned} \quad (34)$$

but from Eq. (14) leaving out the applied external load,

$$W_m^a = \frac{-4\gamma P_i \frac{\epsilon_m}{2\pi} \frac{j^{m+1}}{\chi H'_m(\chi)} - S_m W_m^s}{F_m^L}. \quad (35)$$

Substituting Eq. (35) into Eq. (34) and rearranging:

$$\begin{aligned} S_n W_n^s + F_n^L W_n^s + \frac{F_n^L}{\bar{Z}^c} \frac{\epsilon_n}{2\pi} \sum_m A'_{mn} S_m W_m^s \\ = -4\gamma P_i \frac{\epsilon_n}{2\pi} \frac{j^{n+1}}{\chi H'_n(\chi)}. \end{aligned} \quad (36)$$

If the coating extends over the whole of the circumference, that is,  $|\phi_0| = 0$ , then Eq. (36) simplifies to

$$\left[ S_n + F_n^L + \frac{F_n^L S_n}{\bar{Z}^c} \right] W_n^s = -4\gamma P_i \frac{\epsilon_n}{2\pi} \frac{j^{n+1}}{\chi H'_n(\chi)} \quad (37)$$

which is similar to Eq. (2), again with the exception of displacements to velocities.

Equations (28) and (36) are a set of linear equations, where the unknowns are the amplitudes of the shell surface displacement for a set of circumferential modes. Note that the presence of the partial coating couples the circumferential modes. For full coatings the circumferential modes are uncoupled [Eqs. (29) and (37)]. A solution to these linear sets of equations can be obtained by rewriting Eqs. (28) and (36) in a matrix form and the solution obtained using matrix algebra.

Analyzing Eq. (29), the coating impedance modifies the influence of the fluid-loading. If the coating covers the whole of the shell surface and if  $\bar{Z}^c \ll F_n^L$ , then the influence of the coating is significant, and the shell response approaches that for an unloaded shell. If  $\bar{Z}^c \gg F_n^L$ , then the coating has no influence and the response of the shell is similar to that of a fluid-loaded shell. Similarly, from Eq. (37) for the completely coated shell, if  $\bar{Z}^c \ll F_n^L$ , then the response of the shell is controlled by the impedance of the coating, while if  $\bar{Z}^c \gg F_n^L$ , then the influence of the coating is insignificant.

In the analysis performed in this paper, the loss factor of the coating [Eq. (26)] is assumed to be equal to  $\eta = 0.01$ , and the ratio of the coating impedance to the fluid acoustic impedance is assumed to be equal to  $\xi = 0.01$ . This value of  $\xi$  may be on low side compared to typical coating stiffness. Both these quantities are assumed constant with frequency.

For the partial coating cases considered the coating is assumed to cover an angular distance of  $0.8\pi$ , that is, a value of  $\phi_0$  in Eqs. (16) and (30) of  $0.4\pi$ .

## RESULTS

Using the above formulation results are generated for the velocity response, and acoustic scattered and radiated pressure from a water-loaded steel cylindrical shell partially coated around its circumference, excited either by an internal force or by an incident plane acoustic wave. With the shell partially coated, the internal applied load is in the middle of the coating [Fig. 3(a)], and for the excitation by the plane acoustic wave, the angle of incident of the acoustic wave is symmetric about the sector of the shell with the partial coating. Furthermore, the partial coating is on the front scattering surface [Fig. 3(b)] of the shell. The shell is assumed to be undamped, while the coating is assumed to have the properties given at the end of the previous section.

The results presented in this paper include:

- The amplitude of the input velocity response at the location of the internal load and at beam aspect when the shell is excited by a plane acoustic wave.
- The real part of the input velocity response also at the location where the internal load is applied.
- The far-field radiated pressure at beam aspect for both the case of internal force excitation and the case of excitation by a plane acoustic wave.
- The circumferential distribution of the velocity response for both internal force excitation and plane wave excitation at an arbitrarily selected frequency.
- The circumferential distribution of the near-field radiated and scattered pressure at an arbitrarily selected frequency, also for both internal force excitation and plane wave excitation.

These results are presented for the following coating conditions:

- uncoated;
- fully coated with an undamped coating ( $\eta=0$ ,  $\xi=0.01$ );
- fully coated with a damped coating ( $\eta=0.01$ ,  $\xi=0.01$ );
- fully coated with a highly compliant undamped coating ( $\eta=0$ ,  $\xi\rightarrow 0.0$ ); and
- partially coated ( $\eta=0.01$ ,  $\xi=0.01$ ).

Shown in Fig. 4 is the magnitude of the velocity response on the surface of the shell for all coating conditions considered, for the location where the internal force is applied. As can be observed irrespective of the condition of the coating, the magnitude of the velocity response at the location where the internal load is applied is generally the same, and dominated by the resonance behavior associated with the quasi-flexural circumferential waves. Some additional resonance peaks can be observed for the case of partial coating, where these additional peaks are introduced by reflections from the coating boundaries and the presence of the slow waves in the coating shell combination. The same general level in the magnitude

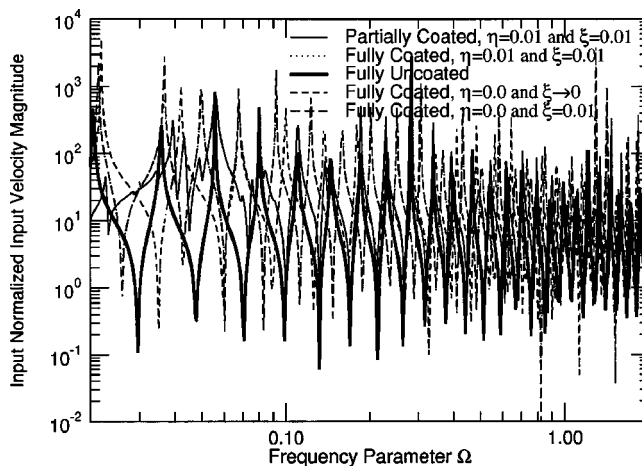


FIG. 4. Magnitude of the velocity response of shell at location of the applied internal force in the center of coating patch—radiation problem.

of the velocity response can be attributed to the fact that the response is dominated by circumferential quasi-flexural waves which are subsonic and thus not significantly influenced by the presence of the fluid. The differences that are noticeable between the responses for the different coating conditions are the location of the peaks. However, for the two fully coated shells, the level of the peaks is lower with the damped coating. This is expected since the damping will dissipate some of the resonant energy. The mean levels of all the responses is, however, different, but this is difficult to observe from the plot in Fig. 4.

Figure 5 shows the real part of the velocity response at the location of the applied load. The real part of the velocity response at the location of the applied internal load is a measure of the input power to the shell. Input power is given by the product of the input force and the input velocity, taking into account the phase difference between the force and the velocity. In the frequency domain the power input is given by the real part of the product of the input force and the conjugate of the velocity. Expressing the velocity response as the velocity per unit load (definition of structural mobility) the input power is given by the product of the real part of the structural mobility and the magnitude squared of the input

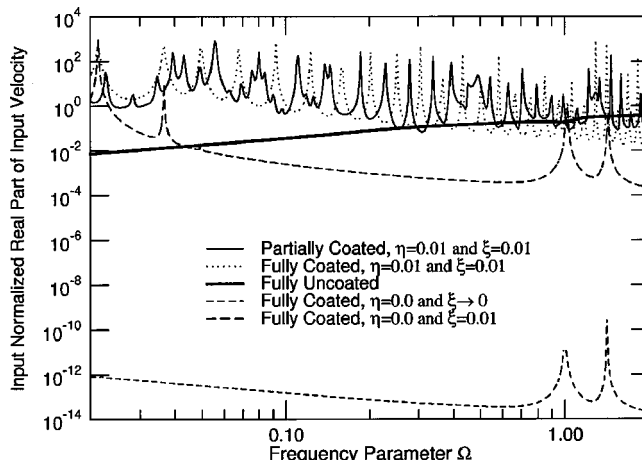


FIG. 5. Real part of the velocity response of shell at location of the applied internal force in the center of coating patch—radiation problem.

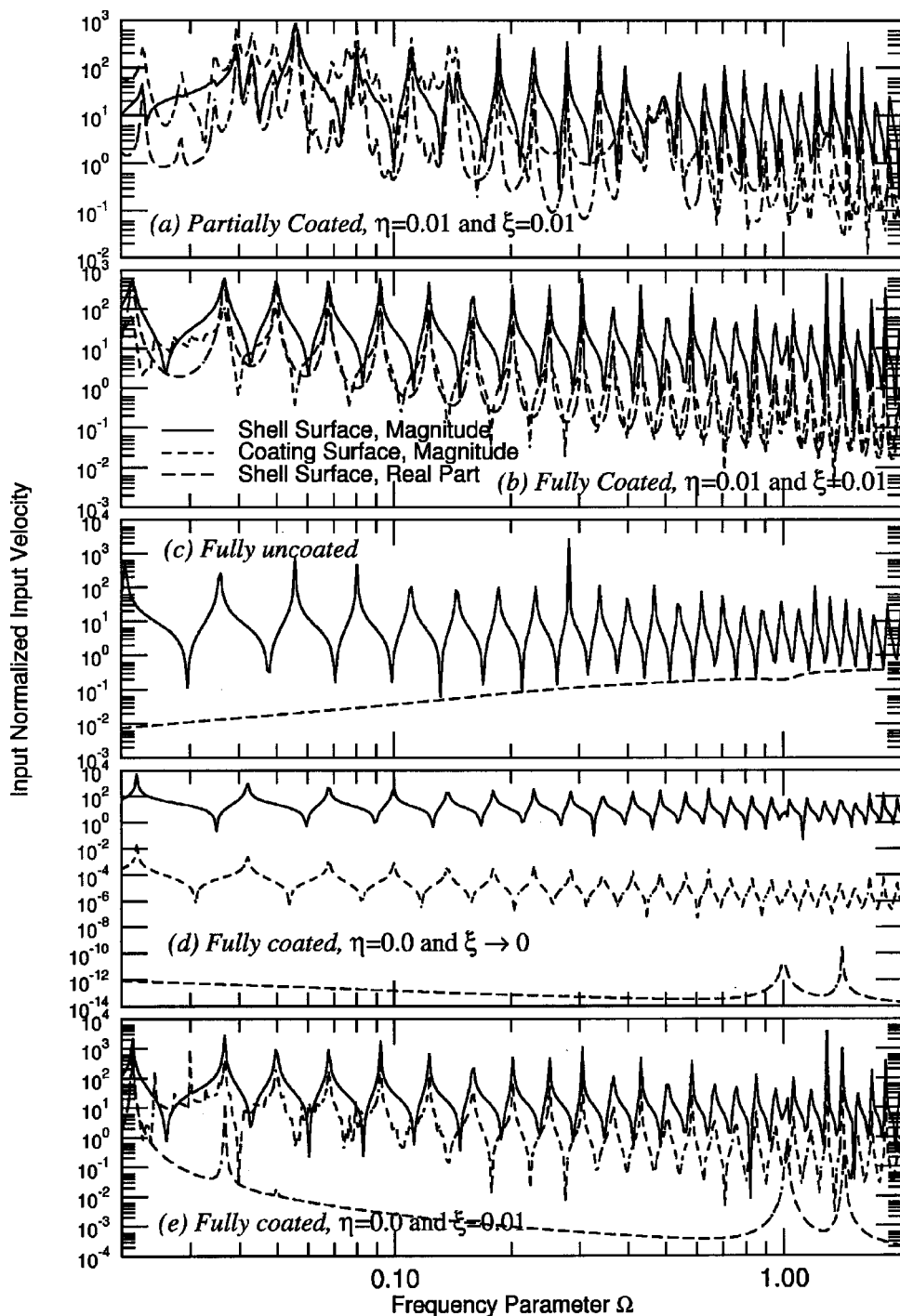


FIG. 6. Velocity response of shell and the coating surfaces at location of the applied internal force in the center of coating patch—radiation problem.

force.<sup>9</sup> For a unit input force, the input power becomes directly proportional to the real part of the input velocity. Through power conservation, the input power is either dissipated due to the damping introduced by the coating (if the coating has damping) or dissipated as acoustic radiation.

With no coating, the shell is strongly coupled to the surrounding fluid, and can accept a relatively high level of input power which is dissipated by the acoustical radiation. The resonant modes of the shell (Fig. 4) cannot accept any power since there is no dissipation (damping) associated with these resonances. These modes are subsonic. With the shell fully coated with a highly compliant but undamped coating, the shell is practically uncoupled from the acoustic

medium and since the shell and coating have no damping, the power introduced in the shell is very low. In fact, theoretically, if the coating totally decouples the shell from the acoustic medium and the coating has no damping, then the input power should be zero. In this case as well the resonant modes of the shell do not accept any power since there is no dissipation mechanism. With the shell fully coated by the undamped coating, the circumferential modes do not have a mechanism to dissipate the energy, and while there are modes present (Fig. 4), the real part of the input velocity response, which represents the power input, does not show the presence of the quasi-flexural resonances of the shell. With the shell fully coated with a 1% damping coating, en-

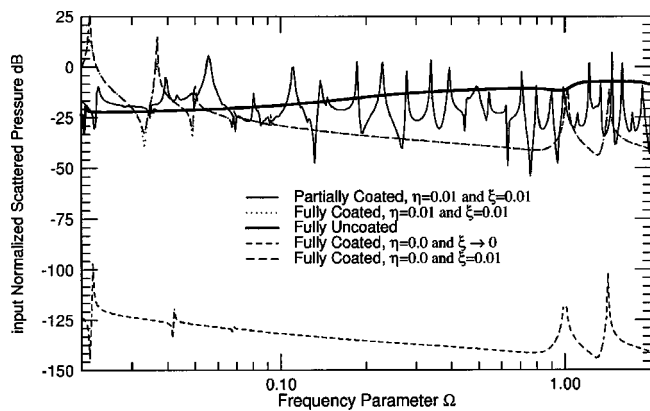


FIG. 7. Far-field radiated pressure from the shell at beam aspect (in line with the location of the applied internal force in the center of coating patch).

ergy can be dissipated by the resonant modes of the shell. These modes are damped by the presence of the damped coating, and thus a resonant behavior is observed in the real part of the velocity response of the shell. Power is transferred to the resonant modes of the shell to be dissipated by the damping of the coating. In this case, even though the shell is mostly uncoupled from the acoustic medium, the damping of the coating can dissipate the input power.

With the shell partially coated by a coating with 1% damping, the real part of the velocity response of the shell also exhibits a number of resonant peaks. In this case, power is dissipated by the damping of the coating and by the radiated acoustical waves. Thus the circumferential modes of the shell and the slow waves in the coating can receive power, which shows as peaks in the real part of the input velocity. Compared to the shell fully coated by a coating with 1% damping, which also had a strong resonant behavior, the partial coating condition has more peaks. These additional peaks can be attributed to the reflections of the quasi-flexural and slow waves at the edges of the partial coating.

Figure 6 shows a comparison of the magnitude of the velocity response on the shell surface and the magnitude of the velocity response on the coating surface. Also shown is the real part of the velocity response on the shell surface, at the location where the internal force is applied. This figure demonstrates the relationship between the amplitude of the response of the shell surface, the coating surface, and the real part of the response for the coating cases considered. Clearly seen is the influence of either damping or acoustic radiation on the real part and the magnitude of the velocity response of the shell and the coating. Some of the characteristics that are observed in this figure have already been commented on when analyzing the results as presented in Figs. 4 and 5. The benefits of Fig. 6 is to show the difference in the response between the shell surface and the coating surface, and also to show the relationship between the magnitude of the response and the real part of the response. Only for the fully coated by a low impedance coating condition there is an appreciable difference in the response between that of the coating and that of the shell. This is attributed to the dominance of the quasi-flexural subsonic waves.

Figure 7 shows the far-field (5 radii away from the shell surface) radiated acoustic pressure at beam aspect. The far-

field pressure result is similar to the result for the real part of the input velocity response, shown in Fig. 5 for the conditions where either the coating is undamped or the shell is uncoated. This is expected since the far-field pressure is a representation of the radiated power, which if the coating and the shell are undamped, from a power balance requirement, must equal to the input power. The two conditions where the far-field pressure results are different from those of the real part of the velocity response are for the partially damped coating and fully damped coating. In the case of the partial coating, the far-field scattered pressure has fewer peaks compared to the real part of the input velocity. Some of the peaks in the real part of the response are attributed to shell resonances, which are subsonic waves and therefore do not radiate. Only the supersonic modes created by the edges of the coating radiate into the far-field. For the shell fully coated with a damped coating, the far-field radiated pressure results also show fewer peaks when compared to the real part of the velocity response results. The peaks in the real part of the velocity response associated with the subsonic shell circumferential quasi-flexural waves do not show in the far-field radiated pressure. Since these waves are subsonic, they do not contribute to the radiated far-field. In fact, the far-field pressure results for the shell fully coated with a damped coating are similar to those of the undamped case except for a reduction in the level of the peaks.

The circumferential distribution of the velocity amplitude response of the shell, when the shell is excited by an internal force, and the radiated near-field sound pressure for an arbitrarily selected frequency parameter of 0.2 are shown in Figs. 8 and 9 for the uncoated, partially coated, and fully coated conditions. Figure 8 shows the distribution of the normalized amplitude of the velocity response at the selected frequency, on both the shell surface and the coating surface normalized with respect to the maximum shell velocity amplitude at the location where the force is applied. This normalization removes the influence of changes in the amplitude due to frequency and any resonant behavior of the coated shell. These figures are presented for comparison of the velocity amplitude distribution between the three conditions.

For the uncoated shell [Fig. 8(a)] there is only one response shown, that of the shell surface, and the predominant response is due to the quasi-flexural shell waves. When the shell is fully coated [Fig. 8(c)] the response of the coating surface is significantly below that of the shell surface. It is this decoupling that reduces the acoustic radiation from the shell. However, there is still some residual velocity associated with the quasi-flexural waves. The slow waves due to the presence of the coating are not apparent. With a partial coating [Fig. 8(b)] over the section of the shell which is coated, a very slow wave, short wavelength (compared to the quasi-flexural wave) can be observed on the coating surface. The amplitude of this wave is controlled by the compliance of the coating and the selected frequency. In the case shown, the selected frequency parameter ( $\Omega \cong 0.2$ ) is close to a resonance of the slow waves in the coating introduced by reflections from the edges of the coating.

The normalized radiated pressure results (Fig. 9) follow from the velocity response results of Fig. 8. In this case as

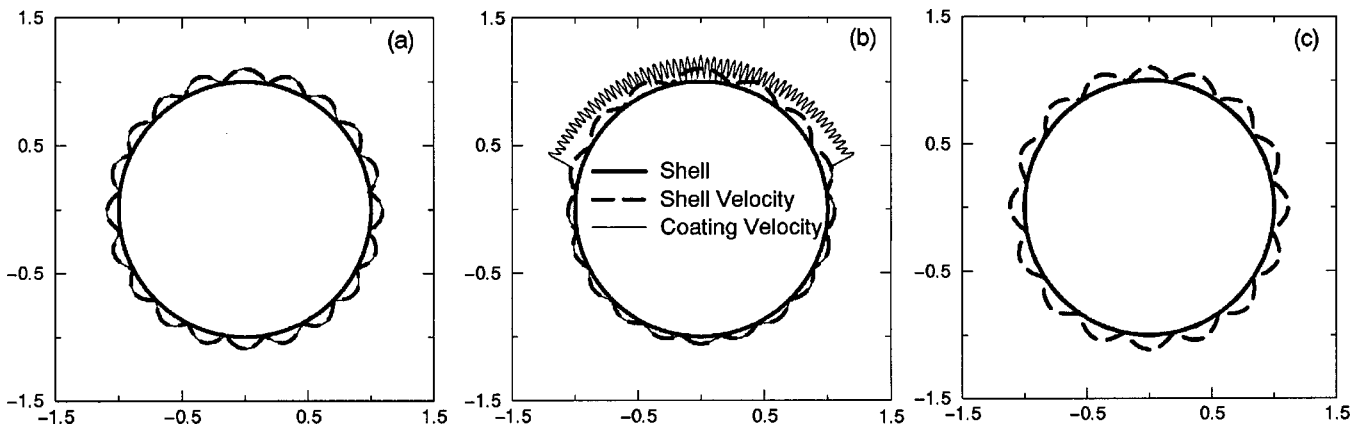


FIG. 8. Normalized circumferential distribution of velocity response for a frequency parameter of  $\Omega=0.2$ , for excitation by internal force—radiation problem. Force located at  $\phi=0$ . (a) Uncoated; (b) partially coated within  $\pm 0.4\pi$ ; (c) fully coated. —: Coating surface response; - - -: shell surface response.

well, because of the presence of resonances and the arbitrarily selected frequency, only a comparison of the distribution of the near-field pressure, and not the amplitude, can be made between the components of Fig. 9. The quasi-flexural waves are subsonic and therefore only contribute in the very near field. As can be observed from Fig. 9(a) and (c), a strong acoustic field exists close to the shell/coating surface, which quickly dissipates with the radial distance from the shell. The radiated sound field in the fully coated case is generally low due to the decoupling provided by the coating (see also Fig. 7). The radiated near field is modified by the partial coating, especially over the uncoated section and near the edges of the coating. This is consistent with the surface velocity results since over the coated section the waves are short wavelength slow waves, which would imply inefficient acoustic sources. However, most of the observed scattering on the uncoated side of the shell has more of a global feature

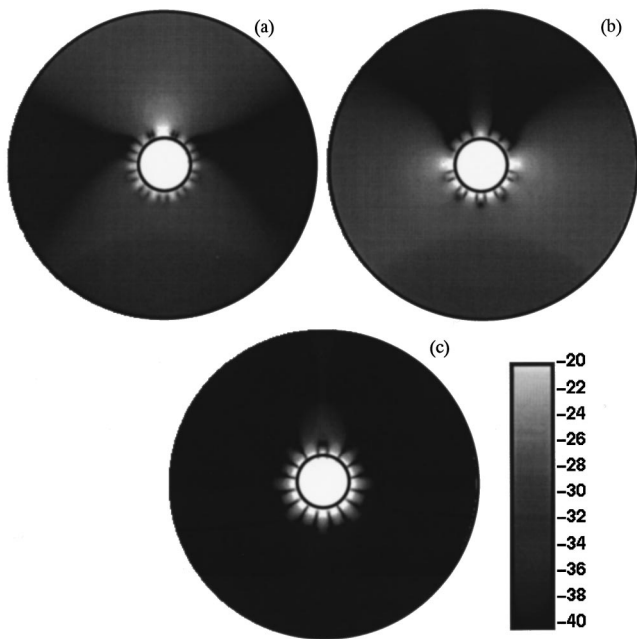


FIG. 9. Scattered pressure for a frequency parameter of  $\Omega=0.2$ , for excitation by internal force—radiation problem. Force located at  $\phi=0$ . (a) Uncoated; (b) partially coated within  $\pm 0.4\pi$ ; (c) fully coated.

and is not directly related to the circumferential quasi-flexural waves. This is why this radiation is attributed to the contributions from the edges of the coating. The edges of the coating create discontinuities which can indirectly enhance the acoustic radiation. Smoothing of the edges of the partial coating could have a significant influence on the strength of the radiated acoustic field, as suggested by Eq. (33).

The scattering patterns are, however, very much dependent on the values of  $ka(\Omega M)$  and  $[ka(2\phi_0)]$ ; the latter represent the relative extent of the partial coating to an acoustic wavelength. The scattering patterns are partially due to the resonances induced in the shell/coating combination.

Figure 10 shows the magnitude of the velocity response at shell beam aspect, as a function of the frequency parameter, when the shell is excited by an incident plane acoustic wave. The results shown are for the uncoated, fully coated with an undamped coating ( $\eta=0$ ,  $\xi=0.01$ ), fully coated with a damped coating ( $\eta=0.01$ ,  $\xi=0.01$ ), fully coated with a highly compliant undamped coating ( $\eta=0$ ,  $\xi \rightarrow 0$ ), and partially coated ( $\eta=0.01$ ,  $\xi=0.01$ ) conditions. As can be observed from these results, for the case of fully uncoated shell, the shell receives high energy which is acoustically scattered. Because of the strong coupling between the shell and the surrounding fluid in this fully uncoated condition, energy introduced in the shell is immediately scattered back into the

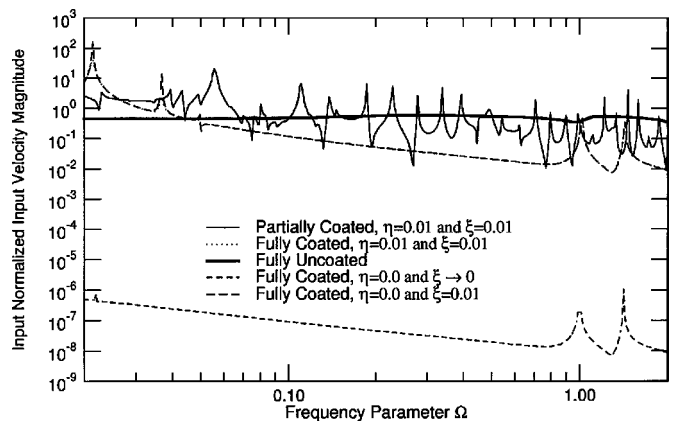


FIG. 10. Normalized velocity response magnitude at the center of the coating patch for excitation by an incident acoustic wave—scattering problem.

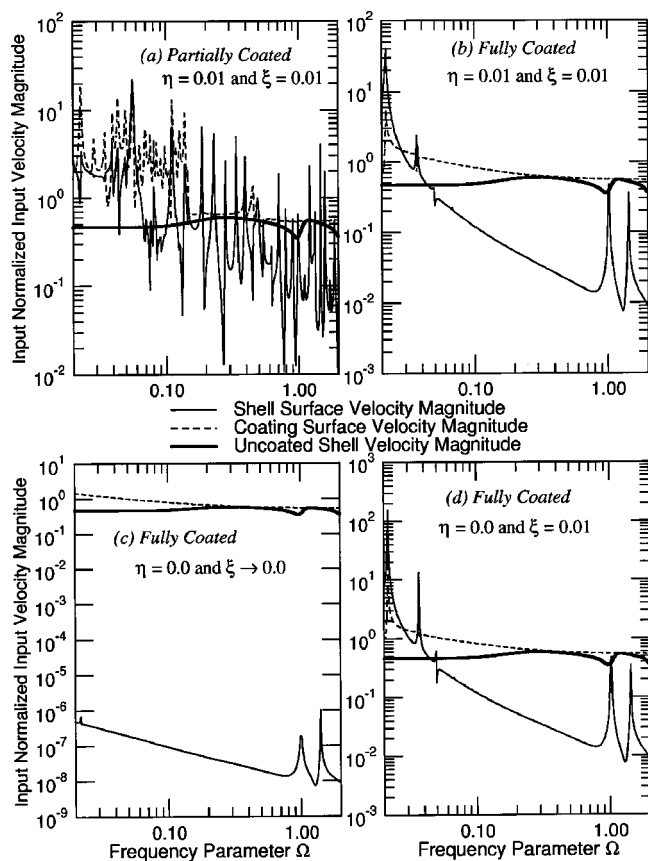


FIG. 11. Velocity response of shell and the coating surfaces in the center of the coating patch—scattering problem.

fluid precluding the formation of resonant behavior.

With the shell fully coated with a highly compliant undamped coating, the shell is practically uncoupled from the acoustic medium. In this case the response of the shell to acoustic wave excitation is significantly reduced. For the coated shell ( $\xi=0.01$ ), and with acoustic wave excitation, the damping does not play a significant role, as can be observed from the results for the velocity response of the fully coated shell with an undamped coating and a 1% damping coating. For these two conditions, the responses are practically identical. This would be expected since the coating is providing the decoupling and resonances (where damping would be important) are not introduced in the shell, except at a few selected frequencies. For the partial coating condition, a resonant behavior is introduced in the shell, with the resonant peaks associated with the shell quasi-flexural waves and the coating slow waves, with reflection from the edges of the coating. In this case it is hypothesized that the interaction of the acoustic waves with the edges of the coating generates the forces that excite the shell circumferential modes.

Figure 11 shows, in separate graphs, the magnitude of the velocity response on the shell surface and the magnitude of the velocity response on the coating surface for the different coating conditions. Also shown on each plot is the magnitude of the velocity response of the uncoated shell. All responses are at the center of the coating patch or its equivalent location for the fully coated cases. This figure shows the relationship between the amplitude of the response on the coating surface and the shell surface, and how the different

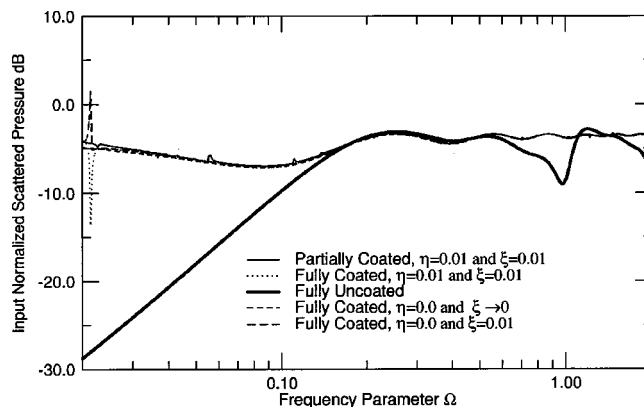


FIG. 12. Far-field scattered pressure from the shell at beam aspect.

coating conditions considered compare to the uncoated case. The benefits of Fig. 11 are to show the difference in the response between the shell surface and the coating surface. For the shell fully coated by a highly compliant coating, there is an appreciable difference in the response of the coating and that of the shell, and as expected the coated shell response is significantly lower than that of the uncoated shell. For the other fully coated conditions the difference in the response increases with frequency, and for high frequencies the response of the coated shell is again lower than the response of the uncoated shell. For the case of the partial coating, a resonant behavior is observed in the response of the shell. This resonant behavior is predominantly caused by quasi-flexural waves and coating/shell interaction waves. The mean level of the response of the shell is comparable to that of the uncoated shell. However, the response at the peak of the resonances is much higher than that of the uncoated shell, especially at low frequencies.

The far-field scattered pressure results at beam aspect for the scattering problem are shown in Fig. 12. The far-field scattered pressure is generally similar for all the coating conditions considered. This is not surprising since, for the frequency range considered, most of the scattering is due to the global behavior of the shell and the coating. Noticeable is the large difference in the scattering between the fully uncoated and the coated conditions at low frequencies. This is attributed to the larger low frequency response of the coating outer surface due to its compliance. The trend of the far-field scattered pressure results matches that of the coating surface velocity response, except for the condition of the partial coating. The far-field pressure does not show any of the resonance peaks observed in the velocity response. This is why these peaks (Fig. 11) are attributed to quasi-flexural and shell/coating subsonic waves.

The shell circumferential velocity amplitude distribution and the scattered near-field sound pressure distribution when the shell is excited by an incident plane acoustic wave, for an arbitrary frequency parameter of 0.34, are shown in Figs. 13 and 14 for uncoated, partial coated, and fully coated conditions. The amplitude of the velocity response on both the shell surface and the coating surface are shown. In this case as well, the velocity amplitude curves and the scattered pressure have been normalized with respect to the velocity amplitude and scattered pressure at the shell beam aspect, to



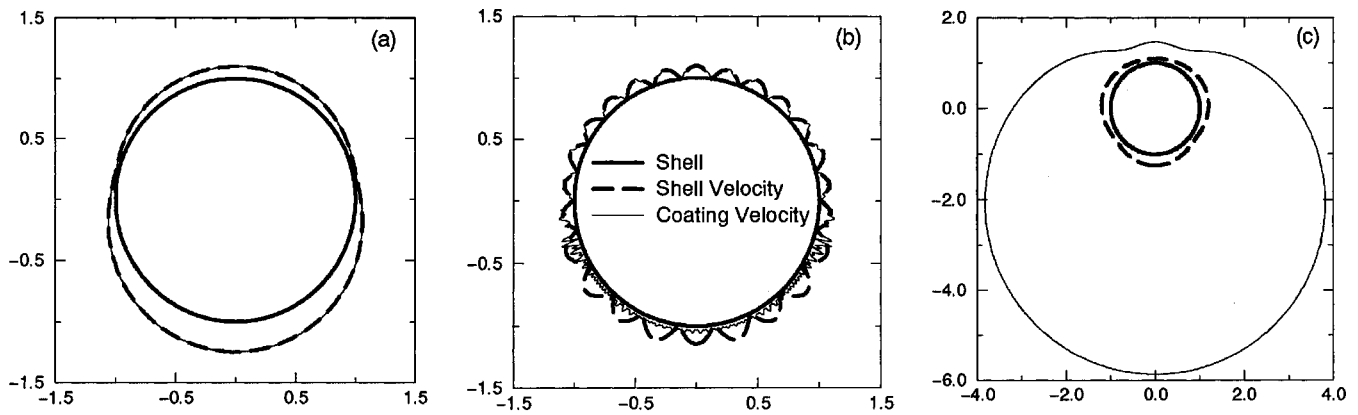


FIG. 13. Circumferential distribution of velocity response for a frequency parameter of  $\Omega=0.34$ , for excitation by incident acoustic wave—scattering problem. Angle of incidence  $\phi=\pi$  (a) uncoated; (b) partially coated within  $\phi=\pi\pm 0.4\pi$ ; (c) fully coated. —: Coating surface response; - - -: shell surface response.

remove the influence of frequency and the presence of the resonances. From these figures a comparison can be made of the distribution of the velocity amplitude and the scattered pressure for the coating conditions considered.

Figure 13(a) shows the shell surface response for the uncoated case. Predominant in this response is the  $n=1$  mode. When fully coated [Fig. 13(c)] the coating surface response is significantly higher than that of the shell surface. Hardly any energy reaches the shell. For the partially coated [Fig. 13(b)] over the section of the shell which is coated, the shell has a response induced by the waves from the uncoated section. The coating surface, on the other hand, exhibits hardly any response and the predominant response of the coating surface is that associated with the slow, short wavelength waves. The slow waves are a consequence of the interaction between the shell and the coating. The response of

the coating over this section is lower than that of the shell. This implies that the energy induced in the shell from the uncoated section is decoupled and not allowed to find itself back into the fluid.

The circumferential distribution of the near-field scattering results (Fig. 14) follow the circumferential distribution of the velocity response results of Fig. 13. The scattering is dominated by the global shell/coating motion. The contribution from the subsonic quasi-flexural waves is insignificant. In fact for all three coating conditions considered, the distribution of the scattered pressure around the shell circumference is, in very general terms, somewhat similar with the fully uncoated case, showing weaker scattering in the backward  $\pm 45^\circ$  directions. The scattering is strongest in the back scattering direction and weakest in the forward scattering direction. These results follow because of the relatively low frequency selected. At high frequencies, when the acoustic wavelength is shorter, the contribution from the shell quasi-flexural waves can become apparent.

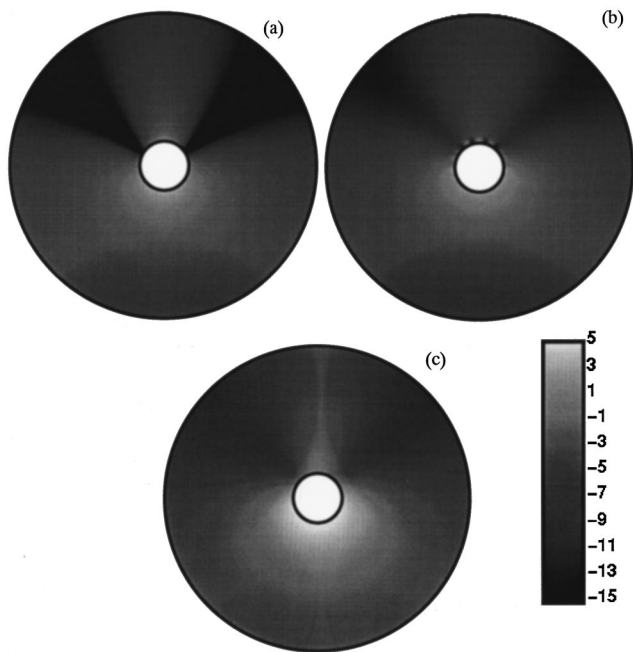


FIG. 14. Scattered pressure for a frequency parameter of  $\Omega=0.2$ , for excitation by incident acoustic wave—scattering problem. Angle of incidence  $\phi=\pi$  (a) fully uncoated; (b) partially coated within  $\phi=\pi\pm 0.4\pi$ ; (c) fully coated.

## SUMMARY AND CONCLUSIONS

In this paper a model and analysis of a shell with no axial dependency and circumferential partial coating has been developed. In the analysis it is shown that, with a partial coating, the circumferential modes of the shell become coupled. This conclusion is similar to that found in Ref. 5. Also, from the formulation [Eq. (33)], one can observe that the strength of the cross coupling between the modes is a function of the edge “smoothness” of the coating layer. Using the developed formulation, results have been generated for both radiation and scattering conditions. Computed results for the response of the shell, as well as the near-field scattered pressure, when the fluid-loaded shell is either loaded by an applied internal load or by an incident plane acoustic wave, are presented. Comparison is made between the three cases when the shell is fully coated, partially coated over a sector of  $0.8\pi$  radians and fully uncoated. From the radiation results the following conclusions can be drawn:

- (a) Due to the direct application of the internal force on the shell, a resonant behavior is obtained. The resonant be-

havior is predominantly circumferential (quasi-flexural) subsonic waves which only have an influence on the acoustic near field.

- (b) With the partial coating a slow velocity wave (short wavelength) is introduced in the portion of the shell which is coated.
- (c) The edges of the coating in the partial coating case create strong radiation zones. It would be anticipated that the edge “smoothness” of the partial coating would play a significant role in the strength of the radiated acoustic field.
- (d) The damping of the coating plays a significant role in the response of the shell.

From the scattering results, the following observations can be made:

- (a) At low frequencies, the scattering patterns with and without a coating are not significantly different. This is attributed to the generally similar deformation of the shell. At low frequencies, the only difference in the response is the onset of the subsonic waves which do not influence the scattering.
- (b) A slow velocity wave (short wavelength) is apparent in the coating. The amplitude of this wave is low compared to the shell resonant response.
- (c) A partial coating induces a resonant behavior in the response of the shell. These resonant waves, which are subsonic, can modify the near-field scattering at high frequencies.
- (d) With full coating, the damping of the coating does not play a significant role in the velocity response of the shell.

The general conclusion is that radiation and/or scattering may increase due to the partial coating. This conclusion is in agreement with the results obtained in Ref. 5. In this paper only results for  $\Omega \approx 0.2$  are presented, however, a complete

sequence of results for the frequency parameter range 0.02–2.0 have been generated and presented.<sup>10</sup> The conclusions drawn here apply equally well within this frequency range. As a final comment, the results presented here are for a shell with no axial dependency or a shell with an internal line force source or scattering at beam. Results taking into account the shell axial dependency will be the subject of a future paper.

## ACKNOWLEDGMENT

This work was partially sponsored by NSWC-CD and ONR.

<sup>1</sup>D. Feit and J. M. Cuschieri, “Scattering of sound by a fluid-loaded plate with a distributed mass-inhomogeneity,” *J. Acoust. Soc. Am.* **99**, 2686–2700 (1996).

<sup>2</sup>G. Maidanik and A. J. Tucker, “Acoustic properties of coated panels immersed in fluid media,” *J. Sound Vib.* **34**, 519–550 (1974).

<sup>3</sup>G. Maidanik, R. Biancardi, and T. Eisler, “Use of decoupling to reduce the radiated noise generated by panels,” *J. Sound Vib.* **81**, 165–185 (1982).

<sup>4</sup>A. A. Ferri, J. H. Ginsberg, and P. H. Rogers, “Scattering of plane waves from submerged objects with partially coated surfaces,” *J. Acoust. Soc. Am.* **92**, 1721–1728 (1992).

<sup>5</sup>B. Laulagnet and J. L. Guyader, “Sound radiation from finite cylindrical shells, partially covered with longitudinal strips of compliant layer,” *J. Sound Vib.* **186**, 723–742 (1995).

<sup>6</sup>J. M. Cuschieri and D. Feit, “Full numerical solution for the far field and near field scattering from a fluid-loaded elastic plate with distributed mass or stiffness inhomogeneity,” *J. Acoust. Soc. Am.* **104**, 915–925 (1998).

<sup>7</sup>J. M. Cuschieri and D. Feit, “Response Green’s function and acoustic scattering from a fluid-loaded cylindrical shell with discontinuities,” *The International Journal of Computers and Structures, Special Issue “Structural Acoustics,”* November 1997, **65**, 337–384 (1997).

<sup>8</sup>M. C. Junger and D. Feit, *Sound Structures and Their Interaction* (Acoustical Society of America, New York, 1995).

<sup>9</sup>R. J. Pinnington and R. G. White, “Power flow through machine isolators to resonant and nonresonant beams,” *J. Sound Vib.* **75**, 179–197 (1981).

<sup>10</sup>J. M. Cuschieri and D. Feit, “Influence of partial coatings on the acoustic radiation from a fluid-loaded structure,” *J. Acoust. Soc. Am.* **101**, 3174(A) (1997).

# Acoustic scattering from fluid-loaded stiffened cylindrical shell: Analysis using elasticity theory

André Baillard, Jean-Marc Conoir, Dominique Décultot, and Gérard Maze  
*Laboratoire d'Acoustique Ultrasonore et d'Electronique UPRESA CNRS 6068, Université du Havre,  
place Robert Schuman, 76610 Le Havre, France*

Aleksander Klauson and Jaan Metsaveer  
*Department of Mechanics, Tallinn Technical University, Ehitajate tee 5, 19086 Tallinn, Estonia*

(Received 23 November 1999; revised 7 February 2000; accepted 10 February 2000)

The acoustic scattering from a fluid-loaded stiffened cylindrical shell is described by using elasticity theory. The cylindrical shell is reinforced by a thin internal plate which is diametrically attached along the tube. In this model, cylindrical shell displacements and constraints expressed from elasticity theory are coupled to those of the plate at the junctions, where plate vibrations are described by using plate theory. The present model is first validated at low frequency range ( $k_1 a \sim 5-40$ ) by comparison with a previous model based on the Timoshenko–Mindlin thin shell theory and by experimental results. Theoretical and experimental resonance spectra are then analyzed in a high frequency range ( $k_1 a \sim 120-200$ ). Only resonances due to the  $S_0$  wave are clearly observed in this frequency range, and their modes of propagation are identified. Furthermore,  $A_0$  wave propagation is detected, because of the presence of the reflection of this wave at the shell-plate junctions. © 2000 Acoustical Society of America. [S0001-4966(00)03205-7]

PACS numbers: 43.40.Fz, 43.20.Fn, 43.20.Tb [CBB]

## INTRODUCTION

The scattering problem for an infinitely long cylindrical shell, stiffened by an internal lengthwise substructure, based on a thin shell theory, has provided an active research topic in recent years.<sup>1-7</sup> In Refs. 8–9, the authors have validated their model with an experimental study and have identified the observed resonances. Previous studies have shown the influence of ribs on the sound scattering by stiffened shells at normal incidence. The main contribution of the ribs is to generate flexural type waves even at very low frequencies, whereas only the membrane type waves are generated in the unstiffened shell having the same characteristics. The generation and the radiation of flexural waves are observed when a junction is directly insonified by the incident wave. Moreover, the presence of the rib can be revealed even if the joint is in the shadow zone.<sup>8</sup>

These theoretical and experimental studies of lengthwise reinforced cylindrical shells have been made in the low and mid-frequency ranges, when wavelengths in the structure and in fluid are larger than the shell thickness ( $k_1 h < 1$ , where  $k_1$  is the wave number in water, and  $h$  is the shell thickness). In that case, the use of thin shell theory is fully justified.

In order to analyze the vibrations of reinforced shells in the high frequency range, a new model is presented in the present paper. In this model, the shell displacements are described by using elasticity theory and the influence of the plate is taken into account by the constraints at the shell-plate junctions. Of course, one needs to know the reactions of the plate extremities when one of them is submitted to unitary displacements. These are obtained from equations of plate theory and they correspond to the bending moment, normal and tangential forces. In order to couple the two sub-

structures, the plate is assumed to be pinned to the shell, so that bending moments are not transmitted. However, it has been shown<sup>3</sup> that only a small difference exists between scattering by clamped and pinned plates in the frequency range of all the previous studies.

Sound scattering from an internally loaded cylindrical shell has also been examined by Pathak and Stepanishen<sup>10</sup> using elasticity theory. More accurately, they have analyzed acoustic radiation from a fluid-loaded infinite cylindrical shell submitted to arbitrary loading, as ring force.

Section I describes the studied object and the new model using elasticity theory.<sup>11,12</sup> Basic equations are recalled, as well as expressions of scalar and vector potentials which are used to obtain constraints and displacements. Displacements at the junctions are then calculated, on the one hand, from those of the shell and, on the other hand, from those of the plate, and the coupling between the two substructures is taken into account by equalizing the constraints at the internal interface. For brevity, only some general expressions are given here, and details of calculation are explained in Appendices.

Section II presents the experimental setup and the object used to validate the model. Experimental and theoretical backscattered spectra are first compared in a low frequency range ( $k_1 a \sim 5-40$ ) where numerous scalloplike variations are observed which make comparison rather difficult. Angular diagrams are easier to compare. Good agreement between the model based on elasticity theory and the experimental results has been observed.<sup>13</sup> These comparisons are also made with results obtained from a model based on thin shell theory.<sup>14</sup> The third part of this section analyzes experimental and theoretical results in a high frequency range ( $k_1 a \sim 120-200$ ). A good agreement between the model established using elasticity theory and the experimental results is

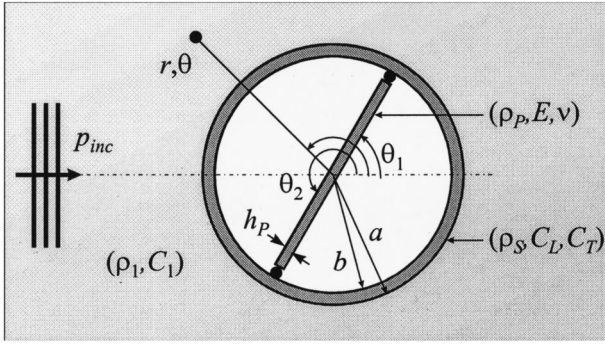


FIG. 1. Geometry of the problem.

observed whereas the limit of validation of thin shell model is exceeded. The  $S_0$  wave resonances are the only type of resonances clearly observed in the resonance spectra and the identification of the concerned mode of propagation is explained. The second predominating phenomena in the resonance spectra is due to the  $A_0$  wave which is reflected at the junctions.

## I. SOLUTION TO THE PROBLEM

Consider an elastic cylindrical shell stiffened by a diametrical lengthwise soldered plate. The shell is made of an isotropic material characterized by density ( $\rho_s$ ), and by longitudinal and transversal wave velocities ( $C_L$ ,  $C_T$ ). Its internal to external radius ratio is  $b/a=0.98$ . The plate is made of an elastic material, of density  $\rho_p$  and is characterized by Young's modulus  $E$  and Poisson ratio  $\nu$ . Its length is  $l_p=2b$  and its thickness is  $h_p$ . The plate is pinned to the shell and the junctions are considered in the calculations as lines along the shell. The object is immersed in an unbounded ideal fluid medium with a density  $\rho_1$  and a sound velocity  $C_1$ . The fluid loading inside the shell is considered absent (vacuum). A plane acoustic wave at normal incidence excites the object. As the shell is considered to be of infinite length and the incident plane wave front is parallel to the  $z$  axis, there is no influence of the  $z$  axis on the calculations. In order to effectively describe the given two-dimensional (2D) problem, polar coordinates ( $r, \theta$ ) are used and the geometry of the problem is presented in Fig. 1.

The incident pressure is described by

$$p_{\text{inc}} = p_0 \exp i(k_1 x - \omega t), \quad (1)$$

where  $k_1$  is the wave number in water,  $\omega$  the angular frequency,  $t$  time and  $p_0$  a constant. As the problem is linear, the harmonic time dependence  $e^{-i\omega t}$  in all quantities is henceforth omitted. Furthermore, these quantities can be decomposed into angular harmonics in  $\theta$  for each reduced frequency  $k_1 r$ :

$$p_{\text{inc}} = p_0 \sum_{n=-\infty}^{n=+\infty} i^n J_n(k_1 r) e^{in\theta}, \quad (2)$$

$J_n$  is the Bessel function of the  $n$ th order. The scattered pressure can be written as

$$p_{\text{scat}} = p_0 \sum_{n=-\infty}^{n=+\infty} i^n a_{1n} H_n^{(1)}(k_1 r) e^{in\theta}, \quad (3)$$

where  $H_n^{(1)}$  is the  $n$ th order Hankel function of the first kind. The total pressure  $p$  in the fluid must satisfy the wave equation, radiation condition at infinity and boundary conditions at the shell interfaces. Henceforth in this paper the bold characters shall denote vectors and matrix. The expression of the displacement in the fluid is

$$\mathbf{u}^{(f)} = \frac{1}{\rho_1 \omega^2} \nabla p. \quad (4)$$

Displacements and constraints into the shell are expressed in the theory of elasticity by the scalar potential  $\phi$  and the vector potential  $\boldsymbol{\psi}$  which satisfy the following equations:

$$\nabla^2 \phi + k_L^2 \phi = 0, \quad (5)$$

$$\nabla^2 \boldsymbol{\psi} + k_T^2 \boldsymbol{\psi} = 0. \quad (6)$$

The solution of Eq. (5) is

$$\phi = \phi_0 \sum_{n=-\infty}^{n=+\infty} i^n [a_{2n} J_n(k_L r) + a_{3n} N_n(k_L r)] e^{in\theta}, \quad (7)$$

where  $N_n$  is the Neumann function of the  $n$ th order. The vector potential has no component in the  $r$ -direction or in  $\theta$  and therefore the solution of Eq. (6) can be written

$$\boldsymbol{\psi}_z = i \psi_0 \sum_{n=-\infty}^{n=+\infty} i^n [a_{4n} J_n(k_T r) + a_{5n} N_n(k_T r)] e^{in\theta}. \quad (8)$$

The factors  $\mathbf{a}_n = (a_{1n} a_{2n} a_{3n} a_{4n} a_{5n})^t$  are the unknown coefficients which are evaluated by applying the boundary conditions at the surfaces of the cylindrical shell ( $r=a$  and  $r=b$ ). Equations (9), (10) and (11) are derived from the boundary conditions at the interface of the shell and the external fluid:

$$u_r = u_r^{(f)} \quad \text{at } r=a, \quad (9)$$

$$\sigma_{rr} = -(p_{\text{inc}} + p_{\text{scat}}) \quad \text{at } r=a, \quad (10)$$

$$\sigma_{r\theta} = 0 \quad \text{at } r=a. \quad (11)$$

The plate is taken into consideration with constraint expressions at the internal surface of the shell. As the fluid loading inside the shell is assumed to be vacuum, these constraints are equal to zero except at the plate-shell junctions ( $\theta_1$  and  $\theta_2$ ). The plate motion is described by the equations of the plate theory.<sup>15,16</sup> The reactions of the plate extremities for unitary displacements are expressed in Appendix A. As the plate has pinned-pinned attachments to the shell, moment reaction does not exist so that the cylindrical shell and the plate are coupled with only normal and tangential components. The following boundary conditions express the coupling between the shell and the plate:

$$\sigma_{rr} = - \sum_{k,l=1}^2 \delta(\theta - \theta_k) F_{k,l}^N u_r(b, \theta_l) \quad \text{at } r=b, \quad (12)$$

$$\sigma_{r\theta} = - \sum_{k,l=1}^2 \delta(\theta - \theta_k) F_{k,l}^T u_\theta(b, \theta_l) \quad \text{at } r=b, \quad (13)$$

where  $F_{k,l}^N$  and  $F_{k,l}^T$  are, respectively, normal and tangential reactions of the plate for unitary displacements, and  $\delta$  is Dirac-delta-function.

The unknown modal coefficients  $a_n$  are solutions of the equation:

$$D_n a_n = b_n, \quad (14)$$

where  $D_n$  is a  $5 \times 5$  matrix obtained previously from the consideration of the boundary conditions. Expressions of  $D_n$  and  $b_n$  are given in Appendix B. As the system is linear, it can be decomposed into three subsystems:

$$a_n = a_n^0 + a_n^N + a_n^T, \quad (15)$$

$$b_n = b_n^0 + b_n^N + b_n^T, \quad (16)$$

where  $a_n^0$  is the well-known solution for an empty shell (without stiffener). Solutions  $a_n^N$  and  $a_n^T$  are obtained by considering, respectively, normal and tangential contributions of the plate.

All coefficients of matrix  $D_n$  can be calculated, but  $b_n^N$  and  $b_n^T$  depend on the normal and tangential displacements at the junction lines (Appendix B). So the first step of the calculation is the determination of these displacements at  $r=b$  and  $\theta = \theta_k$  ( $k=1,2$ ).

$$u_r = \frac{\partial \phi}{\partial r} + \frac{1}{r} \frac{\partial \psi_z}{\partial \theta}, \quad (17)$$

$$u_\theta = \frac{1}{r} \frac{\partial \phi}{\partial \theta} - \frac{\partial \psi_z}{\partial r}. \quad (18)$$

In order to determine the normal displacements  $u_r(b, \theta_k)$ , Eq. (17) is developed as a modal sum of the coefficients  $a_n$  and is given in Appendix C. Since system Eq. (15) is linear, normal displacements are obtained from the summation of modal series due to the empty shell and those of the plate contributions:

$$u_r(b, \theta_k) = u_r^0(b, \theta_k) + u_r^N(b, \theta_k) + u_r^T(b, \theta_k). \quad (19)$$

It is easy to calculate the well-known  $u_r^0(b, \theta_k)$ . Moreover, because of diametrical position of the plate, there is no normal displacement if only tangential reactions of the plate are considered, such that

$$u_r^T(b, \theta_k) = 0. \quad (20)$$

The unknown coefficients  $a_n^N$  are solutions of Eq. (14), hence

$$a_{in}^N = \frac{|D_n^{N[i]}|}{|D_n|}, \quad (21)$$

where  $|D_n|$  is the determinant of matrix  $D_n$  and  $|D_n^{N[i]}|$  is the determinant of the matrix obtained from the matrix  $D_n$  when the  $i$ -column is replaced by the corresponding  $b_n^N$  vector. As  $a_n^N$  and  $b_n^N$  are linearly dependent on  $u_r(\theta_1)$  and  $u_r(\theta_2)$ , these displacements can be extracted from the summation in Eq. (17). The normal displacements are then calculated from the following system:

TABLE I. Materials parameters.

Parameter	Stainless steel	Water
Density (kg/m <sup>3</sup> )	7900	1000
Longitudinal velocity (m/s)	5790	1470
Transversal velocity (m/s)	3100	
Young's modulus (GPa)	197.39	
Poisson ratio	0.3	

$$\begin{bmatrix} u_r(\theta_1) \\ u_r(\theta_2) \end{bmatrix} = \begin{bmatrix} u_r^0(\theta_1) \\ u_r^0(\theta_2) \end{bmatrix} + \alpha \begin{bmatrix} S_1^N & S_2^N \\ S_2^N & S_1^N \end{bmatrix} \begin{bmatrix} F_1^N & F_2^N \\ F_2^N & F_1^N \end{bmatrix} \begin{bmatrix} u_r(\theta_1) \\ u_r(\theta_2) \end{bmatrix}, \quad (22)$$

where  $\alpha$ ,  $S_1^N$  and  $S_2^N$  are developed in Appendix C.

The tangential displacements at  $\theta_k$  are obtained in a similar way:

$$\begin{bmatrix} u_\theta(\theta_1) \\ u_\theta(\theta_2) \end{bmatrix} = \begin{bmatrix} u_\theta^0(\theta_1) \\ u_\theta^0(\theta_2) \end{bmatrix} - \alpha \begin{bmatrix} S_1^T & S_2^T \\ S_2^T & S_1^T \end{bmatrix} \begin{bmatrix} F_1^T & F_2^T \\ F_2^T & F_1^T \end{bmatrix} \begin{bmatrix} u_\theta(\theta_1) \\ u_\theta(\theta_2) \end{bmatrix}. \quad (23)$$

Normal and tangential displacements are then derived, and the second step of the calculation is the resolution of the system Eq. (14), where coefficients  $b_n$  are now determined. When the coefficients  $a_{1n}$  of the scattered pressure Eq. (3) are calculated, the far field form function can be expressed as

$$F_\infty = \frac{2}{\sqrt{\pi k_1 a}} \left| \sum_{n=-\infty}^{+\infty} a_{1n} e^{in\theta} \right|. \quad (24)$$

## II. SPECTRAL ANALYSIS

### A. Experimental setup

The studied cylindrical shell is made of stainless steel and is characterized by its internal to external radius ratio  $b/a=0.98$  with  $a=25.5$  mm and  $b=25$  mm. Its total length is 200 mm. The internal plate is made of the same material and is diametrically soldered along the tube. Its thickness is  $h_p=0.03b$ . The object is closed by two flat circular lids and is vertically immersed into a large water filled tank. The internal fluid loading is air. The characteristics of the material used in the calculations are listed in Table I.

The experimental setup is illustrated in Fig. 2. The transducer with its displacements controlled by a computer turns around the target in a plane perpendicular to its axis at a

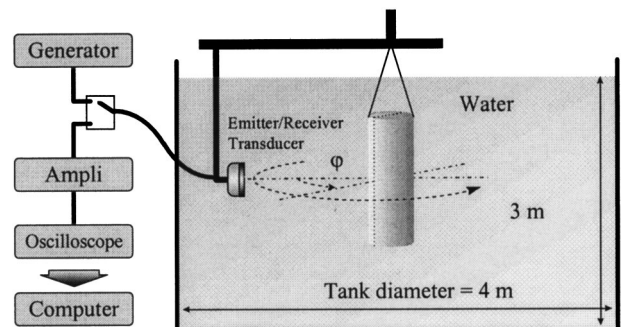


FIG. 2. Experimental setup.

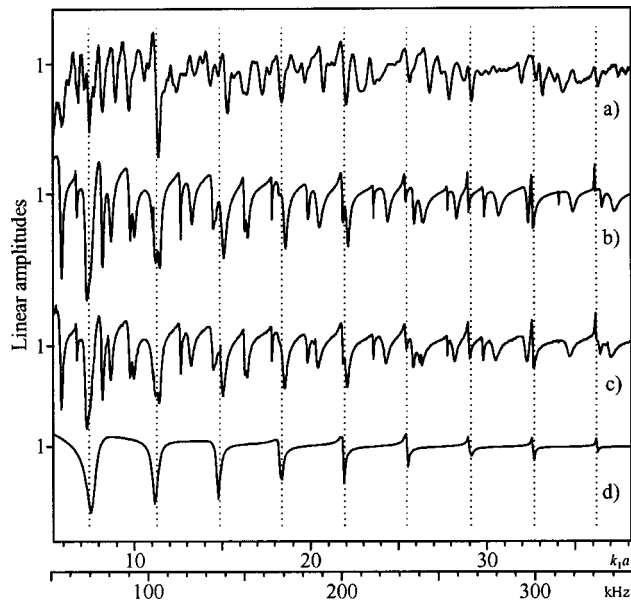


FIG. 3. Backscattering spectra at  $\varphi=0^\circ$ . (a) Experimental. (b) Model using elasticity theory. (c) Thin shell theory. (d) Unstiffened shell using elasticity theory. Dotted lines show  $S_0$  wave resonances.

fixed distance (about 1 m), and the shell is excited by a plane acoustic wave at normal incidence. The angular positions of the plate-shell junctions relative to the incident wave front direction are indicated by  $\varphi$  and  $\varphi+180^\circ$ .

The monostatic quasi-harmonic setup<sup>17,18</sup> is used to obtain experimental spectral responses: the transducer sends a train of sinusoidal signal long enough to establish a steady state vibration on the object. The backscattered signal is received by the same transducer and recorded in the forced excitation zone (backscattering spectrum) or in the free re-emission zone (resonance spectrum).

### B. Low frequency analysis (comparisons)

In order to validate the previous model, experimental measurements<sup>13</sup> and numerical results obtained from a model based on the Timoshenko–Mindlin thin shell theory<sup>6,8,9,14</sup> are compared. Studies are first made in the low frequency range. The central frequency of the used transducer (Panametrics, V3507) is 200 kHz ( $k_1a=21.8$ ) and its frequency band width is over 50–350 kHz ( $k_1a=5-38$ ).

Figure 3 shows experimental and theoretical backscattering spectra at  $\varphi=0^\circ$ . Computational and experimental steps are both at  $0.025 k_1a$ . The two calculations give identical results, but it is difficult to accurately compare them with experimental ones:  $S_0$  wave resonances illustrated by vertical dotted lines in Fig. 3 are always observed in all the cases, with a slight frequency shift when the shell is stiffened, but numerous other scalloplike variations are observed. Indeed, these resonances are due to flexural type waves circumnavigating the shell. Their frequencies are very sensitive to the internal structure,<sup>5</sup> which explains the gap between some experimental and theoretical resonance frequencies. Figure 4 corresponds to the backscattering spectra recorded from  $\varphi=0^\circ$  to  $90^\circ$  at  $2^\circ$  angular intervals. An overall good

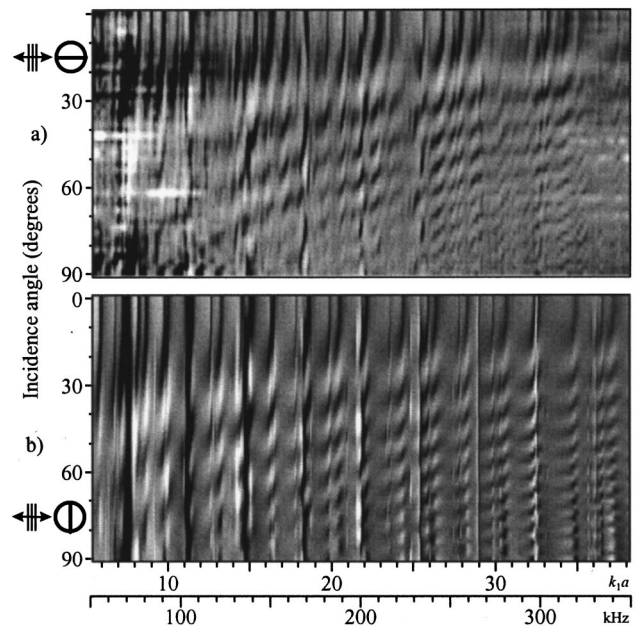


FIG. 4. (a) Experimental and (b) calculated backscattering spectra as functions of the angle  $\varphi$  between the incident wave direction and the position of the internal plate attachments. Amplitude increases from black to white.

agreement is obtained between experimental [Fig. 4(a)] and theoretical results [Fig. 4(b)] calculated from elasticity theory.

To confirm this validation in low frequency range, angular diagrams extracted from Fig. 4 are compared at different frequencies. Some of them are plotted in Figs. 5, 6 and 7, and good agreement is also observed between experimental and theoretical curves, with only light frequency shifts.

Therefore, the model using elasticity theory is validated by the thin shell theory in this low frequency range. Furthermore, the good agreement between models and experiment justifies the use of the pinned attachments of the plate to the shell, that matches with Guo's study.<sup>4</sup> In fact in the experimental object, the plate is soldered to the shell and these attachments do not really correspond to pinned junctions or clamped junctions but a mixture of both.

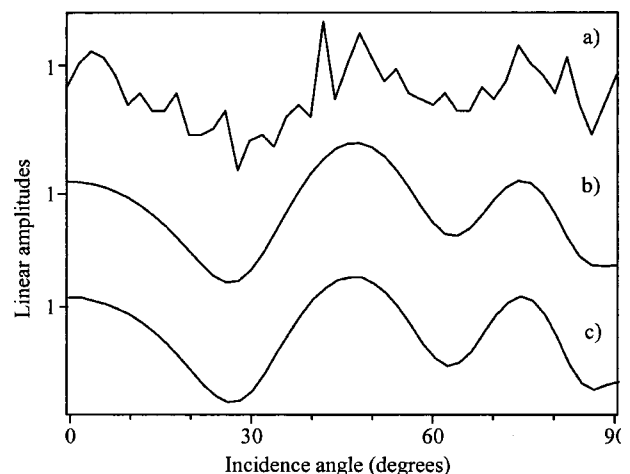


FIG. 5. Angular diagrams. (a) Experimental at  $k_1a=6.68$ . (b) Obtained from model using elasticity theory at  $k_1a=6.78$ . (c) Corresponds to thin shell theory at  $k_1a=6.82$ .

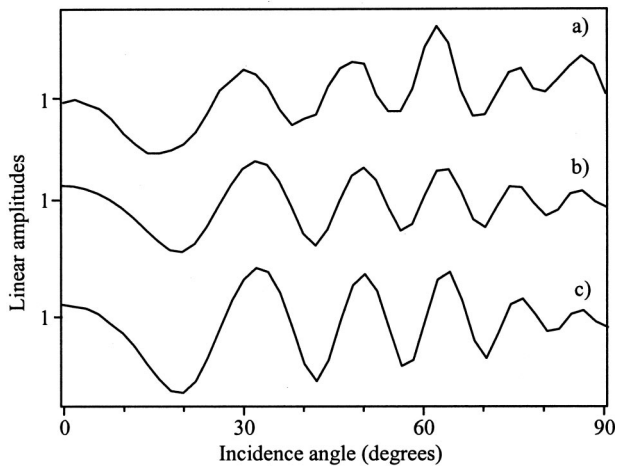


FIG. 6. Angular diagrams. (a) Experimental at  $k_1a=16.05$ . (b) Obtained from model using elasticity theory at  $k_1a=16.20$ . (c) Corresponds to thin shell theory at  $k_1a=16.23$ .

### C. High frequency analysis

The use of thin shell theory is justified as long as the typical acoustic wavelength in fluid and in structure does not exceed the thickness of the shell and the plate: considering the dimensions of the studied object, this frequency range must be below  $k_1a=70-80$ . The use of elasticity theory allows an analysis of the stiffened shell in a high frequency range. In the present model, the descriptions of the shell and the longitudinal contributions of the plate have no frequency limit, but Timoshenko thin plate theory which describes the transversal contributions is limited: the first limitation is analytical and is given in Eq. (A6) in Appendix A. The second limitation is that the theory does not describe wave propagation into the thickness of the plate, thus only antisymmetrical flexural waves are considered.

To carry out analysis in the high frequency range, the same experimental object is used. This time, the transducer is a broad band Panametrics, model V395 with central frequency equal to 2.25 MHz ( $k_1a=245$ ) and its frequency band width is over 0.5–3.5 MHz ( $k_1a=55-380$ ).

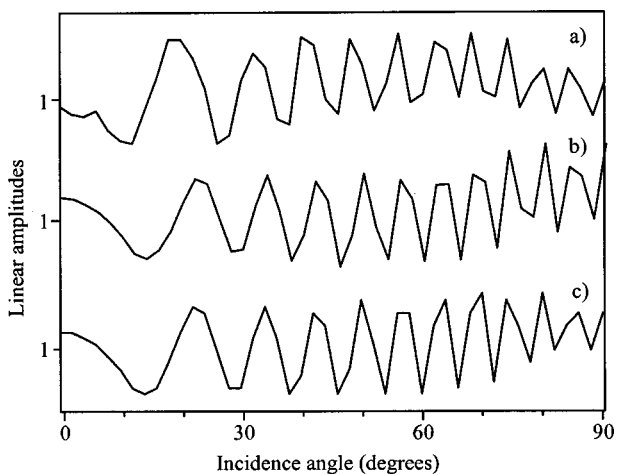


FIG. 7. Angular diagrams. (a) Experimental at  $k_1a=34.09$ . (b) Obtained from model using elasticity theory at  $k_1a=34.05$ . (c) Corresponds to thin shell theory at  $k_1a=34.05$ .

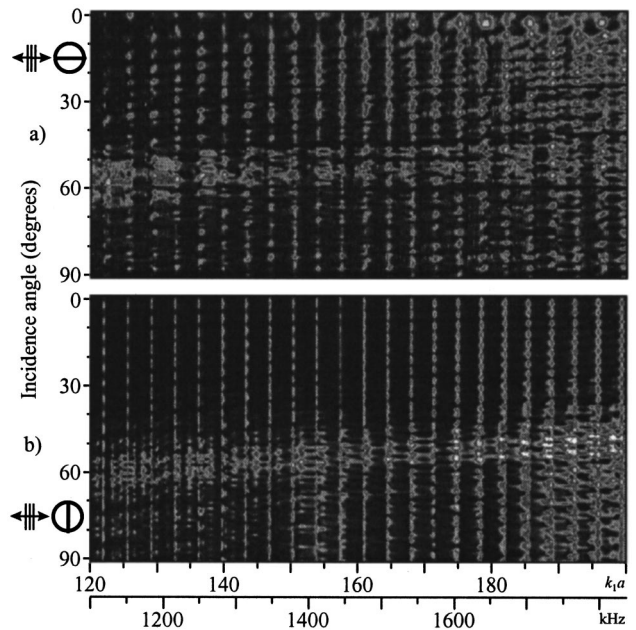


FIG. 8. (a) Experimental and (b) calculated resonance spectra as a function of the angle  $\varphi$  between incident wave direction and position of the internal plate. Amplitude increases from black to white.

In the experimental setup, the pulse monostatic method is used:<sup>19</sup> a pulse is sent to the target and the same transducer receives a backscattered time domain signal which is recorded. To obtain resonance spectra, a Fast Fourier Transform (FFT) is evaluated from the signal recorded without the specular reflection, which is the strongest echo corresponding to the direct reflection from the object. Figure 8 shows resonance spectra recorded from  $\varphi=0^\circ$  to  $90^\circ$  at  $1^\circ$  angular intervals: experimental [Fig. 8(a)] and obtained from model using elasticity theory [Fig. 8(b)]. Computational and experimental steps are both at  $0.1 k_1a$ . This figure shows that experimental results are sufficiently well described by the model using elasticity theory. Vertical dotted lines correspond to  $S_0$  wave resonance peaks. Experimental and theoretical resonance frequencies of the  $S_0$  wave are in a good agreement, and they are close to those of the theoretical unstiffened shell, as shown in Table II. While the monostatic setup for the unstiffened shell does not permit the identification of the modes of resonance, the presence of the internal plate in the stiffened shell makes possible the identification of the modes of the  $S_0$  wave vibrations. Figure 9 presents experimental angular diagram at  $k_1a=129.3$  (bold curve) and theoretical angular diagram, with better precision, at  $k_1a=129.0$  (theoretical angular step is 0.2 degrees). From  $\varphi=0^\circ$  to  $30^\circ$ , six regular periods are observed, which correspond to the 36th mode. In fact, identification of the mode is made possible not by the presence of the plate, but rather by the presence of the junctions themselves, where the  $S_0$  wave is reflected. An analogy between the monostatic method for the stiffened shell and the bistatic method for the unstiffened shell is shown in Ref. 20.

In the frequency window presented in Fig. 8,  $S_0$  wave resonances are the only type of resonances clearly observed. Some variations appear between these peaks, but only when  $\varphi$  is higher than  $50^\circ$ . In fact, there is a global increase of the

TABLE II.  $S_0$  wave resonance frequencies.

Stiffened shell (Experimental)		Resonance frequency				Mode
$ka$	kHz	Stiffened shell (Theoretical)		Unstiffened shell (Theoretical)		
$ka$	kHz	$ka$	kHz	$ka$	kHz	
122.0	1119	121.9	1119	121.9	1119	34
125.6	1152	125.5	1151	125.5	1151	35
129.3	1186	129.0	1184	129.0	1184	36
132.7	1217	132.6	1216	132.6	1216	37
136.4	1251	136.1	1249	136.1	1249	38
140.0	1284	139.7	1281	139.7	1281	39
143.3	1314	143.1	1313	143.2	1314	40
147.0	1348	146.7	1346	146.7	1346	41
150.4	1379	150.3	1379	150.3	1379	42
154.0	1412	153.8	1411	153.8	1411	43
157.5	1445	157.3	1443	157.3	1443	44
160.9	1476	160.8	1476	160.8	1476	45
164.5	1509	164.4	1508	164.4	1508	46
168.1	1542	167.9	1540	167.9	1540	47
171.4	1573	171.3	1572	171.4	1573	48
175.0	1604	174.8	1604	174.9	1605	49
178.5	1638	178.3	1636	178.4	1637	50
182.0	1670	181.8	1668	181.9	1669	51
185.6	1703	185.3	1700	185.4	1701	52
188.9	1733	188.8	1732	188.9	1733	53
192.6	1767	192.3	1764	192.4	1765	54
195.9	1797	195.7	1795	195.8	1796	55
199.5	1831	199.2	1828	199.3	1829	56

amplitude when  $\varphi$  is close to  $50^\circ-60^\circ$  as it can be seen in Fig. 9; this zone corresponds to the angular window where the  $A_0$  wave generated on the shell can be reflected at the junctions without undergoing strong attenuation.<sup>20</sup> These two phenomena are predominant in the resonance spectra, and resonances due to wave propagation into the plate are not clearly detected. Some peaks are theoretically observed but they are thin, their amplitude are small and further, they are not experimentally detected.

### III. CONCLUSIONS

In this paper, a new model describing the acoustic scattering from fluid-loaded stiffened cylindrical shell has been presented. The studied object is a thin cylindrical shell strengthened by an internal diametrical plate. In order to solve the problem, cylindrical shell displacements and con-

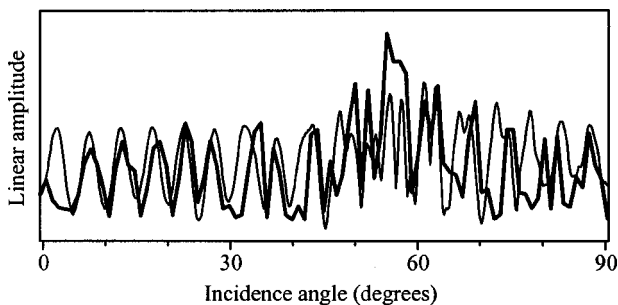


FIG. 9. Angular diagrams: The bold curve is experimental at  $k_1a=129.3$ , the solid curve is obtained from model using elasticity theory at  $k_1a=129.0$ . The 36th mode of  $S_0$  wave is identified from  $\varphi=0^\circ$  to  $30^\circ$ .

straints obtained from elasticity theory are coupled to those of the plate at the junctions, whereas plate vibrations are described by using plate theory.

In the model, the plate is pinned to the shell, thus bending moments are not transmitted. This is verified by both previous studies and in the present case, with the good agreement between theoretical calculations and experimental results. This model is also compared to a model using Timoshenko thin shell theory and then validated in low frequency range ( $k_1a \sim 5-40$ ) with good precision. After that, theory and experiments are studied in the high frequency range ( $k_1a \sim 120-200$ ). Good agreement in the resonance spectra is again observed between the theoretical object and the experimental target. In this frequency window, two phenomena stand out clearly: resonances due to  $S_0$  wave are the only kind of resonances clearly observed and their modes of propagation have been identified. Furthermore, the  $A_0$  wave propagation is detected, largely due to the presence of junctions which allow reflection of the  $A_0$  wave. In the future, time domain signals could be analyzed, and displacements at the shell interfaces could be investigated in order to better understand wave propagation phenomena on such objects.

### APPENDIX A: REACTIONS OF THE PLATE FOR UNITARY DISPLACEMENT

Consider an elastic plate of thickness  $h_P$ , length  $l_P$  in the  $x$ -direction and unitary dimension in the  $z$ -direction, with pinned-pinned attachments. The material is characterized by its density  $\rho_P$ , and its elastic constants  $E$  and  $\nu$ .

The longitudinal vibrations of the plate are described by the equation of motion,<sup>15</sup> where stress  $\sigma$  is related to the longitudinal displacement  $u = ue^{-i\omega t}$ :

$$\frac{Eh_P}{1-\nu^2} \frac{\partial^2 u}{\partial x^2} + \rho_P \omega^2 h_P u = 0. \quad (A1)$$

Solution of Eq. (A1) has the form

$$u = \alpha_1 \cos(k_{LI}x) + \alpha_2 \sin(k_{LI}x), \quad (A2)$$

where  $k_{LI}^2 = \omega^2 \rho_P (1 - \nu^2) / E$ . Coefficients  $\alpha_1$  and  $\alpha_2$  are given by initial conditions. Reactions of the plate extremities for unitary displacement are searched, and longitudinal strains are determined when an extremity of the plate has an unitary displacement. The coefficients  $\alpha_1$  and  $\alpha_2$  are solutions of the following system:

$$u(x=0) = 1, \quad u(x=l_P) = 0. \quad (A3)$$

Expressions of longitudinal strains at the plate extremities for unitary displacement are then

$$F_1^N = F^N(x=0) = -\frac{Eh_P}{1-\nu^2} k_{LI} \cot(k_{LI}l_P),$$

$$F_2^N = F^N(x=l_P) = -\frac{Eh_P}{1-\nu^2} k_{LI} \frac{1}{\sin(k_{LI}l_P)}. \quad (A4)$$

The following notations are used in Eq. (12):

$$F_{k,l}^N = \begin{cases} F_1^N & \text{if } k=l, \\ F_2^N & \text{if } k \neq l. \end{cases}$$



The flexural vibrations of the plate are described by the Timoshenko equation of motion,<sup>16</sup> and the corresponding displacement component is noted  $w$ . Shear deformation and rotatory inertia are not neglected in this equation, which becomes

$$\frac{\partial^4 w}{\partial x^4} + \Omega(\eta + r^2) \frac{\partial^2 w}{\partial x^2} - \Omega(1 - \Omega \eta r^2) w = 0, \quad (\text{A5})$$

where the following notations are used:

$$r^2 = I/h_P, \quad \text{rotatory inertia,}$$

$$\eta = \frac{EI}{1 - \nu^2} \frac{1}{k' h_P G}, \quad \text{shear deformation,}$$

$$\Omega = \omega^2 \rho_P h_P \frac{1 - \nu^2}{EI},$$

with  $I$  the moment of inertia of the cross section,  $G$  the shear modulus and the factor  $k' = 5/6$  representing the distribution of shear stress.

As long as the rotatory inertia and the shear are soft, the following condition is satisfied:

$$\Omega \eta r^2 < 1, \quad (\text{A6})$$

Then the solution of Eq. (A5) takes the form

$$w = \beta_1 \sin(k_{BI}x) + \beta_2 \cos(k_{BI}x) + \beta_3 \sinh(k_{BII}x) + \beta_4 \cosh(k_{BII}x), \quad (\text{A7})$$

where

$$k_{BI} = (\sqrt{\Omega^2(\eta - r^2)^2 + 4\Omega} + \Omega(\eta + r^2))/2,$$

$$k_{BII} = (\sqrt{\Omega^2(\eta - r^2)^2 + 4\Omega} - \Omega(\eta + r^2))/2.$$

Coefficients  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  and  $\beta_4$  are given by initial conditions: transversal strains are determined for an unitary transversal displacement at one extremity, and the plate has pinned–pinned attachments so bending moments at extremities are equal to zero. These coefficients  $\beta_i$  are then solutions of the system:

$$w(x=0) = 1, \quad w(x=l_P) = 0, \quad (\text{A8})$$

$$M(x=0) = 0, \quad M(x=l_P) = 0.$$

Expressions of transversal constraints at the plate extremities for unitary displacement are then

$$\begin{aligned} F_1^T &= F^T(x=0) \\ &= \frac{EI}{(1 - \nu^2)} \frac{1}{1 - \Omega \eta r^2} \\ &\quad \times \frac{\gamma_1 k_{BI}^2 \sin(k_{BI}l_P) \cosh(k_{BII}l_P)}{(k_{BI}^2 + k_{BII}^2) \sin(k_{BI}l_P) \sinh(k_{BII}l_P)} \\ &\quad - \frac{\gamma_2 k_{BII}^2 \cos(k_{BI}l_P) \sinh(k_{BII}l_P)}{\sinh(k_{BII}l_P)}, \end{aligned} \quad (\text{A9})$$

$$\begin{aligned} F_2^T &= F^T(x=l_P) \\ &= \frac{EI}{(1 - \nu^2)} \frac{1}{1 - \Omega \eta r^2} \\ &\quad \times \frac{\gamma_1 k_{BI}^2 \sin(k_{BI}l_P)}{(k_{BI}^2 + k_{BII}^2) \sin(k_{BI}l_P) \sinh(k_{BII}l_P)} \\ &\quad - \frac{\gamma_2 k_{BII}^2 \sinh(k_{BII}l_P)}{\sinh(k_{BII}l_P)}, \end{aligned}$$

with  $\gamma_1 = k_{BII}(k_{BI}^2 - \Omega \eta)$  and  $\gamma_2 = k_{BI}(k_{BII}^2 + \Omega \eta)$ . The following notations are used in Eq. (13):

$$F_{k,l}^T = \begin{cases} F_1^T & \text{if } k=l, \\ F_2^T & \text{if } k \neq l. \end{cases}$$

## APPENDIX B: COEFFICIENTS OF $D_n$ AND $b_n$

$D_n$  is a  $5 \times 5$  matrix obtained from boundary conditions, and its coefficients are

$$D_{11n} = x_T^2 H_n^{(1)}(x_1),$$

$$D_{12n} = (2n^2 - x_T^2) J_n(x_L) - 2x_L J_n'(x_L),$$

$$D_{14n} = 2n(x_T J_n'(x_T) - J_n(x_T)),$$

$$D_{21n} = -x_1 \frac{\rho_S}{\rho_1} H_n^{(1)'}(x_1),$$

$$D_{22n} = x_L J_n'(x_L), \quad D_{24n} = n J_n(x_T),$$

$$D_{31n} = 0, \quad D_{32n} = -2n(x_L J_n'(x_L) - J_n(x_L)),$$

$$D_{34n} = (x_T^2 - 2n^2) J_n(x_T) + 2x_T J_n'(x_T),$$

$$D_{41n} = 0, \quad D_{42n} = (2n^2 - y_T^2) J_n(y_L) - 2y_L J_n'(y_L),$$

$$D_{44n} = 2n(y_T J_n'(y_T) - J_n(y_T)),$$

$$D_{51n} = 0, \quad D_{52n} = -2n(y_L J_n'(y_L) - J_n(y_L)),$$

$$D_{54n} = (y_T^2 - 2n^2) J_n(y_T) + 2y_T J_n'(y_T)$$

and  $D_{i3n}$  and  $D_{i5n}$  are given by the substitutions noted below. Both  $D_{i3n}$  and  $D_{i5n}$  have the same form as, respectively,  $D_{i2n}$  and  $D_{i4n}$  when  $J_n$  replaced by  $N_n$  and  $J_n'$  replaced by  $N_n'$ .

Coefficients  $b_n$  correspond to the left side term of Eq. (14), and are decomposed in three parts:

$$b_n = b_n^0 + b_n^N + b_n^T$$

with

$$b_{1n}^0 = -x_1^2 J_n(x_1), \quad b_{2n}^0 = x_1 \frac{\rho_S}{\rho_1} J_n'(x_1),$$

$$b_{3n}^0 = b_{4n}^0 = b_{5n}^0 = 0,$$

$$b_{1n}^N = b_{2n}^N = b_{3n}^N = b_{5n}^N = 0,$$

$$b_{4n}^N = i^{-n} (f_1^N e^{-in\theta_1} + f_2^N e^{-in\theta_2}),$$

$$b_{1n}^T = b_{2n}^T = b_{3n}^T = b_{4n}^T = 0,$$

$$b_{5n}^T = i^{-n} i (f_1^T e^{-in\theta_1} + f_2^T e^{-in\theta_2}),$$

where the following notations are used:

$$x_j = k_j a, \quad j = 1, L, T,$$

$$y_j = k_j b, \quad j = 1, L, T,$$

$$f_1^N = \frac{y_T^2}{2\pi b} (F_1^N u_r(b, \theta_1) + F_2^N u_r(b, \theta_2)),$$

$$f_2^N = \frac{y_T^2}{2\pi b} (F_1^N u_r(b, \theta_2) + F_2^N u_r(b, \theta_1)),$$

$$f_1^T = \frac{y_T^2}{2\pi b} (F_1^T u_\theta(b, \theta_1) + F_2^T u_\theta(b, \theta_2)),$$

$$f_2^T = \frac{y_T^2}{2\pi b} (F_1^T u_\theta(b, \theta_2) + F_2^T u_\theta(b, \theta_1)).$$

### APPENDIX C: NORMAL AND TANGENTIAL DISPLACEMENTS AT THE JUNCTIONS $\theta = \theta_1$ AND $\theta = \theta_2$

Equation (17) is developed as a modal summation of  $\mathbf{a}_n$  at  $r = b$ :

$$\begin{aligned} u_r(\theta) = & \frac{1}{\rho_S \omega^2 b} \sum_{n=-\infty}^{n=+\infty} i^n y_L (J_n'(y_L) a_{2n}) \\ & + N_n'(y_L) a_{3n} e^{in\theta} + \frac{1}{\rho_S \omega^2 b} \\ & \times \sum_{n=-\infty}^{n=+\infty} i^n n (J_n(y_T) a_{4n} + N_n(y_T) a_{5n}) e^{in\theta}. \end{aligned} \quad (C1)$$

Coefficient  $\mathbf{a}_n$  is decomposed in three parts, which are the solutions for each contribution as it has been shown in Eqs. (15) and (19):

$$\begin{aligned} u_r(\theta) = & u_r^0(\theta) \\ & + \frac{1}{\rho_S \omega^2 b} \sum_{n=-\infty}^{n=+\infty} i^n y_L (J_n'(y_L) a_{2n}^N + N_n'(y_L) a_{3n}^N) e^{in\theta} \\ & + \frac{1}{\rho_S \omega^2 b} \sum_{n=-\infty}^{n=+\infty} i^n n (J_n(y_T) a_{4n}^N + N_n(y_T) a_{5n}^N) e^{in\theta}. \end{aligned} \quad (C2)$$

The only unknown coefficients are now  $\mathbf{a}_n^N$ , which are solutions of Eq. (14). Equation (21) can be written as

$$a_{in}^N = \frac{|D_n^{N[i]*}|}{|D_n|} = (f_1^N e^{-in\theta_1} + f_2^N e^{-in\theta_2}) \frac{|D_n^{N[i]*}|}{|D_n|}, \quad (C3)$$

where  $|D_n^{N[i]*}|$  is the determinant of  $D_n$  in which the  $i$ -column is replaced by  $\mathbf{b}_n^{N*}$  vector instead of  $\mathbf{b}_n^N$  vector, with:

$$\mathbf{b}_n^N = (f_1^N e^{-in\theta_1} + f_2^N e^{-in\theta_2}) \mathbf{b}_n^{N*}.$$

The coefficients  $a_{in}^N$  expressed in Eq. (C3) are then replaced in Eq. (C2). Relations  $f_1^N$  and  $f_2^N$  which contain the displacement terms  $u_r(b, \theta_k)$  are independent of  $n$ , and can be extracted from the summation. The following system is then obtained:

$$\begin{bmatrix} u_r(\theta_1) \\ u_r(\theta_2) \end{bmatrix} = \begin{bmatrix} u_r^0(\theta_1) \\ u_r^0(\theta_2) \end{bmatrix} + \alpha \begin{bmatrix} S_1^N & S_2^N \\ S_2^N & S_1^N \end{bmatrix} \begin{bmatrix} F_1^N & F_2^N \\ F_2^N & F_1^N \end{bmatrix} \begin{bmatrix} u_r(\theta_1) \\ u_r(\theta_2) \end{bmatrix}, \quad (C4)$$

where  $\alpha = (1/\rho_S \omega^2 b)(y_T^2/2\pi b)$ ,

$$\begin{aligned} S_1^N = & \sum_{n=-\infty}^{n=+\infty} i^n y_L (J_n'(y_L) a_{2n}^{N*} + N_n'(y_L) a_{3n}^{N*}) \\ & + \sum_{n=-\infty}^{n=+\infty} i^n n (J_n(y_T) a_{4n}^{N*} + N_n(y_T) a_{5n}^{N*}), \\ S_2^N = & \sum_{n=-\infty}^{n=+\infty} (-i)^n y_L (J_n'(y_L) a_{2n}^{N*} + N_n'(y_L) a_{3n}^{N*}) \\ & + \sum_{n=-\infty}^{n=+\infty} (-i)^n n (J_n(y_T) a_{4n}^{N*} + N_n(y_T) a_{5n}^{N*}), \end{aligned}$$

and

$$a_{in}^{N*} = \frac{|D_n^{N[i]*}|}{|D_n|}.$$

The expressions of the tangential displacements are obtained in a similar way, and Eq. (18) is also developed as a modal summation of  $\mathbf{a}_n$  at  $r = b$ :

$$\begin{aligned} u_\theta(\theta) = & \frac{1}{\rho_S \omega^2 b} \sum_{n=-\infty}^{n=+\infty} i^n n (J_n(y_L) a_{2n} + N_n(y_L) a_{3n}) e^{in\theta} \\ & - \frac{1}{\rho_S \omega^2 b} \sum_{n=-\infty}^{n=+\infty} \frac{i^n}{i} n (J_n'(y_T) a_{4n} + N_n'(y_T) a_{5n}) e^{in\theta}. \end{aligned} \quad (C5)$$

When  $\mathbf{a}_n$  are decomposed in three parts, Eq. (C5) becomes

$$\begin{aligned} u_\theta(\theta) = & u_\theta^0(\theta) + \frac{1}{\rho_S \omega^2 b} \sum_{n=-\infty}^{n=+\infty} i^n n (J_n(y_L) a_{2n}^T \\ & + N_n(y_L) a_{3n}^T) e^{in\theta} \\ & - \frac{1}{\rho_S \omega^2 b} \sum_{n=-\infty}^{n=+\infty} \frac{i^n}{i} n (J_n'(y_T) a_{4n}^T + N_n'(y_T) a_{5n}^T) e^{in\theta}. \end{aligned} \quad (C6)$$

The only unknown coefficients are now  $\mathbf{a}_n^T$ , which are solutions of Eq. (14):

$$a_{in}^T = \frac{|D_n^{T[i]}|}{|D_n|} = (f_1^T e^{-in\theta_1} + f_2^T e^{-in\theta_2}) \frac{|D_n^{T[i]*}|}{|D_n|}, \quad (C7)$$

where  $|D_n^{T[i]*}|$  is the determinant of  $D_n$  where the  $i$ -column is replaced by the  $b_n^{T*}$  vector, with

$$b_n^T = (f_1^T e^{-in\theta_1} + f_2^T e^{-in\theta_2}) b_n^{T*}.$$

The coefficients  $a_{in}^T$  expressed in Eq. (C7) are then replaced in Eq. (C6). Relations  $f_1^T$  and  $f_2^T$  which contain the displacement terms  $u_\theta(b, \theta_k)$  are independent of  $n$ , and can be extracted from the summation. The following system is then obtained:

$$\begin{bmatrix} u_\theta(\theta_1) \\ u_\theta(\theta_2) \end{bmatrix} = \begin{bmatrix} u_\theta^0(\theta_1) \\ u_\theta^0(\theta_2) \end{bmatrix} - \alpha \begin{bmatrix} S_1^T & S_2^T \\ S_2^T & S_1^T \end{bmatrix} \begin{bmatrix} F_1^T & F_2^T \\ F_2^T & F_1^T \end{bmatrix} \begin{bmatrix} u_\theta(\theta_1) \\ u_\theta(\theta_2) \end{bmatrix} \quad (C8)$$

where

$$\begin{aligned} S_1^T &= \sum_{n=-\infty}^{n=+\infty} i^n n (J_n(y_L) a_{2n}^{T*} + N_n(y_L) a_{3n}^{T*}) \\ &\quad + \sum_{n=-\infty}^{n=+\infty} i^n y_T (J_n'(y_T) a_{4n}^{T*} + N_n'(y_T) a_{5n}^{T*}), \\ S_2^T &= \sum_{n=-\infty}^{n=+\infty} (-i)^n n (J_n(y_L) a_{2n}^{T*} + N_n(y_L) a_{3n}^{T*}) \\ &\quad + \sum_{n=-\infty}^{n=+\infty} (-i)^n y_T (J_n'(y_T) a_{4n}^{T*} + N_n'(y_T) a_{5n}^{T*}), \end{aligned}$$

and

$$a_{in}^{T*} = \frac{|D_n^{T[i]*}|}{|D_n|}.$$

- <sup>1</sup>A. Klauson and J. Metsaveer, "Sound scattering by a lengthwise reinforced cylindrical shell," *Sov. Phys. Acoust.* **35**, 42–47 (1989).  
<sup>2</sup>Y. P. Guo, "Sound scattering from an internally loaded cylindrical shell," *J. Acoust. Soc. Am.* **91**, 926–938 (1992).  
<sup>3</sup>Y. P. Guo, "Sound scattering from cylindrical shells with internal elastic plates," *J. Acoust. Soc. Am.* **93**, 1936–1946 (1993).

- <sup>4</sup>Y. P. Guo, "Acoustic radiation from cylindrical shells due to internal forcing," *J. Acoust. Soc. Am.* **99**, 1495–1505 (1996).  
<sup>5</sup>Y. P. Guo, "Acoustic scattering from cylindrical shells with deck-type internal plate at oblique incidence," *J. Acoust. Soc. Am.* **99**, 2701–2713 (1996).  
<sup>6</sup>A. Klauson and J. Metsaveer, "Sound scattering by a cylindrical shell reinforced by lengthwise ribs and walls," *J. Acoust. Soc. Am.* **91**, 1834–1843 (1992).  
<sup>7</sup>R. C. Loftman and D. B. Bliss, "Analytical/numerical matching for efficient calculation of scattering from cylindrical shells with lengthwise constraints," *J. Acoust. Soc. Am.* **103**, 1885–1896 (1998).  
<sup>8</sup>A. Klauson, G. Maze, and J. Metsaveer, "Acoustic scattering by submerged cylindrical shell stiffened by an internal lengthwise rib," *J. Acoust. Soc. Am.* **96**, 1575–1581 (1994).  
<sup>9</sup>A. Klauson, J. Metsaveer, D. Décultot, G. Maze, and J. Ripoche, "Identification of the resonances of a cylindrical shell stiffened by an internal lengthwise rib," *J. Acoust. Soc. Am.* **100**, 3135–3143 (1996).  
<sup>10</sup>A. G. Pathak and P. R. Stepanishen, "Acoustic harmonic radiation from fluid-loaded infinite cylindrical elastic shells using elasticity theory," *J. Acoust. Soc. Am.* **96**, 573–582 (1994).  
<sup>11</sup>J. J. Faran, Jr., "Sound scattering by solid cylinders and spheres," *J. Acoust. Soc. Am.* **23**, 405–418 (1951).  
<sup>12</sup>L. Flax, L. R. Dragonette, and H. Überall, "Theory of elastic resonance excitation by sound scattering," *J. Acoust. Soc. Am.* **63**, 723–731 (1978).  
<sup>13</sup>A. Baillard, J.-M. Conoir, D. Décultot, G. Maze, A. Klauson, and J. Metsaveer, "Sound scattering by a fluid-loaded cylindrical shell with a diametral wall," *Transport Noise and Vibration 98*, Tallinn, Estonia, Conference Proceedings (1998), pp. 251–254.  
<sup>14</sup>A. Klauson, J. Metsaveer, A. Baillard, D. Décultot, and G. Maze, "Acoustic scattering by a fluid loaded cylindrical shell with an internal plate," *J. Acoust. Soc. Am.* **99**, 2544 (1996).  
<sup>15</sup>M. C. Junger and D. Feit, *Sound, Structures, and their Interaction* (Massachusetts Institute of Technology, Cambridge, MA, 1972).  
<sup>16</sup>W. Soedel, *Vibrations of Shells and Plates* (Marcel Dekker, New York, 1993).  
<sup>17</sup>G. Maze and J. Ripoche, "Méthode d'Isolement et d'Identification des résonances (M.I.I.R.) de cylindres et de tubes soumis à une onde acoustique plane dans l'eau," *Rev. Phys. Appl.* **18**, 319–326 (1983).  
<sup>18</sup>G. Maze, "Acoustic scattering from submerged cylinders. M.I.R. Im/Re: Experimental and theoretical study," *J. Acoust. Soc. Am.* **89**, 2559–2566 (1991).  
<sup>19</sup>P. Pareige, P. Rembert, G. Maze, and J. Ripoche, "Signature spectrale et identification des résonances d'objets immergés par une méthode impulsionnelle numérisée (M.I.I.R. impulsionnelle)," *J. Acoust.* **3**, 101–106 (1990).  
<sup>20</sup>A. Baillard, J.-M. Conoir, D. Décultot, G. Maze, A. Klauson, and J. Metsaveer, "A<sub>0</sub> wave in a submerged stiffened cylindrical shell," *1999 ASME Design Engineering Technical Conference*, Las Vegas, Nevada, Conference Proceedings (1999).

# Refined multiload method for measuring acoustical source characteristics of an intake or exhaust system

Seung-Ho Jang and Jeong-Guon Ih

*Center for Noise and Vibration Control, Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Science Town, Taejon 305-701, Korea*

(Received 17 November 1999; accepted for publication 20 February 2000)

The one-port source characteristics in a duct system, viz., source impedance and strength, can be determined by using the four-load method. In this paper, to avoid the instability problem of the conventional four-load method, a new formulation for the multiload method has been proposed, which employs an error function based on the linear, time-invariant source model. It is shown that the method is less sensitive to input errors compared to the previous methods. For a 10% input error, the proposed method yields a relative error in the source resistance that is about 1/100 times smaller than for the conventional method. The effectiveness of the present method is demonstrated by two test examples, a loudspeaker and a blower, each operating in a duct. It is observed that the conventional and least-squares methods result in large errors, whereas the present method yields far better agreement with the actual source parameters, as measured by the direct method. The present method is then used to obtain the source parameters on the exhaust side of an operating internal combustion engine. The radiated sound spectrum from the exhaust opening is predicted by using the measured source parameters and the calculated result agrees very well with the measured one. © 2000 Acoustical Society of America. [S0001-4966(00)00906-1]

PACS numbers: 43.50.Gf, 43.20.Mv, 43.20.Ye [MRS]

## INTRODUCTION

The acoustic performance of a ducted silencing system depends upon the acoustic characteristics of the source as well as that of the silencer itself. Examples include blowers, HVAC duct systems, and the intake and exhaust systems of typical internal combustion engines. For a source located at one side of the duct, the acoustic one-port source model can be used by assuming plane wave propagation and a linear, time-invariant system. In this case, the source can be characterized by two complex parameters, viz., the source strength and the source impedance. Without having these source parameters, one cannot predict the radiated sound pressure level from the duct opening. For the prediction of insertion loss of any silencing system, the information on the source impedance is essentially required.

The four-load method has been proposed as one of indirect measurement techniques for estimating the characteristics of a source in a duct.<sup>1</sup> The advantages of the method are that there is no need for a secondary source, because the response is measured from the actual test source, and the sound power spectrum measured by one microphone outside the duct is the only measurement required. These features make the four-load method preferable to the so-called direct method, especially in the case of an internal combustion engine source. This is a very favorable feature when the pressure measurement is done in the duct where the turbulent flow and the hot and corrosive gas are involved.<sup>2</sup> In addition, the four-load method does not need a reference signal that is always required for using the two-load method<sup>3</sup> as an indirect measurement method. Direct methods include the standing wave ratio method using a traversing microphone,<sup>4,5</sup> the two-microphone method,<sup>6,7</sup> and the multiple microphone method.<sup>8,9</sup> In these techniques, the measurement should be

made inside the duct and a secondary external source is required which can emit a far higher level sound than the tested source. However, the four-load method has been reported to be very sensitive to errors in the input data and this error sensitiveness sometimes give erroneous results.<sup>10-12</sup> Desmons and Hardy<sup>11</sup> suggested a least-squares method in order to overcome the problem. However, the method requires complex pressure measurements that can be more influenced from the effects of reflecting surfaces than power spectrum measurements with one microphone. In addition, it is very hard to find the phase reference signal in some machines.<sup>12</sup> In order to refine the conventional four-load method, Bodén<sup>12</sup> proposed a direct least-squares multiload method that is less sensitive to the input errors, but which still results in appreciable errors.

In this paper, the conventional four-load method is modified to overcome the instability problem. The relationship between source properties and load properties is expressed by the convective acoustic parameters.<sup>13</sup> The method uses an error function that is based on the linear, time-invariant source model. For demonstrating the effectiveness of the present method, the source parameters are obtained for a loudspeaker, a centrifugal blower, and the exhaust system of an internal combustion engine.

## I. THEORY

### A. Source model

Assuming linear and time-invariant, an acoustic source in a duct with mean flow can be represented by using the electroacoustic analogy as shown in Fig. 1. In the frequency

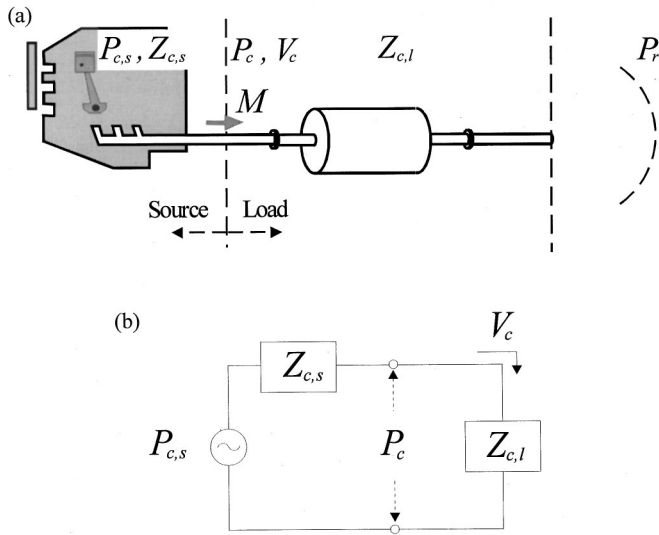


FIG. 1. Source-load model. (a) Acoustic source-load system in the duct, (b) analogous electroacoustic circuit.

domain, the equation for representing the source-load interaction can be written using the convective source parameters as<sup>13</sup>

$$P_c = P_{c,s} Z_{c,l} / (Z_{c,l} + Z_{c,s}), \quad (1)$$

where  $P_{c,s}$  and  $Z_{c,s}$  are the convective source pressure and the convective source impedance, respectively. In Fig. 1, the convective pressure  $P_c$ , convective volume velocity  $V_c$  at the source-load junction, and convective impedance  $Z_{c,l}$  of the load are defined as

$$P_c = P(1 + MZ_0/Z_l), \quad (2a)$$

$$V_c = V + MP/Z_0, \quad (2b)$$

$$Z_{c,l} = (Z_l + MZ_0)/(1 + MZ_l/Z_0), \quad (2c)$$

where  $M$  is the flow Mach number and  $Z_0$  is the characteristic impedance of the medium. Here,  $P$ ,  $V$ , and  $Z_l$  are the acoustic pressure, the acoustic volume velocity at the source-load junction, and the acoustic impedance of the load, respectively. The acoustic pressure  $P$  at the source-load interface can be obtained from the Thévenin theorem. By using the convective parameters, one can easily predict the acoustic performances of actual duct systems such as the radiated acoustic power and the insertion loss of the muffler in the presence of mean flow. From Eq. (1), the radiated acoustic power can be derived as

$$W = \frac{|P_c|^2}{2} \operatorname{Re} \left( \frac{1}{Z_{c,l}} \right) = \frac{|P_{c,s}|^2 \operatorname{Re}(Z_{c,l})}{2|Z_{c,l} + Z_{c,s}|^2}, \quad (3)$$

where  $\operatorname{Re}(x)$  denotes the real part of a complex quantity  $x$ . One can see that the real part of the convective load impedance should be positive, in order to have a positive sound power. In the same context, the real part of the convective source impedance describing the passive properties of the source is expected to be positive. In terms of the acoustic state variables, this condition can be written as

$$\begin{aligned} \operatorname{Re}(Z_{c,s}/Z_0) &= \operatorname{Re} \left( \frac{Z_s/Z_0 - M}{1 - MZ_s/Z_0} \right) \\ &= \frac{\operatorname{Re}(Z_s/Z_0)(1 + M^2) - M(|Z_s|^2/Z_0^2 + 1)}{|1 - MZ_s/Z_0|^2} > 0. \end{aligned} \quad (4)$$

This inequality gives the following condition for the range of convective source:

$$\operatorname{Re}(Z_s/Z_0)(1 + M^2) - M(|Z_s|^2/Z_0^2 + 1) > 0. \quad (5)$$

If one differentiates Eq. (3) with respect to  $Z_{c,l}$ , one can derive the following equation:

$$dW = |P_{c,s}|^2 \frac{1}{|Z_{c,s} + Z_{c,l}|^2} \operatorname{Re} \left[ \frac{Z_{c,s} - Z_{c,l}^*}{4(Z_{c,s} + Z_{c,l})} dZ_{c,l} \right]. \quad (6)$$

It is noted that the radiated acoustic power has its extreme value only when  $Z_{c,l} = Z_{c,s}^*$ , where \* means the complex conjugate. The condition can be rewritten by using the acoustic parameters as

$$Z_l = \frac{2MZ_0 - Z_s^*(1 + M^2)}{2MZ_s^*/Z_0 - 1 - M^2}, \quad (7)$$

which corresponds to the result of the approach using the acoustic state variables.<sup>14</sup> If Eq. (6) is differentiated once more, one can get

$$d^2W = -\frac{1}{16} |P_{c,s}|^2 \frac{|dZ_{c,l}|^2}{\operatorname{Re}(Z_{c,s})|Z_{c,s} + Z_{c,l}|^2} < 0, \quad (8)$$

which is the condition for maximum radiation of sound power. Figure 2 shows calculated results of the variation of the radiated sound power by changing the load impedance. One can observe that the maximum sound power is radiated when the load impedance equals the complex conjugate of the source impedance. However, it is unfortunately obvious that there is no local minimum of sound power, which is of concern in the design of silencers.

## B. Refined new multiloading method

The original four-load method developed by Prasad<sup>1</sup> for the evaluation of the source properties has an instability problem, which needs to be overcome in order to obtain accurate results. As mentioned earlier, Bodén<sup>12</sup> proposed a direct *least-squares* method that is less sensitive to input errors, but still results in appreciable errors. Hereafter, Prasad's original four-load method will be called the "*conventional method*" and Bodén's numerical fitting method will be called the "*direct least-squares method*" in this paper.

In order to avoid the instability problems of the previous algorithms, one can start from the modification of Eq. (1) as

$$|P_{c,s}|^2 = |P_{c,n}|^2 \frac{|Z_{c,n} + Z_{c,s}|^2}{|Z_{c,n}|^2} \quad (n = 1, 2, 3, \dots, N), \quad (9)$$

where  $N$  is the number of applied loads. It is noted that the equation deals with the energy quantity rather than the physical parameter itself, as employed in the conventional method. From Eq. (9), motivated by the fact that  $|P_{c,s}|^2$

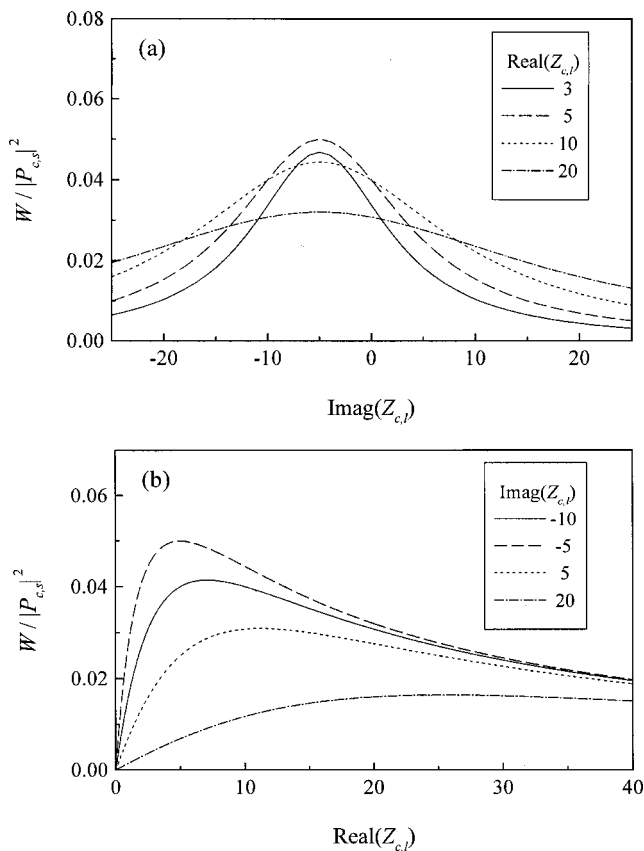


FIG. 2. Change of radiated sound power due to the variation of load impedance ( $Z_{c,s} = 5 + 5i$ ). (a) Change of load resistance, (b) change of load reactance.

should be invariable no matter what load is connected, an error function  $F(Z)$  can be defined as

$$F(Z) = \sum_{m < n} \left\{ \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} |Z_{c,m} + Z|^2 - \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} |Z_{c,n} + Z|^2 \right\}^2 \quad (m, n = 1, 2, 3, \dots, N), \quad (10)$$

where  $P_{c,n}, Z_{c,n}$  denote the convective acoustic pressure and impedance of the  $n$ th load, respectively. The choice of the  $m$ th and  $n$ th load pair is based on every possible combination of load sets, thus enabling the maximized collection of information. The source impedance  $Z_{c,s}$  is the value of  $Z$  which satisfies  $F(Z) = 0$ , in principle. However, any actual measured input data would be subject to small errors and, in this case, the value of  $Z$  minimizing  $F$  is chosen as  $Z_{c,s}$ . Differentiating  $F$  with respect to  $Z$ , one can obtain

$$dF = \sum_{m < n} 2 \left\{ \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} |Z_{c,m} + Z|^2 - \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} |Z_{c,n} + Z|^2 \right\} \times \text{Re} \left[ \left( \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} (Z_{c,m}^* + Z^*) - \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} (Z_{c,n}^* + Z^*) \right) dZ \right]. \quad (11)$$

Because  $dZ$  is arbitrary, the necessary condition for the minimum  $F$  is  $dF = 0$ , or

$$G(Z) \equiv \sum_{m < n} \left\{ \frac{|P_{c,m}|^4}{|Z_{c,m}|^4} |Z_{c,m} + Z|^2 (Z_{c,m} + Z) - \frac{|P_{c,m}|^2 |P_{c,n}|^2}{|Z_{c,m}|^2 |Z_{c,n}|^2} |Z_{c,n} + Z|^2 (Z_{c,m} + Z) + \frac{|P_{c,n}|^4}{|Z_{c,n}|^4} |Z_{c,n} + Z|^2 (Z_{c,n} + Z) - \frac{|P_{c,m}|^2 |P_{c,n}|^2}{|Z_{c,m}|^2 |Z_{c,n}|^2} |Z_{c,m} + Z|^2 (Z_{c,n} + Z) \right\} = 0. \quad (12)$$

Equation (12) can be arranged into separate real and imaginary parts to give a pair of nonlinear cubic equations in the two variables of resistance and reactance. Iteration methods, such as the Newton–Raphson method, are not appropriate for solving this problem because the calculated results depend on the initial starting values. In this study, the polar representation of the impedances is employed in order to get the solution suitably. After substituting  $re^{i\theta}$  for  $Z$ , one can obtain the two cubic polynomial equations (see Appendix), so that the angle  $\theta$  and the corresponding magnitude  $r$  satisfying the equations can be found in an iterative manner. Since Eq. (12) is the necessary condition for  $F$  having a minimum value, the value of  $Z$  minimizing  $F$  among several solutions should be chosen as the source impedance. After obtaining the source impedance, the source strength can be easily obtained from Eq. (1) as

$$|P_{c,s}| = \frac{\sum_m |P_{c,m}| \left| \frac{Z_{c,m}}{Z_{c,m} + Z_{c,s}} \right|}{\sum_m \left| \frac{Z_{c,m}}{Z_{c,m} + Z_{c,s}} \right|^2}. \quad (13)$$

The present method employs the squared quantities as in the least-squares method,<sup>12</sup> but the two multiloading methods are basically different from each other. The present method is based on the error function of Eq. (10), whereas the least-squares method is based on a direct numerical fit of the data to the nonlinear model by using the least-squares method.

## II. SIMULATION OF SENSITIVITY TO INPUT ERROR

In actual measurements, the measured quantities are subject to some error, no matter how small. The effects of errors in the input data on the calculated source parameters have been studied numerically. If the source parameters are determined, the pressures at the source–load interface for the different loads can be calculated using Eq. (1) and they are utilized as the input data to perform the subsequent simulations.

Normalized source resistance of practical fluid machines is usually within the range of 0–2 and the normalized source reactance is in the  $\pm 2$  range.<sup>2,5,12,15–17</sup> Ten values with the same interval were chosen from these data ranges and every possible combination of the resistance and reactance was used in the error analysis. The analyzed data were then averaged. The same source strength was used for all calculations. The load impedance data were normalized to the characteristic impedance. The normalized load resistance was selected from the 0.1–0.2 range and the normalized load

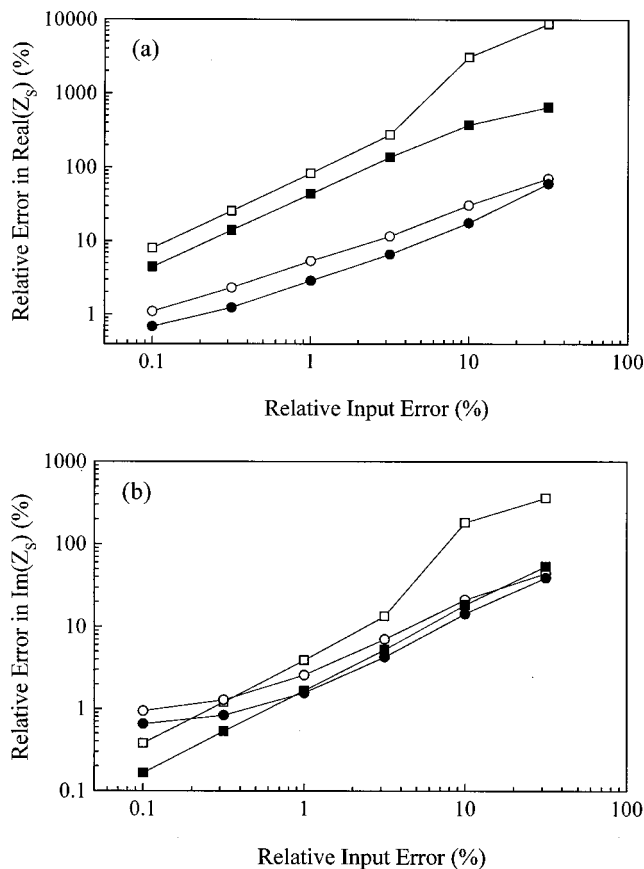


FIG. 3. Sensitivity of the source impedance to the relative input error:  $\square$ ,  $\blacksquare$  denote the conventional method and  $\circ$ ,  $\bullet$  denote the present method; data points with blank symbols are obtained by using four loads and solid symbols by using eight loads. (a) Resistance, (b) reactance.

reactance from the  $\pm 2$  range, which is usual for open pipes. First, the radiated pressure spectrum was calculated by using the known source impedance, source strength, and load impedance. Second, in order to create a required relative input error, a random noise with uniform distribution was added to the radiated sound spectrum. Third, the load impedance data and the corresponding radiated pressure spectra with input noise were used for the conventional and the present multiloading method. Here, the conventional four-load method was extended to use the multiloading data. Finally, by comparison with the *a priori* known source impedance, the relative error was calculated. As can be seen in Fig. 3(a), the present method yields far smaller errors in the resistance value than the conventional method. The effectiveness of the present method is particularly superior to the conventional method when the relative input error is large. For a 10% input error, the proposed method yields a relative error in the source resistance that is about 1/100 times smaller than for the conventional method. In particular, when an overdetermined set of eight loads is used, the present method yields a definite improvement in estimating the source resistance. Compared to the resistance case, a somewhat different trend can be seen in the error sensitivity of the estimated source reactance as depicted in Fig. 3(b). Using only four loads, the present method provides a better estimation of source reactance in the presence of input errors than the conventional method. For a 10% input error, the proposed method yields a relative

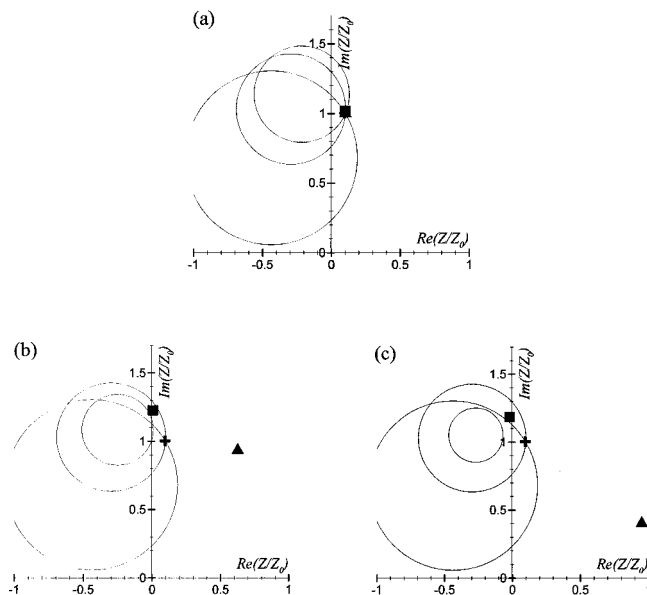


FIG. 4. The effect of errors in the input data  $|P_{c,i}|$  on the calculated source impedance using four loads.  $+$ , true value ( $Z_{c,s}/Z_0=0.1+1i$ );  $\blacktriangle$ , direct least-squares method;  $\blacksquare$ , present method. (a) No input error, (b) 10% input errors, (c) 20% input errors.

error in the source reactance that is about 1/10 times smaller than for the conventional method. Although one can find a range of input error in which the conventional method provides better results than the present one, the input error in that range is quite small and thus impractical for most real measurement situations. When overdetermined eight loads are used, the present method can attain marginal improvements of error sensitivity to the conventional method.

Figure 4 illustrates the graphical representation of the source impedance calculated using four loads by three different methods: conventional four-load method,<sup>1</sup> direct least-squares method,<sup>12</sup> and present method. Using Eq. (1), one can obtain the three equations of circles.<sup>1,10</sup> By eliminating the quadratic terms in the equations, the conventional four-load method calculates the source impedance of which the real and imaginary parts correspond to the coordinates of a point where the three circles meet together. Under the condition of no errors in the input data, one can observe that the three circles meet at one point and the correct results are expected for any method. However, in the cases with some errors included in the input data  $|P_{c,m}|$ , the three circles do not intersect at a common point and the original four-load method fails to find the proper solution. On the other hand, the present and the direct least-squares methods yield equally good results when the input error is quite small. However, it is noted that the present method is far less sensitive to the relatively large errors in the measured input data than the direct least-squares method. In implementing the conventional method, one should bear in mind that there might exist a pair of circles that do not intersect with each other as shown in Fig. 4(b) and (c). The source impedance calculated using eight loads is shown graphically in Fig. 5. It can be observed that the circles that are utilized by the extended conventional method can hardly be exploited to get accurate results. This is because the effect of random errors on the

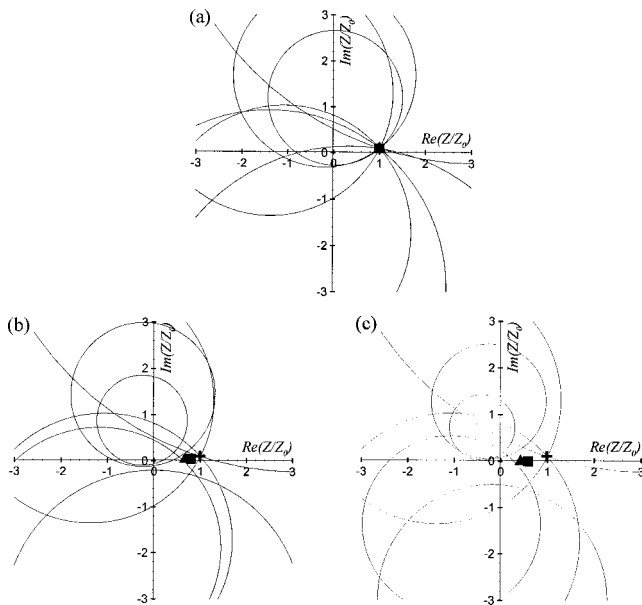


FIG. 5. The effect of errors in the input data  $|P_c|$  on the calculated source impedance using eight loads. +, true value ( $Z_{c,s}/Z_0 = 1 + 0.1i$ ); ▲, direct least-squares method; ■, present method. (a) No input error, (b) 10% input errors, (c) 20% input errors.

results by the least-squares method and the present method is not severe, whereas the conventional method can be singular and very easily influenced by the random error. In this highly overdetermined case, the present method seems to be slightly more accurate than the direct least-squares method.

In Fig. 6, a simulated result is shown for the effect of number of loads on the error sensitivity of the source impedance to input error when the present method is utilized. Starting from the four loads which were selected arbitrarily from the usual range of tube impedance values, other loads are added one by one to the formerly chosen loads. Although a definite improving trend in sensitivity can be observed with an increase of the number of loads, the amount of decrease in relative error is small for each increment of number of loads. However, for a 10% input error and using five loads, the relative error in the source resistance is expected to be about 10% smaller than the case using four loads. The reduction of error in reactance in the same condition is about 5%. It seems that the improvement in source impedance estimation is marginal for using more than five or six loads. The fluctuation of data with increase of number of loads is within  $\pm 0.5\%$  and it is thought that this is caused from the arbitrariness in choosing the loads and the statistical variation in the subsequent simulation.

### III. EXPERIMENTAL RESULTS

Two acoustic sources were taken as examples for comparing the performance in estimating the source properties. The most simple, but easily controllable, test example is the loudspeaker, because there are no mean flow and temperature gradient effects. In addition, a centrifugal blower operating at a constant rotating speed was used as an example of a practical fluid machine having actual fluid flow. For comparative purposes, four different measurement techniques were employed in order to obtain the source parameters. The

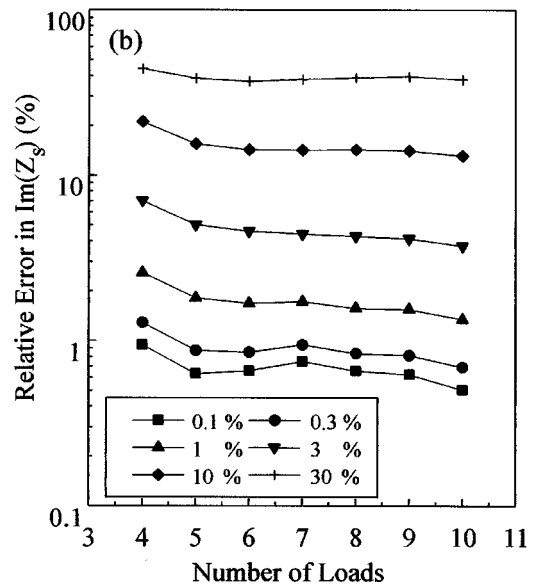
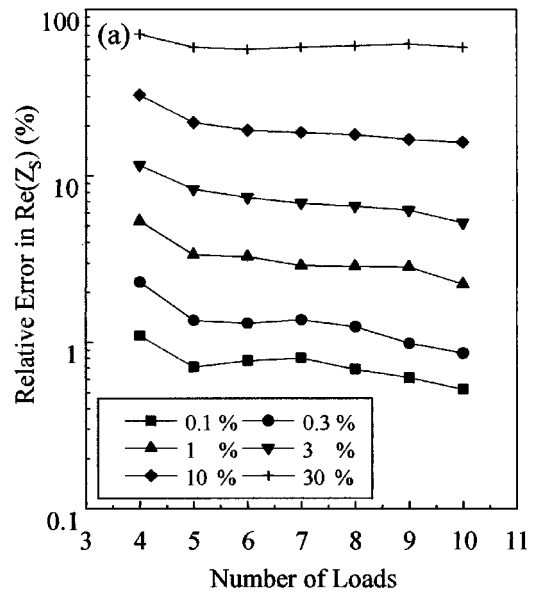


FIG. 6. Effect of number of loads on the error sensitivity of the resultant source impedance to the relative input error. Each symbol represents a magnitude of input error percentage in the simulation of the present method. (a) Resistance, (b) reactance.

multiple microphone method,<sup>9</sup> which is one of the direct measurement methods, was selected as the reference. Of course, other measuring techniques can be used for this purpose such as laser vibrometry and in-duct measurement.<sup>18</sup> The superiority of the present method compared to other previous methods is confirmed with these example applications. To demonstrate the practical applicability of the proposed method, the method was finally applied to measuring the source parameters on the exhaust side of an IC (internal combustion) engine.

#### A. Loudspeaker in a duct

The source parameters of a loudspeaker, which was connected to an acrylic circular tube of 176 mm in length, 40.0 mm in diameter, and 5 mm in thickness, were measured by using eight open-ended pipes with different lengths as loads. The pipes used as loads had the following lengths in m:



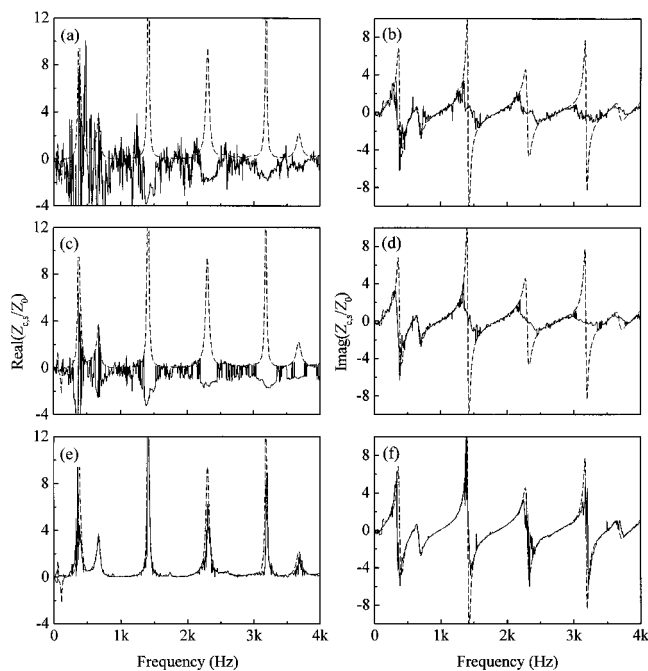


FIG. 7. Comparison of measured source impedance for a loudspeaker in duct: —, load method; ---, multiple-microphone method. (a), (b) Conventional method; (c), (d) least-squares method; (e), (f) present method.

0.227, 0.35, 0.444, 0.645, 0.842, 0.937, 1.329, 1.833. The acoustic pressure signals were measured at the interface between source and load by using a flush mounted 1/4-in microphone (B&K 4135). After the same number of averaging for each loading condition, the acoustic power spectra for the input data were obtained by the signal analyzer (HP 35670A).

The measured source impedance of the loudspeaker source is shown in Fig. 7. It can be seen that the conventional method produces large errors, while the present method yields the best agreement with the actual source impedance of the three indirect techniques. The direct least-squares method gives better results than the conventional one, but still produces a large error compared to the present one. Unlike the present method, the direct least-squares method uses the equation for estimating the source strength in order to find the source impedance value.<sup>12</sup> In this case, the accuracy of the calculated source impedance appears to be seriously affected by that of the source strength estimation.

## B. Centrifugal blower in a duct

A centrifugal blower attached to a PVC pipe of 1.32 m in length and 40.0 mm in diameter was also tested in an operating condition with mean flow of  $M=0.082$ . Eight open pipes with different lengths were used as loads. The lengths of loads consisted of the following in m: 0.541, 0.663, 0.758, 0.959, 1.156, 1.251, 1.643, 2.147. Two types of measurements were performed in this case: in-duct and out-duct measurements. For in-duct measurement for both direct and indirect methods, the correlation technique<sup>19</sup> was adopted to suppress the flow-generated noise from the turbulent flow. For out-duct measurements for implementing the load meth-

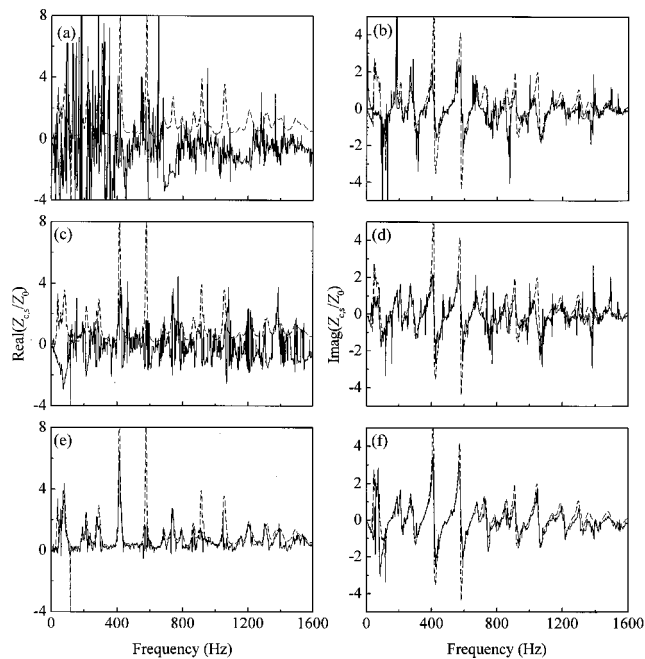


FIG. 8. Comparison of measured source impedance for a centrifugal blower in duct (in-duct measurement): —, load method; ---, multiple-microphone method. (a), (b) Conventional method; (c), (d) least-squares method; (e), (f) present method.

ods, i.e., indirect methods, the radiated sound pressure was measured at a microphone (B&K 4135) located 0.1 m from the pipe opening at  $90^\circ$  to the duct axis. Assuming monopole radiation from the pipe end, the measured pressure was then converted into the load pressure by accounting for both the reflection coefficient of an unflanged pipe and the transfer matrix of a straight pipe in the presence of mean flow.<sup>20,21</sup>

The source properties of the centrifugal blower, obtained by using the in-duct and out-duct pressure measurements, are shown in Figs. 8 and 9, respectively. It can be seen that the conventional method results in large errors, especially in the real part of the source impedance, and the present method yields more enhanced and reliable results than the other load methods. Large ghost peaks appearing in the results of the least-squares method do not appear in those of the present method. Large errors at about 120 Hz in the direct method are due to the poor signal-to-noise ratio at the measuring point, which is caused by the pressure node generated from the additional external source excitation.

## C. Exhaust system of an IC engine

Finally, the source characteristics for the exhaust of a V6 gasoline engine were measured using the present method. The acoustic load set consisted of 14 open pipes with different lengths between 0.15 and 2.5 m. An acoustic power spectrum for each load condition was measured outside the exhaust pipe with a half-inch microphone (B&K 4134) in an anechoic chamber. The engine speed and load were controlled by a dynamometer installed in a separate room annexed to the anechoic chamber. In addition, the temperatures at several points in each pipe were measured using thermocouples and the flow velocity at the open end was measured as well. These data were required to calculate the load im-

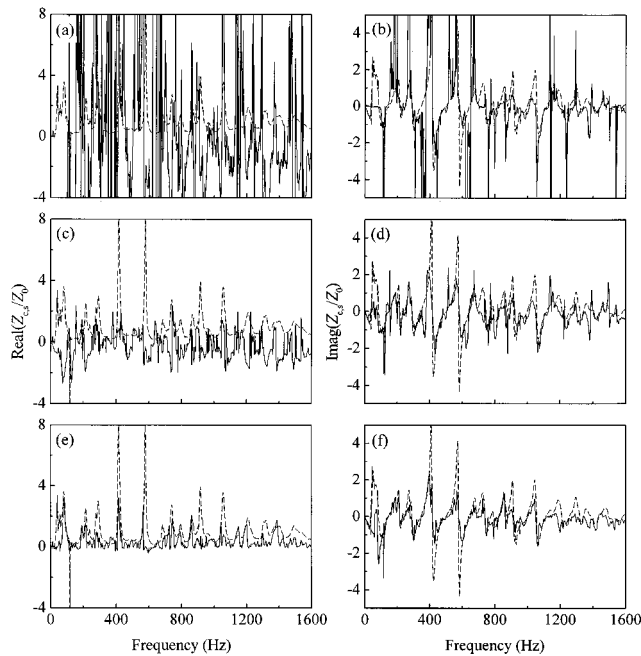


FIG. 9. Comparison of measured source impedance for a centrifugal blower in duct (out-duct measurement): —, load method; ---, multiple-microphone method. (a), (b) Conventional method; (c), (d) least-squares method; (e), (f) present method.

pedance and to transfer the measured out-duct sound pressure to the pressure at the interface between source and load.

In Fig. 10, the measured source characteristics for the exhaust of the engine operating at 2500 rpm are shown as an order plot. The gas temperature was 170 °C and the flow speed was 46 m/s ( $M=0.11$ ) at the junction of source and load. One can see that the real part of the convective source impedance has negative values at almost all orders. This seems to be unphysical and is against the fact that the convective source resistance should be positive for the present source model as mentioned earlier. This strange result has been reported in previous works<sup>1,12,22,23</sup> and the cause of this phenomenon has been ascribed to the nonlinearity and the time-varying nature of the actual source. In addition, it has been claimed that the presence of negative resistance does not harm the accuracy of the final predicted result of insertion loss of the whole silencing system: most of the previous results showed good agreements with measured insertion losses. This fact can also be confirmed in this study. The sound pressure level at 0.5 m from the outlet pipe opening was predicted when an open pipe of 1.05 m in length was connected to the system. Figure 11 shows the comparison of predicted and measured radiated sound spectrum. It can be observed that predicted and measured sound pressure level agree very well except for very high order components that have negligibly small magnitudes. This reveals that the present load method can be a very good practical tool in predicting the acoustic performance of the intake and exhaust systems of IC engines.

#### IV. CONCLUSIONS

In order to avoid the instability problem of the conventional four-load method, a new formulation for the multiloading

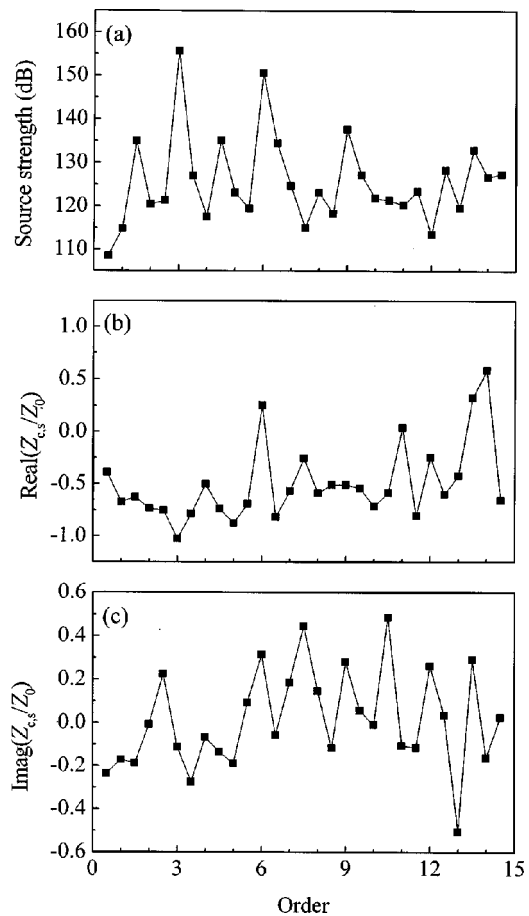


FIG. 10. Measured source parameters of the exhaust of an IC engine operating at 2500 rpm. (a) Source strength, (b) source resistance, (c) source reactance.

method has been proposed, which employs an error function based on the linear, time-invariant source model. It was numerically shown that the method is less sensitive to input errors compared to the conventional four-load method as well as the least-squares multiloading method. When the relative input error is 10%, the relative error in the resistance of the source, as calculated by the conventional method, is about 100 times larger than that by the proposed method. In particular, when an overdetermined system of eight loads is used for the measurements, the relative error in the resistance of the source by the proposed method maintains a very small value. The effectiveness of the present method was demonstrated by two test examples, a loudspeaker and a blower,

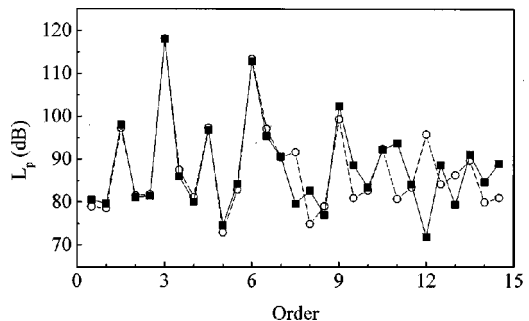


FIG. 11. Comparison of predicted and measured sound spectrum at 0.5 m from the exhaust opening: —■—, predicted; - -○- -, measured.

operating in a duct. The measured source impedances by the present method were compared with those by direct method, conventional method, and least-squares method. The proposed ‘‘refined’’ multiloading method was used to obtain the source parameters for the exhaust of an internal combustion engine. It is observed that the source resistance for the exhaust of the internal combustion engine has a negative value, although the convective effect is considered. In spite of this fact, the predicted radiated sound pressure level spectrum agreed very well with the measured spectrum.

Although various types of loads are preferable from the viewpoint of accuracy, the overdetermined set of open pipes with different lengths can be used as an alternative. However, further study is needed for determining what combinations of the load impedances will yield the best result. Although the predicted sound pressure spectrum agrees well with the measured spectrum, the origin of the negative source resistance should be uncovered in further studies. For this purpose, the basic assumptions for the linear, time-invariant sources should be reconsidered and a detailed full simulation of the wave and gas propagation processes such as in Ref. 23 should be performed. Although the foregoing features are not clarified yet, the present method would be useful in predicting the insertion loss or the radiated sound level from intake or exhaust systems of practical applications such as engines, fans, compressors, etc.

## ACKNOWLEDGMENTS

The authors would like to thank Hyundai Motor Co. for allowing them to use test facilities and for supporting the experiments. The authors also thank Dr. Keith S. Peat at Loughborough University in the U.K. for his discussion and help. This work was partially supported by the Regional Research Center for Advanced Climate Control Technology at Sonmoon University in Korea.

## APPENDIX: GENERATING EQUATIONS FOR SOURCE IMPEDANCE

Substituting  $re^{i\theta}$  for  $Z$ , Eq. (12) can be expressed by two cubic polynomial equations as

$$A_1 r^3 + B_1 r^2 + C_1 r + D_1 = 0, \quad (A1)$$

$$A_2 r^3 + B_2 r^2 + C_2 r + D_2 = 0, \quad (A2)$$

where the coefficients are given by

$$A_1 = \sum_{m < n} \left( \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} - \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} \right)^2 (\cos^3 \theta + \sin^2 \theta \cos \theta), \quad (A3)$$

$$B_1 = \sum_{m < n} \left\{ \frac{|P_{c,m}|^4}{|Z_{c,m}|^4} (2R_m \cos^2 \theta + 2X_m \cos \theta \sin \theta + R_m) + \frac{|P_{c,n}|^4}{|Z_{c,n}|^4} (2R_n \cos^2 \theta + 2X_n \cos \theta \sin \theta + R_n) - \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} [2(R_m + R_n) \cos^2 \theta + 2(X_m + X_n) \cos \theta \sin \theta + R_m + R_n] \right\}, \quad (A4)$$

$$C_1 = \sum_{m < n} \left\{ \frac{|P_{c,m}|^4}{|Z_{c,m}|^4} (3R_m^2 \cos \theta + 2R_m X_m \sin \theta + X_m^2 \cos \theta) + \frac{|P_{c,n}|^4}{|Z_{c,n}|^4} (3R_n^2 \cos \theta + 2R_n X_n \sin \theta + X_n^2 \cos \theta) - \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} [(R_m^2 + R_n^2) \cos \theta + 4R_m R_n \cos \theta + (X_m^2 + X_n^2) \cos \theta + 2(R_m X_n + R_n X_m) \sin \theta] \right\}, \quad (A5)$$

$$D_1 = \sum_{m < n} \left\{ \frac{|P_{c,m}|^4}{|Z_{c,m}|^4} (R_m^3 + R_m X_m^2) + \frac{|P_{c,n}|^4}{|Z_{c,n}|^4} (R_n^3 + R_n X_n^2) - \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} (R_m R_n^2 + R_m X_n^2 + R_m^2 R_n + R_n X_m^2) \right\}, \quad (A6)$$

$$A_2 = \sum_{m < n} \left( \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} - \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} \right)^2 (\sin^3 \theta + \cos^2 \theta \sin \theta), \quad (A7)$$

$$B_2 = \sum_{m < n} \left\{ \frac{|P_{c,m}|^4}{|Z_{c,m}|^4} (2X_m \sin^2 \theta + 2R_m \cos \theta \sin \theta + X_m) + \frac{|P_{c,n}|^4}{|Z_{c,n}|^4} (2X_n \sin^2 \theta + 2R_n \cos \theta \sin \theta + X_n) - \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} [2(X_m + X_n) \sin^2 \theta + 2(R_m + R_n) \cos \theta \sin \theta + X_m + X_n] \right\}, \quad (A8)$$

$$C_2 = \sum_{m < n} \left\{ \frac{|P_{c,m}|^4}{|Z_{c,m}|^4} (3X_m^2 \sin \theta + 2R_m X_m \cos \theta + R_m^2 \sin \theta) + \frac{|P_{c,n}|^4}{|Z_{c,n}|^4} (3X_n^2 \sin \theta + 2R_n X_n \cos \theta + R_n^2 \sin \theta) - \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} [(X_m^2 + X_n^2) \sin \theta + 4R_m R_n \sin \theta + (R_m^2 + R_n^2) \sin \theta + 2(R_m X_n + R_n X_m) \cos \theta] \right\}, \quad (A9)$$

$$D_2 = \sum_{m < n} \left\{ \frac{|P_{c,m}|^4}{|Z_{c,m}|^4} (X_m^3 + R_m^2 X_m) + \frac{|P_{c,n}|^4}{|Z_{c,n}|^4} (X_n^3 + R_n^2 X_n) - \frac{|P_{c,m}|^2}{|Z_{c,m}|^2} \frac{|P_{c,n}|^2}{|Z_{c,n}|^2} (X_m X_n^2 + R_n^2 X_m + X_m^2 X_n + R_m^2 X_n) \right\}. \quad (A10)$$

Here,  $R_n, X_n$  denote the real and the imaginary part of  $Z_{c,n}$ , respectively. Note that Eqs. (A3)–(A10) are functions of the measured power spectrum of the pressure and acoustic impedance for each load as well as the angle  $\theta$ . One can divide the  $0-2\pi$  range into  $N_\theta$  intervals in order to calculate the coefficients, and a  $(r, \theta)$  set minimizing  $F(Z)$  should be se-

lected from the solutions that satisfy both Eqs. (A1) and (A2).

- <sup>1</sup>M. G. Prasad, "A four load method for evaluation of acoustical source impedance in a duct," *J. Sound Vib.* **114**, 347–356 (1987).
- <sup>2</sup>M. G. Prasad and M. J. Crocker, "On the measurement of the internal source impedance of a multi-cylinder engine exhaust system," *J. Sound Vib.* **90**, 479–508 (1983).
- <sup>3</sup>M. L. Kathuriya and M. L. Munjal, "Experimental evaluation of the aeroacoustic characteristics of a source of pulsating gas flow," *J. Acoust. Soc. Am.* **65**, 240–278 (1979).
- <sup>4</sup>A. G. Galaitis and E. K. Bender, "Measurement of the acoustic impedance of an internal combustion engine," *J. Acoust. Soc. Am. (Suppl. 1)* **58**, S8 (A) (1975).
- <sup>5</sup>D. F. Ross and M. J. Crocker, "Measurement of the acoustical internal impedance of an internal combustion engine," *J. Acoust. Soc. Am.* **74**, 18–27 (1983).
- <sup>6</sup>A. F. Seybert and D. F. Ross, "Experimental determination of acoustic properties using a two-microphone random-excitation technique," *J. Acoust. Soc. Am.* **61**, 1362–1370 (1977).
- <sup>7</sup>M. L. Munjal and A. G. Doige, "The two-microphone method incorporating the effects of mean flow and acoustic damping," *J. Sound Vib.* **137**, 135–138 (1990).
- <sup>8</sup>T. Fujimori, S. Sato, and H. Miura, "An automated measurement system of complex sound pressure reflection coefficients," *Proc. Inter Noise* **84**, 1009–1014 (1984).
- <sup>9</sup>S.-H. Jang and J.-G. Ih, "On the multiple microphone method for measuring in duct acoustic properties in the presence of mean flow," *J. Acoust. Soc. Am.* **103**, 1520–1526 (1998).
- <sup>10</sup>B. S. Sridhara and M. J. Crocker, "Error analysis for the four-load method used to measure the source impedance in ducts," *J. Acoust. Soc. Am.* **92**, 2924–2931 (1992).
- <sup>11</sup>L. Desmons and J. Hardy, "A least squares method for evaluation of characteristics of acoustical sources," *J. Sound Vib.* **175**, 365–376 (1994).
- <sup>12</sup>H. Bodén, "On multi-load methods for determination of the source data of acoustic one-port sources," *J. Sound Vib.* **180**, 725–743 (1995).
- <sup>13</sup>M. L. Munjal and A. G. Doige, "On the relation between convective source characteristics and their acoustic counterparts," *J. Sound Vib.* **136**, 343–346 (1990).
- <sup>14</sup>H. Bodén and M. Åbom, "Maximum sound power from in-duct sources with applications to fans," *J. Sound Vib.* **187**, 543–550 (1995).
- <sup>15</sup>A. G. Doige and H. S. Alves, "Experimental characterization of noise sources for duct acoustics," *Trans. ASME, J. Vib. Acoust. Stress Rel. Design* **111**, 108–114 (1989).
- <sup>16</sup>B. S. Sridhara and M. J. Crocker, "Prediction of sound pressure radiated from open end of a pipe and muffler insertion loss using a single, efficient scheme: Application to a vacuum pump," *J. Acoust. Soc. Am.* **99**, 1333–1338 (1996).
- <sup>17</sup>J.-G. Ih, S.-H. Jang, S.-J. Kim, and J.-S. Shim, "Measurement of the acoustic source characteristics of the intake port in the refrigerator compressor," *Proceedings of the 1998 International Compressor Engineering Conference, Purdue, 1998*, pp. 561–564.
- <sup>18</sup>J.-S. Lee and J.-G. Ih, "On the method for estimating the volume velocity of an acoustic source in a chamber," *J. Sound Vib.* **182**, 505–522 (1995).
- <sup>19</sup>A. Michalke, F. Arnold, and F. Holste, "On the coherence of the sound field in a circular duct with uniform mean flow," *J. Sound Vib.* **190**, 261–271 (1996).
- <sup>20</sup>M. L. Munjal, *Acoustics of Ducts and Mufflers* (Wiley-Interscience, New York, 1987).
- <sup>21</sup>R. M. Munt, "Acoustic transmission properties of a jet pipe with subsonic jet flow: I. The cold jet reflection coefficient," *J. Sound Vib.* **142**, 413–436 (1990).
- <sup>22</sup>H. Bodén and F. Albertson, "In-duct acoustic one-port sources, linear or nonlinear?" *Proc. Inter Noise* **98**, 227–230 (1998).
- <sup>23</sup>V. H. Gupta and M. L. Munjal, "On numerical prediction of the acoustic source characteristics of an engine exhaust system," *J. Acoust. Soc. Am.* **92**, 2716–2725 (1992).

# Control performance and robustness of an active noise control system for uncertain primary sound fields

Doo-Byung Yoon<sup>a)</sup> and Yang-Hann Kim<sup>b)</sup>

*Center for Noise and Vibration Control, Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Science Town, Taejeon 305-701, Korea*

(Received 13 September 1998; revised 29 April 1999; accepted 11 February 2000)

In this work, the control performance and robustness of an active noise control system subjected to uncertain primary sound fields are investigated. For this purpose, the performance index, residual potential energy in a desired quiet zone, is derived as a function of sound field variables, quiet zone variables, and control system variables. In the presence of uncertainty, typical measures of the robustness and performance of a control system, maximum, minimum, mean, and variance of the performance index are derived theoretically. In addition, based on the least-squares orthogonality principle, the condition for implementing the best-oriented control system, which is robust and can maximize the control performance by using a given number of control sources and sensors, is investigated. Finally, numerical simulations are carried out to illustrate and verify the proposed theory. © 2000 Acoustical Society of America. [S0001-4966(00)03505-0]

PACS numbers: 43.50.Ki [MRS]

## INTRODUCTION

Active noise control is a method used to reduce unwanted noise in a desired quiet zone by generating a control sound by means of control sources (Fig. 1). In active noise control, the cost function is typically constructed as the sum of the squared errors from a number of discrete sensor positions; for example, see Refs. 1–8. The performance of an active control system can be evaluated by a performance index, which is a measure of the integrated resulting noise in a desired quiet zone. Typically, the ratio of the performance index before and after control can be used as a measure of the control performance.

The sound field in a desired quiet zone within an interior space can be expressed as a linear superposition of the eigenfunctions (acoustic modes) that satisfy the boundary conditions of the sound field. If one wants to perfectly control any noise in a desired quiet zone, then one must have a sufficient number of sensors and control sources, so that the control system is perfectly observable and controllable.<sup>3–5</sup> Typical examples of this kind can be found in Refs. 3 and 4. In these studies, the first four propagating modes in a square duct have been controlled using four sensors and control sources.

However, when the number of eigenfunctions is relatively large, it is impractical to utilize a control system that consists of a large number of sensors and control sources. In practice, one has to decrease significantly the noise in a desired quiet zone by using a few suitably placed control sources and sensors. In this case, the primary sound field, the location of control sources and sensors can affect the performance of the control system (Fig. 2).

Recently, the optimal control source and sensor location problem has been investigated using optimization techniques that employ, for example, genetic algorithms to minimize the

squared difference between the unwanted sound field and the control source-generated sound field at the discrete sensor positions.<sup>6–8</sup> However, even though the residual pressure at the sensing points may be small, there is no guarantee that the integrated residual noise in a desired quiet zone will be globally reduced.

When we sample the primary sound field at discrete sensing points using fewer sensors than the number of eigenfunctions necessary to represent the sound field, we obtain only partial information about the entire continuous field. In other words, we do not know the primary sound field completely. Also, the actual primary sound field might vary from case to case. This lack of knowledge of the primary sound field is the main uncertainty<sup>9</sup> which the control system must handle. A control system is good if the performance index is less than that of uncontrolled value regardless of what primary sound field actually occurs in practice. Such a control system is robust respect to the uncertainty in the sound field. On the other hand, if the small variation in the primary sound field can cause the substantial increase of the performance index, then the control system is fragile with respect to uncertainty, and this is undesirable.

These considerations motivated the present study which considers the robustness and performance of a control system, when the sum of squared pressure at discrete sensor positions is minimized. Our first objective is to investigate the control performance and robustness of an active control system subjected to uncertain primary sound fields. The second objective is to find a control strategy for a desired quiet zone, whose number of eigenfunctions is known, but whose modal participation is unknown, by using a given number of sensors and control sources.

To achieve these objectives, the first thing to do will be a mathematical formulation associated with the problem. In Sec. I, equations of the performance index (the acoustic potential energy in a desired quiet zone) are derived as a function of sound field variables, quiet zone variables, and con-

<sup>a)</sup>Electronic mail: m9900431@hanta.co.kr

<sup>b)</sup>Electronic mail: yhkim@sorak.kaist.ac.kr

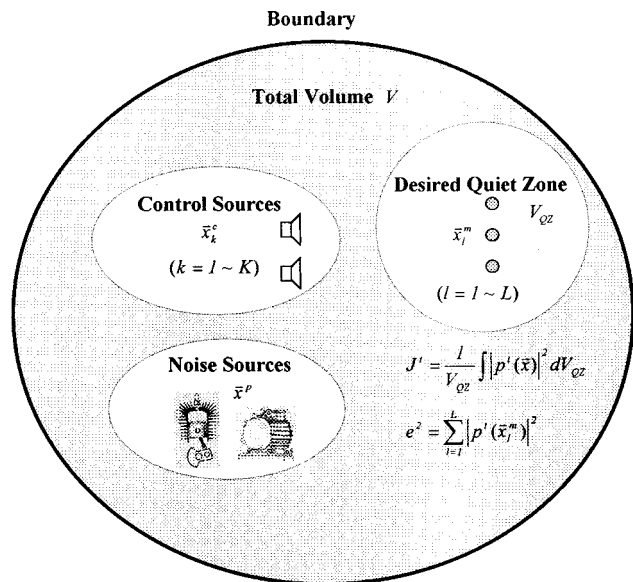


FIG. 1. Schematic diagram of an active noise control system with  $K$  control sources and  $L$  sensors.

control system variables (Fig. 2). The unknown primary sound fields are considered as sound field variables. We consider the size and location of the desired quiet zone as quiet zone variables. The number and locations of control sources and sensors are considered as control system variables.

In Sec. II, control performance of the control system will be theoretically investigated. The representative measures of the robustness and control performance, the maximum, minimum, mean, and variance of the performance index, will be derived for the set of uncertain primary sound fields. In Sec. III, the possibility of implementing the best-oriented control system, which is robust and can maximize control performance by using a given number of control sources and sensors, will be investigated. Finally, in Sec. IV, to illustrate the potential of the proposed theory, numerical simulations will be carried out.

## I. GENERAL FORMULATION

### A. Derivation of the performance index of a control system

The performance of a control system is dependent on the primary sound field, desired quiet zone, and locations of con-

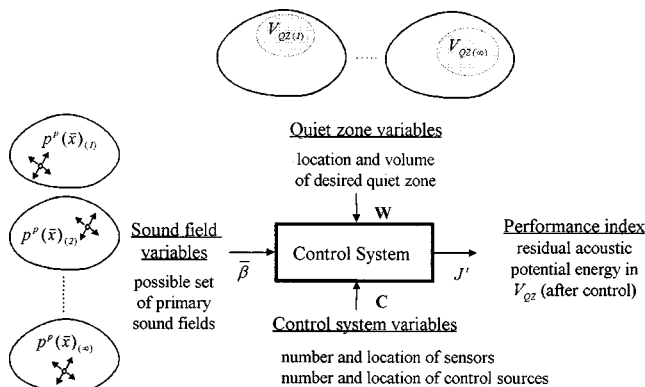


FIG. 2. The relations between sound field variables, control system variables, quiet zone variables, and the performance index.

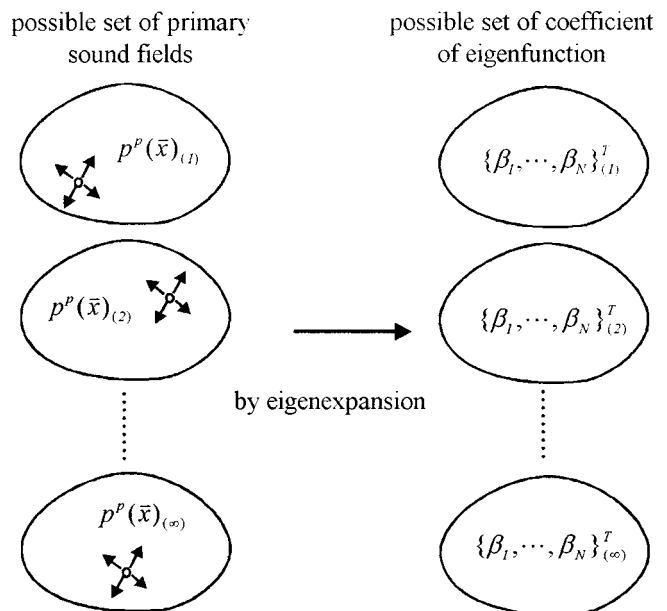


FIG. 3. The relation between primary sound fields  $[p^p(\mathbf{x})]$  and sound field variables ( $\beta$ ).

control sources and sensors (Fig. 2). This section derives the performance index (potential energy in a desired quiet zone) in terms of sound field variables, quiet zone variables, and control system variables when the sum of squared pressure at sensor positions is minimized.

### 1. Residual pressure after application of the control system

Assume that the unknown variable primary sound field within an interior space can be represented as a finite sum of eigenfunctions that satisfies the boundary condition as

$$p^p(\mathbf{x}) = \sum_{n=1}^N \beta_n \psi_n(\mathbf{x}), \quad (1)$$

where  $\psi_n(\mathbf{x})$  and  $\beta_n$  denote the  $n$ th orthonormal eigenfunction of the sound field and its amplitude, respectively (Fig. 3).  $N$  denotes the number of eigenfunctions that is necessary to represent the sound field. The basic assumption here is that  $\psi_n(\mathbf{x})$  and  $N$  are known. If the frequency range of interest is bandlimited, then this assumption is reasonable.

Alternatively, Eq. (1) can be expressed in a matrix form as an inner product of  $\bar{\psi}$  and  $\bar{\beta}$  as

$$p^p(\mathbf{x}) = \bar{\psi} \bar{\beta}, \quad (2)$$

where  $\bar{\psi} = \{\psi_1(\mathbf{x}), \dots, \psi_N(\mathbf{x})\}$  is the  $N$ -length vector that satisfies the orthonormal condition,  $1/V \int_V \bar{\psi}^H \bar{\psi} dV = \mathbf{I}$ , where  $V$  is the total volume of the sound field.  $\bar{\beta} = \{\beta_1, \dots, \beta_N\}^T$  is the  $N$ -length vector, whose amplitudes are unknown and variable.

Let  $p^c(\mathbf{x})$  represent the pressure at  $\mathbf{x}$  due to  $K$  control sources located at  $\mathbf{x}_k^c$  (Fig. 1). Then,  $p^c(\mathbf{x})$  can be expressed as<sup>1,2</sup>

$$p^c(\mathbf{x}) = \bar{\psi} \mathbf{T}^c \Psi^c \bar{q}. \quad (3a)$$

$$\mathbf{T}^c = \begin{bmatrix} T_{11}^c & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & T_{NN}^c \end{bmatrix}, \quad (3b)$$

$$\Psi^c = \begin{bmatrix} \psi_1(\mathbf{x}_1^c)^* & \cdots & \psi_1(\mathbf{x}_K^c)^* \\ \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1^c)^* & \cdots & \psi_N(\mathbf{x}_K^c)^* \end{bmatrix}, \quad (3c)$$

where  $\bar{q} = \{q_1, \dots, q_K\}^T$  is the  $K$ -length vector that consists of the volume velocity ( $q_k$ ) of the control sources. The control sources assumed to be not coupled and modeled as constant volume velocity, the outputs of which are linearly related to their inputs.  $\Psi^c$  is the  $(N \times K)$  matrix that consists of  $\psi_n(\mathbf{x}_k^c)^*$ , where the superscript “\*” denotes complex conjugate.  $\Psi^c \bar{q}$  is the  $N$ -length vector representing the modal volume velocities due to  $K$  control sources.  $\mathbf{T}^c$  is the  $(N \times N)$  diagonal matrix which represents the relation between  $\Psi^c \bar{q}$  and modal pressure amplitudes at  $\mathbf{x}$ .

From Eqs. (2) and (3a), the measured pressure vector at the  $L$  error sensors located at  $\mathbf{x}_l^m$  can be expressed as

$$\bar{p}^m = \Psi^m (\bar{\beta} + \mathbf{T}^c \Psi^c \bar{q}), \quad (4)$$

where  $\Psi^m$  is the  $(L \times N)$  matrix consisting of  $\psi_n(\mathbf{x}_l^m)$ ,

$$\Psi^m = \begin{bmatrix} \psi_1(\mathbf{x}_1^m) & \cdots & \psi_N(\mathbf{x}_1^m) \\ \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_L^m) & \cdots & \psi_N(\mathbf{x}_L^m) \end{bmatrix}. \quad (5)$$

Then, the sum of squared pressure at the error microphone locations ( $e^2$ ) can be expressed as<sup>1,2</sup>  $e^2 = \bar{p}^m H \bar{p}^m$ , where the superscript “ $H$ ” denotes Hermitian transpose.

Assume that  $N \geq L \geq K$  and that rank of  $\Psi^m$  is  $L$  (rows of  $\Psi^m$  is linearly independent), and rank of  $\Psi^c$  is  $K$  (columns of  $\Psi^c$  is linearly independent).<sup>1,2,10</sup> Then, the control source volume velocities ( $\bar{q}$ ) which minimizes  $e^2$  is given by<sup>1,2</sup>

$$\bar{q} = -(\Psi^m \mathbf{T}^c \Psi^c)^+ \Psi^m \bar{\beta}, \quad (6)$$

where the superscript “+” denotes the Moore–Penrose generalized matrix inverse.<sup>11</sup> It should be mentioned that  $\bar{q}$  is dependent on  $\Psi^m$ ,  $\Psi^c$ , and  $\bar{\beta}$ , where the modal amplitudes of primary sound field ( $\bar{\beta}$ ), and the position of the control sources ( $\Psi^c$ ) and sensors ( $\Psi^m$ ) are still not specified.

From Eqs. (2), (3a), and (6), the total pressure field [ $p^t(\mathbf{x})$ ] due to primary and control sources can be expressed as

$$p^t(\mathbf{x}) = \bar{\psi}(\mathbf{I} - \mathbf{T}^c \Psi^c (\Psi^m \mathbf{T}^c \Psi^c)^+ \Psi^m) \bar{\beta}, \quad (7)$$

where  $\mathbf{I}$  is the  $(N \times N)$  identity matrix. For convenience, let the second term in the parentheses be  $\mathbf{C} = \mathbf{T}^c \Psi^c (\Psi^m \mathbf{T}^c \Psi^c)^+ \Psi^m$ ; then, the residual pressure field can be expressed in terms of sound field variables ( $\bar{\beta}$ ) and control system variables ( $\mathbf{C}$ ) as

$$p^t(\mathbf{x}) = \bar{\psi}(\mathbf{I} - \mathbf{C}) \bar{\beta}. \quad (8)$$

It is noteworthy that if  $\mathbf{C}$  is the  $(N \times N)$  identity matrix [ $\text{rank}(\Psi^m) = \text{rank}(\Psi^c) = N$ ],<sup>3–5</sup> then  $p^t(\mathbf{x})$  vanishes for any value of  $\bar{\beta}$ . The aim of this study is to investigate the control performance and robustness<sup>9</sup> of a control system, which con-

sists of fewer sensors and control sources than  $N$ . Thus, the control performance and robustness of such a control system will be investigated in the Sec. II.

## 2. Performance index of the control system

In this section the performance index (residual potential energy in a desired quiet zone) will be derived in terms of sound field variables, quiet zone variables, and control system variables. Let  $J^t$  denote the residual acoustic potential energy in a desired quiet zone. Then,  $J^t$  can be defined as

$$J^t = \frac{1}{V_{\text{QZ}}} \int_{V_{\text{QZ}}} |p^t(\mathbf{x})|^2 dV_{\text{QZ}}, \quad (9)$$

where  $V_{\text{QZ}}$  is the volume of the desired quiet zone (Fig. 1).

Substitution of Eq. (8) into Eq. (9) yields

$$J^t = \bar{\beta}^H (\mathbf{I} - \mathbf{C})^H \left[ \frac{1}{V_{\text{QZ}}} \int_{V_{\text{QZ}}} \bar{\psi}^H \bar{\psi} dV_{\text{QZ}} \right] (\mathbf{I} - \mathbf{C}) \bar{\beta}. \quad (10)$$

If one defines the  $(N \times N)$  matrix within the brackets, which is a function of the location and volume of the desired quiet zone, as  $\mathbf{W}$ ,

$$\mathbf{W} = \frac{1}{V_{\text{QZ}}} \int_{V_{\text{QZ}}} \bar{\psi}^H \bar{\psi} dV_{\text{QZ}}, \quad (11)$$

then the residual acoustic potential energy ( $J^t$ ) can be written in a more compact form; that is,

$$J^t = \bar{\beta}^H (\mathbf{I} - \mathbf{C})^H \mathbf{W} (\mathbf{I} - \mathbf{C}) \bar{\beta}. \quad (12)$$

In Eq. (12), one can see that  $J^t$  can be expressed as a function of sound field variables ( $\bar{\beta}$ ), control system variables ( $\mathbf{C}$ ), and quiet zone variables ( $\mathbf{W}$ ) (Fig. 2).

## B. Modeling of the uncertain primary sound fields

In this section, a method of modeling a set of uncertain primary sound fields that can be produced due to arbitrary primary sources will be investigated.

The acoustic potential energy in a desired quiet zone, due to arbitrary primary sources, can be obtained by substituting  $\mathbf{C} = \mathbf{0}$  into Eq. (12); that is,

$$J^p = \bar{\beta}^H \mathbf{W} \bar{\beta}, \quad (13)$$

where  $J^p$  can have any value depending on the variation of the primary sources. Since we have no information about the primary sound field, the first thing to do is to have a model of the primary sound fields, which can be produced in a desired quiet zone.

As mention in the Introduction, in this work the residual potential energy in a desired quiet zone ( $J^t$ ) is used as a performance index, and will be compared with that of uncontrolled value ( $J^p$ ). Since the ratio of the potential energy before and after control is used as the measure of the control performance, we can assume that  $J^p$  is an arbitrary given constant. Thus, the model employed here is a set of *all* possible primary sound fields, whose acoustic potential energy is  $J^p$  (arbitrary given constant).

Uncertainty of the primary sound field is introduced into the model by considering the modal amplitudes vector ( $\bar{\beta}$ ) to

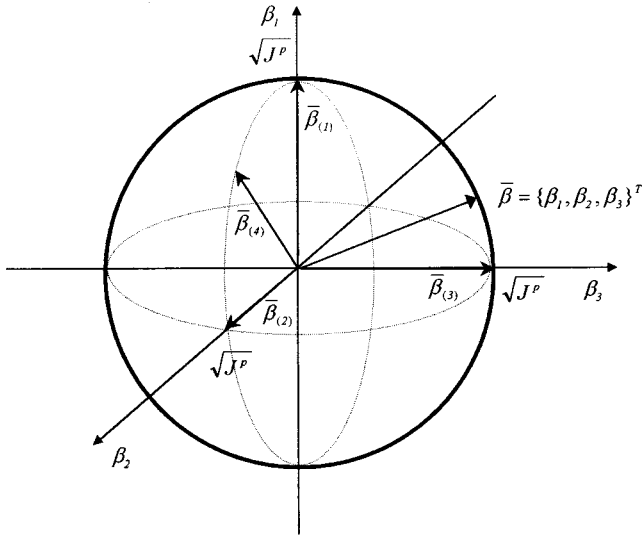


FIG. 4. Example of a set of primary sound fields,  $S(J^p)$ , when  $N=3$  and  $\mathbf{W}=\mathbf{I}$  [in this case, any primary sound field which satisfies the relation,  $\bar{\beta}^H \bar{\beta} = J^p$ , is the element of the  $S(J^p)$ ].

be selected statistically. Mathematically, the uncertainty in the primary sound field can be represented by allowing the modal amplitudes vector ( $\bar{\beta}$ ) to vary on a set of values bounded on an ellipsoid as

$$S(J^p) = \{p^p(\mathbf{x}) = \bar{\psi} \bar{\beta} | \bar{\beta}^H \mathbf{W} \bar{\beta} = J^p\}. \quad (14)$$

In this model, any primary sound field which satisfies the relation,  $\bar{\beta}^H \mathbf{W} \bar{\beta} = J^p$ , is the element of the  $S(J^p)$ . This implies that the  $\bar{\beta}$  can be regarded as a sample of *all* possible events. Each element ( $\bar{\beta}$ ) on an ellipsoid represents a possible realization of an uncertain primary sound field. It is noteworthy that unknown variations of the primary sound field can be represented by  $S(J^p)$ . Figure 4 represents an example of the  $S(J^p)$ , when  $N=3$  and  $\mathbf{W}=\mathbf{I}$ . In this case, any primary sound field ( $\bar{\beta}$ ) which satisfies the relation,  $\beta_1^H \beta_1 + \beta_2^H \beta_2 + \beta_3^H \beta_3 = J^p$ , is the element of the  $S(J^p)$ .

Now, we have an equation of a performance index [Eq. (12)] and an uncertainty model of the primary sound field [Eq. (14)]. Then, the next logical step is to investigate the robustness and performance of a control system under the uncertain primary sound field circumstance.

## II. ROBUSTNESS AND PERFORMANCE OF A CONTROL SYSTEM

This section studies the robustness and performance of a control system subjected to the set of uncertain primary sound fields [ $S(J^p)$ ]. The robustness of the control system can be checked by the maximum value of  $J^t$ . In the presence of uncertainty, the control performance can be described statistically. In the statistical point of view, the mean and variance of  $J^t$  can be used as measures to evaluate the control performance of the control system.

The residual acoustic potential energy in a desired quiet zone for an arbitrary  $\bar{\beta}$  in  $S(J^p)$  can be expressed as

$$J^t = \bar{\beta}^H (\mathbf{I} - \mathbf{C})^H \mathbf{W} (\mathbf{I} - \mathbf{C}) \bar{\beta}, \quad (15a)$$

$$\text{constraint } \bar{\beta}^H \mathbf{W} \bar{\beta} = J^p. \quad (15b)$$

The method of Lagrange multipliers can be used to represent the above-constrained equation<sup>12</sup>

$$\mathbf{L} = \bar{\beta}^H (\mathbf{I} - \mathbf{C})^H \mathbf{W} (\mathbf{I} - \mathbf{C}) \bar{\beta} + \lambda (J^p - \bar{\beta}^H \mathbf{W} \bar{\beta}), \quad (16)$$

where  $\mathbf{L}$  is the Lagrangian and  $\lambda$  is the Lagrange multiplier for the equality constraint [Eq. (15b)].

Differentiating  $\mathbf{L}$  with respect to  $\bar{\beta}$  and  $\lambda$  and setting them to zero, yields necessary conditions for an extremum of  $J^t$  as<sup>9,12</sup>

$$(\mathbf{I} - \mathbf{C})^H \mathbf{W} (\mathbf{I} - \mathbf{C}) \bar{\beta} = \lambda \mathbf{W} \bar{\beta}, \quad (17a)$$

$$\bar{\beta}^H \mathbf{W} \bar{\beta} = J^p, \quad (17b)$$

where Eq. (17a) is a generalized eigenvalue problem, and Eq. (17b) is just the original constraint.

Let  $\bar{\beta}' = \mathbf{W}^{1/2} \bar{\beta}$ , and by premultiplying Eq. (17a) by  $\mathbf{W}^{-(1/2)H}$ , then the necessary conditions for an extremum can be expressed in a more simple eigenvalue problem<sup>9</sup>

$$\mathbf{V} \bar{\beta}' = \lambda \bar{\beta}', \quad (18a)$$

$$\bar{\beta}'^H \bar{\beta}' = J^p, \quad (18b)$$

where  $\mathbf{V} = \mathbf{W}^{-(1/2)H} (\mathbf{I} - \mathbf{C})^H \mathbf{W} (\mathbf{I} - \mathbf{C}) \mathbf{W}^{-1/2}$ , and  $\bar{\beta}'$  represent the  $N$ -length vector, whose elements represent the modal components of  $J^p$ . Equation (18a) show that  $\bar{\beta}'$  must be an eigenvector of the matrix  $\mathbf{V}$  and the Lagrange multiplier ( $\lambda$ ) is the corresponding eigenvalue. Since  $\mathbf{V}$  is the  $(N \times N)$  Hermitian matrix, this eigenvalue problem yields a set of  $N$  orthogonal eigenvectors ( $\bar{\beta}'_n$ ) and associated eigenvalues ( $\lambda_n$ ) satisfying the relations,<sup>9,13</sup>

$$\mathbf{V} \bar{\beta}'_n = \lambda_n \bar{\beta}'_n, \quad (19a)$$

$$\bar{\beta}'_n{}^H \bar{\beta}'_n = J^p. \quad (19b)$$

By premultiplying Eq. (19a) by  $\bar{\beta}'_n{}^H$  and using Eq. (19b), the following relation can be obtained:

$$\lambda_n = \frac{\bar{\beta}'_n{}^H \mathbf{V} \bar{\beta}'_n}{\bar{\beta}'_n{}^H \bar{\beta}'_n} = \frac{J^t}{J^p} \Big|_{\bar{\beta}' = \bar{\beta}'_n}. \quad (20)$$

It is noteworthy that the eigenvector ( $\bar{\beta}'_n$ ) represents the direction of  $\bar{\beta}'$  where extremum of  $J^t$  occurs and the corresponding eigenvalue (Lagrange multiplier) represents the ratio of  $J^t$  to  $J^p$  when  $\bar{\beta}' = \bar{\beta}'_n$ .

Since the eigenvectors ( $\bar{\beta}'_n$ ) are orthogonal to each other, arbitrary sound field variables ( $\bar{\beta}'$ ) in  $S(J^p)$  can be represented as a linear superposition of  $\bar{\beta}'_n$  such that,

$$\bar{\beta}' = \sum_{n=1}^N a_n \bar{\beta}'_n, \quad (21a)$$

$$\text{constraint } \sum_{n=1}^N |a_n|^2 = 1. \quad (21b)$$

By premultiplying Eq. (18a) by  $\bar{\beta}'_n{}^H$  and using Eqs. (20) and (21), the residual potential energy in a desired quiet zone



TABLE I. Sufficient condition for implementing the best-oriented control system.

Desired zone of quiet	$V_{QZ}$	$V$
$\mathbf{W} \left( = \frac{1}{V_{QZ}} \int_{V_{QZ}} \bar{\psi}^H \bar{\psi} dV_{QZ} \right)$	$\mathbf{W}$	$\mathbf{I}$
modal component of $J^p$ (before control)	$\bar{\beta}'$	$\bar{\beta}$
modal component of $J^t$ (after control)	$(\mathbf{I} - \mathbf{W}^{1/2} \mathbf{C} \mathbf{W}^{-1/2}) \bar{\beta}'$	$(\mathbf{I} - \mathbf{C}) \bar{\beta}$
sufficient condition for the best-oriented control system	$\Psi^{mH} = \mathbf{W} \mathbf{T}^c \Psi^c$	$\Psi^{mH} = \mathbf{T}^c \Psi^c$

for an arbitrary sound field ( $\bar{\beta}'$ ) in  $S(J^p)$  can be expressed as

$$J^t = (|a_1|^2 \lambda_1 + |a_2|^2 \lambda_2 + \dots + |a_N|^2 \lambda_N) J^p. \quad (22)$$

In Eq. (22), we can see that  $J^t$  is determined by  $a_n$  [determined by sound field variables, Eq. (21)] and  $\lambda_n$  (determined by control system variables and quiet zone variables).

Let the eigenvalues be ordered as follows:  $\lambda_1 \leq \dots \leq \lambda_N$ . Then, from Eqs. (21) and (22), the minimum and maximum value of the performance index subjected to  $S(J^p)$  can be obtained as

$$J_{\min}^t = \lambda_1 J^p \quad [\text{when } \bar{\beta}' = \bar{\beta}'_1 (|a_1|^2 = 1, |a_2|^2 = \dots = |a_N|^2 = 0)], \quad (23a)$$

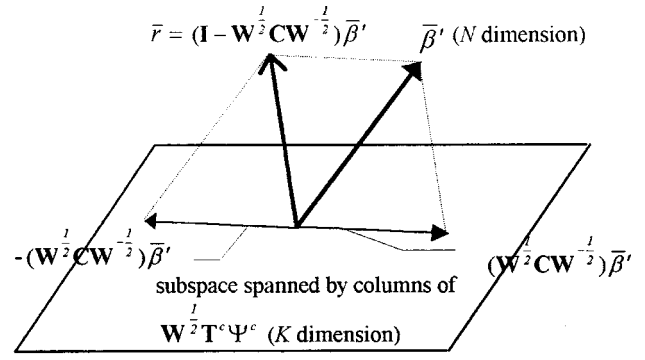
$$J_{\max}^t = \lambda_N J^p \quad [\text{when } \bar{\beta}' = \bar{\beta}'_N (|a_N|^2 = 1, |a_1|^2 = \dots = |a_{N-1}|^2 = 0)], \quad (23b)$$

where  $\lambda_1$  and  $\lambda_N$  denote the minimum and maximum value of  $\lambda_n$ , respectively.

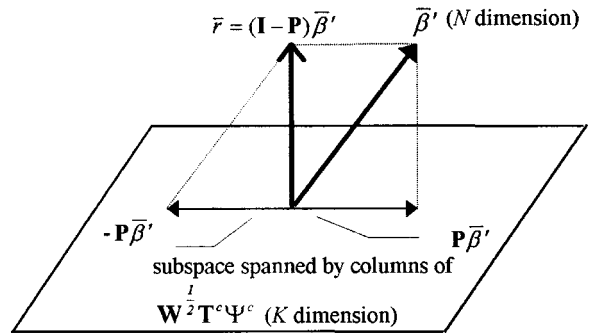
It is noteworthy that  $\lambda_N$  is a measure of the robustness which accounts for the design of the control system and the uncertainty in the primary sound fields. When  $\lambda_N \leq 1$ ,  $J^t$  is always less than  $J^p$ , regardless of the primary sound field in  $S(J^p)$ . On the other hand, when  $\lambda_N > 1$ ,  $J^t$  is larger than  $J^p$  for some  $\bar{\beta}'$  in  $S(J^p)$ .

In the presence of uncertainty, as a typical measure for the evaluation of the expected value and its dispersion of the performance index, the mean and variance are readily available. From Eqs. (21) and (22), the mean ( $J_{\text{mean}}^t$ ) and variance [ $\sigma^2(J^t)$ ] of the residual potential energy subjected to  $S(J^p)$  can be obtained as

$$J_{\text{mean}}^t = E[J^t] = \frac{J^p}{N} \sum_{n=1}^N \lambda_n, \quad (24)$$



(a)



(b)

FIG. 5. Geometric representation of the condition for the best-oriented control system. (a) Geometric representation of the controlled component and residual component of  $\bar{\beta}'$ . (b) Geometric representation of the best-oriented control system.

$$\begin{aligned} \sigma^2(J^t) &= E[(J^t - J_{\text{mean}}^t)^2] \\ &= \left[ \frac{2(N-1)}{N^2(N+2)} \sum_{n=1}^N \lambda_n^2 - \frac{2}{N^2(N+2)} \right. \\ &\quad \left. \times \sum_{n=1}^N \sum_{m=1, m \neq n}^N \lambda_n \lambda_m \right] J^{p^2}, \end{aligned} \quad (25)$$

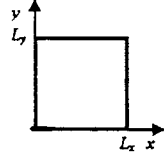
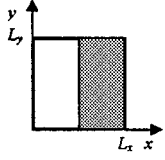
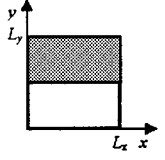
where  $E[\cdot]$  is the statistical expectation operator with respect to the sound field variables ( $\bar{\beta}'$ ) in  $S(J^p)$ . The detail derivations can be found in the Appendix.

From Eqs. (23), (24), and (25), one can see that the typical measures of the performance of a control system subjected to the set of uncertain primary sound fields [ $S(J^p)$ ] can be obtained by the eigenanalysis of the matrix  $\mathbf{V}$ . These measures can be used as useful tools for checking the robustness and performance of a control system, when designing or evaluating an active control system.

### III. CONDITION FOR THE BEST-ORIENTED CONTROL SYSTEM

In this section, for a given number of control sources and sensors, the condition that can maximize the control performance will be investigated based on the least-squares orthogonality principle.<sup>14</sup> This condition will be referred to as the condition for the best-oriented control system.

TABLE II. Eigenfunctions  $[\psi_n(x,y)]$ , its natural frequencies ( $f_n$ ), and  $T_{nn}^c$  of the lightly damped two-dimensional enclosure (Ref. 1).  $L_x=L_y=0.5$  m,  $\omega=2\pi f$ ,  $\rho$ : density of air,  $c$ : speed of sound.  $N_n$ : normalization factor of the acoustic mode ( $N_2=N_3=\sqrt{2}$ ).  $\zeta_n$ : damping ratio of the acoustic mode ( $\zeta_2=\zeta_3=0.01$ ).  $T_{nm}^c$ :  $n$ th diagonal element of  $\mathbf{T}^c$  [Eq. (3b)].

$n$	1	2	3
$\psi_n(x,y)$	 $1$	 $N_2 \cos \frac{\pi x}{L_x}$	 $N_3 \cos \frac{\pi y}{L_y}$
$f_n$	0 Hz	343 Hz	343 Hz
$\omega_n$	0 rad/sec	$2\pi \cdot 343$ rad/sec	$2\pi \cdot 343$ rad/sec
$T_{nn}^c$	$\frac{\rho c^2}{i\omega}$	$\frac{\omega \rho c^2}{i(\omega^2 - \omega_2^2) + 2\zeta_2 \omega \omega_2}$	$\frac{\omega \rho c^2}{i(\omega^2 - \omega_3^2) + 2\zeta_3 \omega \omega_3}$

From Eqs. (12), (13), and (18b), the ratio of  $J^i$  to  $J^p$  for any  $\bar{\beta}'$  in  $S(J^p)$  can be expressed as

$$\frac{J^i}{J^p} = \frac{\bar{\beta}'^H (\mathbf{I} - \mathbf{W}^{1/2} \mathbf{C} \mathbf{W}^{-1/2})^H (\mathbf{I} - \mathbf{W}^{1/2} \mathbf{C} \mathbf{W}^{-1/2}) \bar{\beta}'}{\bar{\beta}'^H \bar{\beta}'} = \frac{\bar{r}^H \bar{r}}{\bar{\beta}'^H \bar{\beta}'}, \quad (26)$$

where  $\bar{\beta}'$  and  $\bar{r}$  are the  $N$ -length vectors, whose elements represent the modal components of  $J^p$  and  $J^i$ , respectively.

Substitution of the relation  $\mathbf{C} = \mathbf{T}^c \Psi^c (\Psi^m \mathbf{T}^c \Psi^c)^+ \Psi^m$  into  $\bar{r}$  yields

$$\bar{r} = (\mathbf{I} - \mathbf{W}^{1/2} \mathbf{T}^c \Psi^c (\Psi^m \mathbf{T}^c \Psi^c)^+ \Psi^m \mathbf{W}^{-1/2}) \bar{\beta}', \quad (27)$$

where the residual modal components of  $J^i$  is related to the modal components of  $J^p$  by the term in the parentheses (Table I).

Assume that  $N \geq L \geq K$  and that columns of  $\mathbf{W}^{1/2} \mathbf{T}^c \Psi^c$  are linearly independent. Then, as shown in Fig. 5(a), these columns span a  $K$ -dimensional controllable subspace of the  $N$ -dimensional vector space spanned by  $\bar{\beta}'$ . This means that if  $\bar{\beta}'$  is a linear combination of the columns of  $\mathbf{W}^{1/2} \mathbf{T}^c \Psi^c$ , then  $\bar{r}$  is a null vector. Since an arbitrary vector ( $\bar{\beta}'$ ) cannot

be represented in terms of a linear combination of the columns of  $\mathbf{W}^{1/2} \mathbf{T}^c \Psi^c$ , the objective of this section is to find out the condition that can minimize the term in the parentheses of Eq. (27) for unknown arbitrary  $\bar{\beta}'$ .

The approach that can be used in this situation is to find the least-square solution.<sup>14</sup> As illustrated in Fig. 5(b), the least-square solution has the property that the residual component ( $\bar{r}$ ) is orthogonal to the controlled component ( $\bar{\beta}' - \bar{r}$ ). From Fig. 5(b), one can deduce that  $\bar{r}$  can be minimized when these two vectors are mutually orthogonal. In this case, using the ideas of the orthogonal projection of a vector on a subspace, the residual vector ( $\bar{r}$ ) can be represented as

$$\bar{r} = (\mathbf{I} - \mathbf{P}) \bar{\beta}', \quad (28)$$

where  $\mathbf{P} = \mathbf{W}^{1/2} \mathbf{T}^c \Psi^c (\mathbf{W}^{1/2} \mathbf{T}^c \Psi^c)^+$  is the orthogonal projection matrix<sup>14</sup> of rank  $K$ ,  $\mathbf{P} \bar{\beta}'$  is a orthogonal projection of  $\bar{\beta}'$  onto the  $K$ -dimensional subspace and represents a least-squares error approximation of  $\bar{\beta}'$ . Since  $\bar{r} = (\mathbf{I} - \mathbf{P}) \bar{\beta}'$  represents a least-squares error minimization,<sup>14</sup> in this case we can deduce that  $J^i(\bar{r}^H \bar{r})$  will be always less than  $J^p(\bar{\beta}'^H \bar{\beta}')$

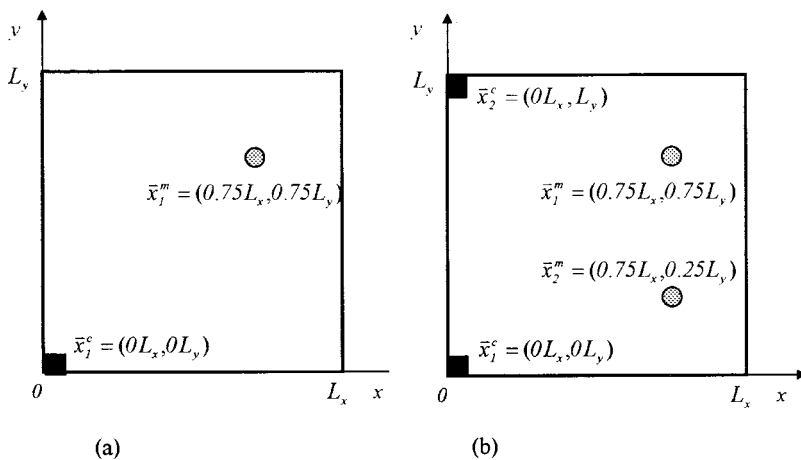


FIG. 6. Examples of the best-oriented control system (two-dimensional enclosure with  $L_x=0.5$  m and  $L_y=0.5$  m,  $N=3$ ,  $f=220$  Hz). (a) When  $K=L=1$ ; (b) when  $K=L=2$ .

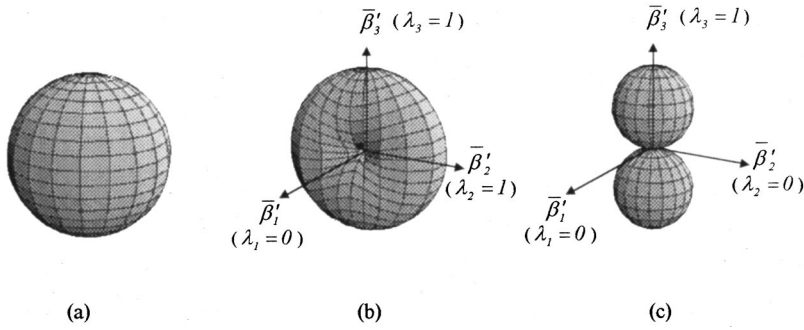


FIG. 7. Locus of  $\sqrt{J^I/J^P}$  of the best-oriented control system as the sound field variable varies (two-dimensional enclosure with  $L_x=0.5$  m and  $L_y=0.5$  m,  $N=3$ ,  $f=220$  Hz). (a) With no control; (b)  $K=L=1$  [Fig. 6(a),  $\bar{\beta}'_1=\{0.58,-0.58,-0.58\}^T$ ,  $\bar{\beta}'_2=\{0,0.71,-0.71\}^T$ ,  $\bar{\beta}'_3=\{0.82,0.41,0.41\}^T$ ]; (c)  $K=L=2$  [Fig. 6(b),  $\bar{\beta}'_1=\{0.58,-0.58,-0.58\}^T$ ,  $\bar{\beta}'_2=\{0.41,-0.41,0.82\}^T$ ,  $\bar{\beta}'_3=\{0.71,0.71,0\}^T$ ].

regardless of the what  $\bar{\beta}'$  actually is in practice.

By comparing Eq. (27) with Eq. (28), one can deduce that if one makes  $(\Psi^m \mathbf{T}^c \Psi^c)^+ \Psi^m \mathbf{W}^{-1/2}$  equal to  $(\mathbf{W}^{1/2} \mathbf{T}^c \Psi^c)^+$ , then the residual component [Eq. (27)] will be minimized. This condition can be satisfied when the control system variables satisfy the following relation:

$$\Psi^{mH} = \mathbf{W} \mathbf{T}^c \Psi^c. \quad (29)$$

It is noteworthy that Eq. (29) is the condition for implementing the best-oriented control system, which is robust and can maximize the control performance in uncertain primary sound field circumstances. Since Eq. (29) is the condition that relates the positions of control sources to the positions of sensors, there is an infinite set of positions of sensors ( $\Psi^m$ ) and control sources ( $\Psi^c$ ) satisfying the relation. Even though  $\Psi^{mH} = \mathbf{W} \mathbf{T}^c \Psi^c$  may not yield an easy solution for the optimal sensor and control source positions, in control system design perspective, it can be used as a useful measure for the evaluation of the control performance and robustness of a designed control system subjected to uncertain primary sound fields.

#### IV. NUMERICAL SIMULATION AND DISCUSSION

For the verification of the theoretical analysis results, the control performance and robustness of an active control system subjected to uncertain primary sound fields was investigated numerically.

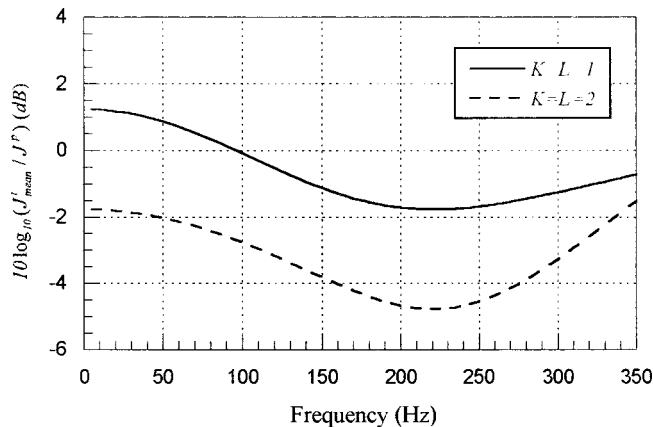


FIG. 8. Control performance ( $J^I_{\text{mean}}/J^P$ ) of the best-oriented control system with varying frequencies {solid line:  $K=L=1$  [Fig. 6(a)], dashed line:  $K=L=2$  [Fig. 6(b)]}.

As a sound field, a lightly damped two-dimensional enclosure sound field that consists of three modes ( $N=3$ ) was considered for the sake of simplicity. Table II summarizes the acoustic modes, natural frequencies, and  $\mathbf{T}^c$  of this sound field. In simulation, the entire region of the sound field was considered as the desired quiet zone ( $\mathbf{W}=\mathbf{I}$ ), and the best-oriented control systems that satisfy Eq. (29) were designed. The design frequency is chosen (220 Hz) that is lower than the natural frequency of the 2nd (3rd) modes (Table II). In Eq. (29), when the design frequency is specified, then the variables one can choose are the location of control source and sensor. For example, when  $N=3$  and  $K=L=1$ , if one selects the position of the control source at  $\mathbf{x}_1^c=(0L_x,0L_y)$ , then the error microphone locations, which satisfy Eq. (29), can be obtained as  $\mathbf{x}_1^m=(0.75L_x,0.75L_y)$ . In the same manner, if one selects the position of the control source at  $\mathbf{x}_2^c=(0L_x,L_y)$ , then the error microphone location can be obtained as  $\mathbf{x}_2^m=(0.75L_x,0.25L_y)$ . The schematic diagram of these control systems is shown in Fig. 6. The following describes the control performance and the robustness of these control systems.

In the presence of the uncertainty in the primary sound field, lower and upper bounds of the performance index have been obtained by numerical simulation. Figure 7 represents the control performance of the control system in terms of the sound field variables. Figure 7(b) represents the locus of  $\sqrt{J^I/J^P}$  of the control system with  $K=L=1$  [Fig. 6(a)] as the sound field variable varies. Figure 7(c) is the simulation result when  $K=L=2$  [Fig. 6(b)]. In Fig. 7(b) and (c), one could confirm that  $J^I$  is always less than  $J^P(\lambda_N=1)$ ; the best-oriented designed control system is robust with respect to the uncertainty in primary sound fields.

Since  $\mathbf{T}^c$  of Eq. (29) is a function of the frequency (Table II), it is necessary to investigate the performance of

TABLE III. Control performance of the control systems with varying microphone position ( $N=3$ ,  $K=1$ ,  $L=1$ ,  $\mathbf{x}_1^c=(0L_x,0L_y)$ ,  $L_x=L_y=0.5$  m,  $f=220$  Hz).

Case	$\mathbf{x}_1^m$	$J^I_{\text{mean}}/J^P$	$\sigma^2(J^I)/J^{P^2}$	$J^I_{\text{max}}/J^P$	$J^I_{\text{min}}/J^P$
1	$(0.75L_x,0.75L_y)$	0.67	0.09	1.00	0
2	$(0.5L_x,0.75L_y)$	0.83	0.15	1.49	0
3	$(0.25L_x,0.75L_y)$	3.29	6.28	8.86	0
4	$(0.25L_x,0.25L_y)$	3.43	6.92	9.28	0
5	$(0.5L_x,0.5L_y)$	1.32	0.60	2.95	0
6	$(1.0L_x,1.0L_y)$	0.68	0.09	1.03	0
7	$(0.0L_x,1.0L_y)$	5.26	18.16	14.77	0

the control system according to the variation of the frequency. Figure 8 is the simulation results which represent  $J^t_{\text{mean}}/J^p$  of the designed control systems [Fig. 6(a) and (b)] as frequency varies. In Fig. 8, one can see that the magnitude of  $J^t_{\text{mean}}/J^p$  increases as the difference between the primary source frequency and the design frequency (220 Hz) increases. On the other hand, in the narrow frequency band centered at the design frequency, one can observe that it is possible to control the potential energy by using fewer control sources ( $K$ ) and sensors ( $L$ ) than  $N$ . In addition, one can observe that the control performance has improved by increasing the number of control sources and sensors.

For the comparison of the control performance of the best-oriented control system with the other control systems that do not satisfy Eq. (29), the maximum ( $J^t_{\text{max}}$ ), mean ( $J^t_{\text{mean}}$ ), and variance [ $\sigma^2(J^t)$ ] of the performance index are obtained as the error microphone location varies. Table III represents the simulation results when  $K=L=1$ . From Table III, one can observe that the control performance of the best-oriented control system [case (1)] is the best, and  $J^t$  is always less than  $J^p$ . In Table III, for the cases that do not satisfy Eq. (29), one can observe that  $J^t_{\text{max}}$  is larger than  $J^p$ , and the control performances are worse than those of the best-oriented control system [case (1)].

## V. CONCLUSIONS

An analysis on the robustness and performance of an active control system subjected to uncertain primary sound fields has been carried out. For uncertain primary field circumstances, the typical measures of the robustness and control performance, maximum, minimum, mean, and variance of the residual potential energy were derived theoretically. Then, the condition for implementing the best-oriented control system, which is robust and can maximize the control performance by using a given number of control sources and sensors, was obtained.

Based on the theoretical analysis results, the control performance and robustness of the best-oriented control system were investigated by numerical simulations. The numerical simulation results demonstrate that it is possible to control the potential energy consisting of higher modes in the narrow frequency band centered at the design frequency by using the best-oriented control system. The results also demonstrate that the best-oriented control system shows better control performance than other control systems that do not satisfy the condition for the best-oriented control system.

## ACKNOWLEDGMENTS

The authors would like to thank Professor Yakov Ben-Haim in Israel Institute of Technology for his helpful discussion about the proposed theory. This work was supported by the Hyundai Marine Research Institute. The authors acknowledge the help and participation of the Noise and Vibration Department of HMRI.

## APPENDIX: DERIVATIONS OF THE MEAN AND VARIANCE OF THE PERFORMANCE INDEX

The mean ( $J^t_{\text{mean}}$ ) of  $J^t$  for the set of primary sound fields [ $S(J^p)$ ] can be expressed as

$$J^t_{\text{mean}} = E[J^t] = E[(|a_1|^2\lambda_1 + \dots + |a_N|^2\lambda_N)J^p], \quad (\text{A1a})$$

$$\text{constraint } \sum_{n=1}^N |a_n|^2 = 1. \quad (\text{A1b})$$

where  $E[\cdot]$  is the statistical expectation operator with respect to sound field variables ( $\beta'$ ) in  $S(J^p)$ .

Since  $\lambda_n$  of Eq. (A1) is constant, Eq. (A1) can be rewritten as

$$E[J^t] = J^p \sum_{n=1}^N \lambda_n E[|a_n|^2], \quad (\text{A2a})$$

$$\text{constraint } \sum_{n=1}^N |a_n|^2 = 1, \quad (\text{A2b})$$

where the  $E[|a_n|^2]$  are unknown. To find out  $E[|a_n|^2]$ , we can use Eq. (A2b). The expectation of Eq. (A2b) can be written as

$$E\left[\sum_{n=1}^N |a_n|^2\right] = E[|a_1|^2] + \dots + E[|a_N|^2] = 1. \quad (\text{A3})$$

Since the probability of  $|a_n|^2$  ( $n=1-N$ ) are equal to each other, we can see that  $E[|a_1|^2] = E[|a_2|^2] = \dots = E[|a_N|^2] = 1/N$ . Thus, the Eq. (A2) can be expressed as

$$J^t_{\text{mean}} = \frac{J^p}{N} \sum_{n=1}^N \lambda_n. \quad (\text{A4})$$

The variance [ $\sigma^2(J^t)$ ] of  $J^t$  subject to  $S(J^p)$  can be expressed as

$$\sigma^2(J^t) = E[(J^t - J^t_{\text{mean}})^2] = E[J^{t^2}] - J^{t^2}_{\text{mean}}, \quad (\text{A5a})$$

$$\text{constraint } \sum_{n=1}^N |a_n|^2 = 1, \quad (\text{A5b})$$

Using Eq. (A1a), the  $E[J^{t^2}]$  in Eq. (A5a) can be written as

$$E[J^{t^2}] = J^{p^2} \sum_{n=1}^N \lambda_n^2 E[|a_n|^4] + J^{p^2} \sum_{\substack{n=1 \\ n \neq m}}^N \sum_{m=1}^N \lambda_n \lambda_m E[|a_n|^2 |a_m|^2], \quad (\text{A6})$$

where  $E[|a_n|^4]$  and  $E[|a_n|^2 |a_m|^2]$  are unknown. To find out  $E[|a_n|^4]$  and  $E[|a_n|^2 |a_m|^2]$ , we can use Eq. (A5b). Since

$$\left(\sum_{n=1}^N |a_n|^2\right)^2 = 1,$$

the

$$E\left[\left(\sum_{n=1}^N |a_n|^2\right)^2\right]$$

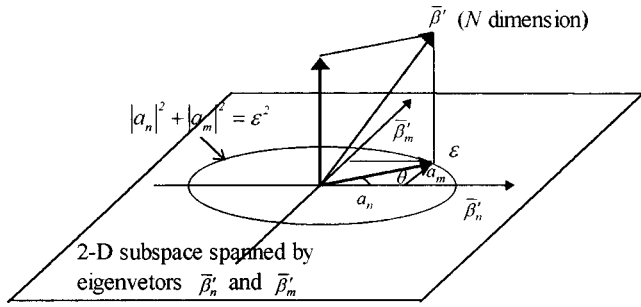


FIG. A1. Geometrical representation of  $|a_n|$  and  $|a_m|$  that satisfies  $|a_n|^2 + |a_m|^2 = \varepsilon^2$  ( $0 \leq \varepsilon \leq 1$ ).

can be written as

$$E\left[\left(\sum_{n=1}^N |a_n|^2\right)^2\right] = \sum_{n=1}^N E[|a_n|^4] + \sum_{\substack{n=1 \\ n \neq m}}^N \sum_{m=1}^N E[|a_n|^2 |a_m|^2] = 1. \quad (\text{A7})$$

Since the probability of  $|a_n|^4$  ( $n=1-N$ ) are equal to each other, one can see that  $E[|a_1|^4] = E[|a_2|^4] = \dots = E[|a_N|^4]$ . Similarly, one can see that  $E[|a_1|^2 |a_2|^2] = \dots = E[|a_{N-1}|^2 |a_N|^2]$ . Thus, the Eq. (A7) can be expressed as

$$NE[|a_n|^4] + N(N-1)E[|a_n|^2 |a_m|^2] = 1. \quad (\text{A8})$$

As shown in Fig. A1, assume that  $|a_n|^2 + |a_m|^2 = \varepsilon^2$  ( $\varepsilon$  is an arbitrary constant satisfying  $0 \leq \varepsilon \leq 1$ ) and amplitude of every other components ( $|a_l|, l \neq n, l \neq m$ ) is fixed. Since the eigenvectors are orthogonal to each other, the variation of  $|a_n|$  and  $|a_m|$  are independent of the value of every other components. Then, in this situation, the ratio of  $E[|a_n|^2 |a_m|^2] / E[|a_n|^4]$  can be obtained as follows:

$$E[|a_n|^2 |a_m|^2] / E[|a_n|^4] = \int_0^{2\pi} \varepsilon^2 \cos^2 \theta \sin^2 \theta d\theta / \int_0^{2\pi} \varepsilon^2 \cos^4 \theta d\theta = 1/3, \quad (\text{A9})$$

where  $|a_n|^2$  and  $|a_m|^2$  are replaced with  $\varepsilon^2 \cos^2 \theta$  and  $\varepsilon^2 \sin^2 \theta$  that satisfy the relation  $|a_n|^2 + |a_m|^2 = \varepsilon^2$  ( $0 \leq \varepsilon \leq 1$ ).

Since the above relation is always satisfied regardless of the value of  $\varepsilon$ , from Eqs. (A8) and (A9) one can obtain

$$E[|a_n|^4] = 3/(N(N+2)), \quad (\text{A10a})$$

$$E[|a_n|^2 |a_m|^2] = 1/(N(N+2)). \quad (\text{A10b})$$

By substituting Eqs. (A4), (A6), and (A10) into Eq. (A5a), the variance  $[\sigma^2(J^t)]$  of the performance index subjected to  $S(J^p)$  can be written as

$$\sigma^2(J^t) = \left[ \frac{2(N-1)}{N^2(N+2)} \sum_{n=1}^N \lambda_n^2 - \frac{2}{N^2(N+2)} \times \sum_{\substack{n=1 \\ n \neq m}}^N \sum_{m=1}^N \lambda_n \lambda_m \right] J^p{}^2. \quad (\text{A11})$$

- <sup>1</sup>P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, New York, 1992), Chaps. 10 and 12.
- <sup>2</sup>P. A. Nelson, A. R. D. Curtis, S. J. Elliott, and A. J. Bullmore, "The Active Minimization of Harmonic Enclosed Sound Fields. Theory," *J. Sound Vib.* **117**(1), 1-13 (1987).
- <sup>3</sup>I.-S. Kim and K.-J. Kim, "Active Control of Cross Mode Waves in Rectangular Ducts," *Mech. System Signal Process.* **8**(1), 93-108 (1994).
- <sup>4</sup>T. Morishita, C. Yamaguchi, T. Tanaka, and M. Taki, "Active Control of Noise including Higher-order Acoustic Modes Propagating in a Duct," *Proceedings of Inter-Noise 94, Yokohama, Japan* (1994), pp. 1373-1376.
- <sup>5</sup>R. L. Clark, "Adaptive Feedforward Modal Space Control," *J. Acoust. Soc. Am.* **98**, 2639-2650 (1995).
- <sup>6</sup>C. H. Hansen, M. T. Simpson, and C. T. Wangler, "Application of Genetic Algorithms to Active Noise and Vibration Control," *Proceedings of the 4th International Congress on Sound and Vibration, St. Petersburg, Russia* (1996), pp. 371-388.
- <sup>7</sup>S. Pottie and D. Botteldooren, "Optimal Placement of Secondary Sources for Active Noise Control using a Genetic Algorithm," *Proceedings of Inter-Noise 96, Liverpool* (1996), pp. 1101-1104.
- <sup>8</sup>D. A. Manolas, T. Gialamas, and D. T. Tsalis, "A Genetic Algorithm for the Simultaneous Optimization of the Sensor and Actuator Positions for an Active Noise and/or Vibration Control System," *Proceedings of Inter-Noise 96, Liverpool* (1996), pp. 1187-1191.
- <sup>9</sup>Yakov Ben-Haim, *Robust Reliability in the Mechanical Sciences* (Springer Berlin, 1996), Chap. 3.
- <sup>10</sup>F. Asano, Y. Suzuki, T. Sone, and D. C. Swanson, "Optimization of Control Source Location in Active Control Systems," *Proceedings of ACTIVE 95, Newport Beach* (1995), pp. 489-498.
- <sup>11</sup>Peter Lancaster and Miron Tismenetsky, *The Theory of Matrices with Applications*, 2nd ed. (Academic, New York, 1985), pp. 432-440.
- <sup>12</sup>J. S. Arora, *Introduction to Optimum Design* (McGraw-Hill, New York, 1989).
- <sup>13</sup>S. Haykin, *Adaptive Filter Theory*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1991), Chap. 4.
- <sup>14</sup>Gilbert Strang, *Linear Algebra and Its Applications*, 3rd ed. (Saunders College Publishing, Philadelphia, 1986), pp. 153-165.

# A relaxed condition for “perfect” cancellation of broadband noise in 3D enclosures

Jing Yuan<sup>a)</sup>

Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

(Received 6 May 1999; accepted for publication 20 February 2000)

This paper presents a relaxed condition for “perfect” cancellation of broadband noise in 3D enclosures. On the basis of a truncated modal model, it can be shown that the primary and secondary paths belong to a same subspace if a certain condition is satisfied. There exists a finite impulse response (FIR) filter transfer function vector for perfect cancellation of the primary paths. The analytical result is verified numerically with an active noise control (ANC) system in a 3D rectangular enclosure. The proposed ANC scheme is shown to fit well into the framework of an existing multichannel least-mean squares (LMS) algorithm for adaptive implementation. © 2000 Acoustical Society of America. [S0001-4966(00)00806-7]

PACS numbers: 43.50.Ki [MRS]

## INTRODUCTION

Active noise control (ANC) is a technique that uses actuators to synthesize destructive interference at selected locations in a sound field.<sup>1</sup> It is most effective in low-frequency ranges where the performance of passive noise control methods tends to deteriorate. This reason makes ANC an important alternative to passive noise control methods.

The objective of ANC is to cancel or minimize noise signals in  $n$  sensor locations. Its mathematical model consists of a primary path vector  $P(z)$  and a secondary path matrix  $\mathbf{G}(z)$ . “Perfect” cancellation requires  $\mathbf{G}(z)H(z) + P(z) = 0$ , where  $H(z)$  denotes the ANC transfer function. An ideal solution may be  $H_I(z) = -\mathbf{G}^{-1}(z)P(z)$  that is not necessarily applicable. Stability of  $H_I(z)$  requires a minimum phase  $\mathbf{G}(z)$  that may be true if the error sensors collocate with the actuators. In practice, error sensors are usually placed away from actuators to avoid the near-field effects. With few exceptions,  $\mathbf{G}(z)$  is usually nonminimum phase and  $H_I(s)$  is not stable. Some researchers<sup>2</sup> derive stable controllers to approximate  $H_I(s)$  under proper criteria. Most others adjust the weights of an FIR filter  $H(z)$  to minimize  $\min_{H(z)} \|\mathbf{G}(z)H(z) + P(z)\|$  in the least-mean squares (LMS) sense.<sup>3,4</sup> It is suggested<sup>5</sup> that an IIR filter  $H(z)$  will minimize  $\min_{H(z)} \|\mathbf{G}(z)H(z) + P(z)\|$  with better effects. These are stable approximations of an unstable, albeit exact, solution  $H_I(s)$ , not perfect cancellation solutions.

In this study, perfect cancellation means a stable and exact solution of  $\mathbf{G}(z)H(z) + P(z) = 0$ , which is achievable by an interesting idea.<sup>6</sup> The method, named “MINT” by its authors, depends on an important result in general control theory.<sup>7</sup> If the Smith form of an  $n \times (n+1)$  polynomial matrix  $\mathbf{G}(z)$  is equivalent to  $[\mathbf{I}_n \ O]$ , then there exists a polynomial matrix  $\mathbf{H}(z)$  with a finite degree such that

$$\mathbf{I}_n = \mathbf{G}(z)\mathbf{H}(z), \quad (1)$$

where  $\mathbf{I}_n$  is an  $n \times n$  identity and  $O$  an all-zero vector. An experimental verification of MINT creates a zone of quiet in a wild frequency range.<sup>8</sup> For ease of reference, the condition on the Smith form of  $\mathbf{G}(z)$  is called the MINT assumption here.

This study uses a truncated modal model to analyze 3D reverberant sound fields. It reveals that the MINT assumption is not easy to satisfy for  $n > 2$  in a 3D reverberant field, though it is generally true for  $n = 1$ . Perfect cancellation is still possible even if the MINT assumption is not valid. The reason is rather simple:  $P(z)$  is not an arbitrary vector. It is subject to the same acoustical constraints as  $\mathbf{G}(z)$ . Under a relaxed condition,  $\mathbf{G}(z)$  will consist of a sufficient number of basis to span a polynomial subspace that includes  $P(z)$ . A multichannel filtered- $x$  algorithm<sup>9</sup> may be applied to implement an FIR filter  $H(z)$  such that  $P(z) + \mathbf{G}(z)H(z) = 0$  for all  $z$ . Contrary to MINT, this approach may be called SCBN for “subspace cancellation of broadband noise” in 3D enclosures. The corresponding condition may be called the SCBN assumption as well.

## I. MATHEMATICAL MODEL

Let  $\mathbf{x}$  denote a 3D coordinate; then, a pressure signal, measured at  $\mathbf{x}$ , can be expressed as

$$y(t, \mathbf{x}) = \sum_{i=1}^{\infty} \alpha_i(t) \phi_i(\mathbf{x}), \quad (2)$$

where  $\phi_i$  is a spatial mode function and  $\alpha_i$  the temporal mode magnitude. The Laplace transform of  $\alpha_i$  is the response of a second-order subsystem

$$\alpha_i(s) = \frac{\kappa_{ij}}{s^2 + 2\xi_i \omega_i s + \omega_i^2} u_j(s), \quad (3)$$

where  $\omega_i$  and  $\xi_i$  are, respectively, the  $i$ th resonance frequency and damping ratio;  $u_j(s)$  is a source. The spatial source effect on the  $i$ th mode is evaluated by

<sup>a)</sup>Electronic mail: mmjyuan@polyu.edu.hk

$$\kappa_{ij} = \int \int_V \int \phi_i(\mathbf{x}) f_j(\mathbf{x}) dv, \quad (4)$$

where  $f_j(\mathbf{x})$  describes the spatial magnitude of the source. The integral reduces to  $\kappa_{ij} = \phi_i(\mathbf{x}_j)$  if the source can be approximated by a point excitation at location  $\mathbf{x}_j$ .

### A. Single-path model

Since ANC is most effective in low-frequency ranges, a measured signal is usually low-pass filtered by a sharp filter  $F(s)$  to cut off the high-frequency components. Let  $y_k(s)$  denote the pressure signal sensed at  $\mathbf{x}_k$  and filtered by  $F(s)$ ; one can write

$$y_k(s) = \left( \varphi_{kj} + \sum_{i=1}^m \frac{\phi_i(\mathbf{x}_k) \phi_i(\mathbf{x}_j)}{s^2 + 2\xi_i \omega_i s + \omega_i^2} \right) u_j(s) \\ = g_{kj}(s) u_j(s),$$

where  $g_{kj}(s)$  denotes the transfer function from source location  $\mathbf{x}_j$  to sensor location  $\mathbf{x}_k$ . It is derived from Eqs. (2), (3), and (4), assuming a point excitation. The 0th mode of  $g_{kj}(s)$ , denoted by

$$\varphi_{kj} = \sum_{i=m+1}^{\infty} \frac{\phi_i(\mathbf{x}_k) \phi_i(\mathbf{x}_j)}{\omega_i^2} \\ \approx \sum_{i=m+1}^{\infty} \frac{\phi_i(\mathbf{x}_k) \phi_i(\mathbf{x}_j)}{s^2 + 2\xi_i \omega_i s + \omega_i^2} F(s),$$

is due to high-order modes in the passband of the sharp low-pass filter.<sup>10</sup>

In control theory, cancellation of the 0th mode is a trivial problem. One may ignore  $\varphi_{kj}$  in  $g_{kj}(s)$  without losing generality while keeping a better focus on the main problem. The path model is therefore simplified to

$$g_{kj}(s) = \sum_{i=1}^m \frac{\phi_i(\mathbf{x}_k) \phi_i(\mathbf{x}_j)}{s^2 + 2\xi_i \omega_i s + \omega_i^2} = \frac{g_{kj}^*(s)}{D(s)}, \quad (5)$$

where  $D(s)$  is the system characteristic polynomial;  $g_{kj}^*(s)$  is the numerator polynomial of  $g_{kj}(s)$ . Equation (5) implies

$$g_{kj}^*(s) = \sum_{i=1}^m \phi_i(\mathbf{x}_k) \phi_i(\mathbf{x}_j) \theta_i(s), \quad (6)$$

where

$$\theta_i(s) = \prod_{l \neq i}^m (s^2 + 2\xi_l \omega_l s + \omega_l^2) \quad \text{for } 1 \leq i \leq m. \quad (7)$$

These  $m$  scalar polynomials form a polynomial diagonal matrix

$$\Theta(s) = \begin{bmatrix} \theta_1(s) & & & \\ & \theta_2(s) & & \\ & & \ddots & \\ & & & \theta_m(s) \end{bmatrix} = \sum_{l=0}^{2m-2} \Theta_l s^l, \quad (8)$$

where the  $i$ th diagonal element of  $\Theta_l$  is the  $l$ th-order coefficient of  $\{\theta_i(s)\}_{i=1}^m$ .

The mode functions may be spatially sampled at a coordinate  $\mathbf{x} = [xyz]$  to form a mode vector

$$\Phi^T(\mathbf{x}) = [\phi_1(\mathbf{x}) \ \phi_2(\mathbf{x}) \ \cdots \ \phi_m(\mathbf{x})]. \quad (9)$$

Equations (8) and (9) allow one to express Eq. (6) in a concise form

$$g_{kj}^*(s) = \Phi^T(\mathbf{x}_k^s) \Theta(s) \Phi(\mathbf{x}_j^a), \quad (10)$$

where superscripts “ $s$ ” and “ $a$ ” indicate that the mode vectors are sampled at sensor location  $\mathbf{x}_k^s$  and actuator location  $\mathbf{x}_j^a$ , respectively.

### B. Multiple-path model

Generally, an ANC system has  $n_s$  sensors and  $n_a$  actuators. Locations of actuators and sensors affect the transfer functions in terms of an actuation matrix

$$\mathbf{M}_a = \begin{bmatrix} \phi_1(\mathbf{x}_1^a) & \phi_1(\mathbf{x}_2^a) & \cdots & \phi_1(\mathbf{x}_{n_a}^a) \\ \phi_2(\mathbf{x}_1^a) & \phi_2(\mathbf{x}_2^a) & \cdots & \phi_2(\mathbf{x}_{n_a}^a) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_m(\mathbf{x}_1^a) & \phi_m(\mathbf{x}_2^a) & \cdots & \phi_m(\mathbf{x}_{n_a}^a) \end{bmatrix} \\ = [\Phi(\mathbf{x}_1^a) \ \Phi(\mathbf{x}_2^a) \ \cdots \ \Phi(\mathbf{x}_{n_a}^a)], \quad (11)$$

and a sensor matrix

$$\mathbf{M}_s^T = \begin{bmatrix} \phi_1(\mathbf{x}_1^s) & \phi_2(\mathbf{x}_1^s) & \cdots & \phi_m(\mathbf{x}_1^s) \\ \phi_1(\mathbf{x}_2^s) & \phi_2(\mathbf{x}_2^s) & \cdots & \phi_m(\mathbf{x}_2^s) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_{n_s}^s) & \phi_2(\mathbf{x}_{n_s}^s) & \cdots & \phi_m(\mathbf{x}_{n_s}^s) \end{bmatrix} = \begin{bmatrix} \Phi^T(\mathbf{x}_1^s) \\ \Phi^T(\mathbf{x}_2^s) \\ \cdots \\ \Phi^T(\mathbf{x}_{n_s}^s) \end{bmatrix}. \quad (12)$$

The system has  $n_s \times n_a$  transfer functions representing paths from  $n_a$  actuators to  $n_s$  sensors. Similar to the derivation of Eqs. (5) and (10), one can express these transfer functions in a concise  $n_s \times n_a$  matrix

$$\mathbf{G}(s) = \frac{1}{D(s)} \mathbf{M}_s^T \Theta(s) \mathbf{M}_a. \quad (13)$$

The effect of the primary source can be modeled by  $\kappa_p^T = [\kappa_{1p} \ \kappa_{2p} \ \cdots \ \kappa_{mp}]$ . The  $i$ th element of  $\kappa_p$  is an integral similar to Eq. (4), with  $f_j(\mathbf{x})$  replaced by  $f_p(\mathbf{x})$ . Since the primary source is not necessarily a point excitation,  $\kappa_p$  is not necessarily  $\Phi$  sampled at a specific coordinate. Let  $P(s)$  denote the primary path vector; then, it differs from Eq. (13) in terms  $\kappa_p$ . The sensor matrix  $\mathbf{M}_s^T$  remains unchanged because the same set of sensors measure a sound field due to both primary source and secondary actuators. It then follows

$$P(s) = \frac{1}{D(s)} \mathbf{M}_s^T \Theta(s) \kappa_p. \quad (14)$$

Combining Eqs. (13) and (14), one expresses the sensor signals as

$$Y(s) = \frac{1}{D(s)} \mathbf{M}_s^T \Theta(s) [\kappa_p r(s) + \mathbf{M}_a U(s)],$$

where  $r(s)$  is a scalar transfer function of the primary source and  $U(s)$  is a vector transfer function of  $n_a$  secondary actuators. Cancellation errors are measured by the sensors and expressed in Laplace transform domain as an  $n_s$ -dimensional vector  $Y(s)$ .

### C. Definition of perfect cancellation

Since the primary path vector and the secondary path matrix share the same denominator  $D(s)$ , only the numerators are of interest here. If the actuator signals are synthesized by  $U(s) = H(s)r(s)$ , then the problem becomes

$$\mathbf{M}_s^T \Theta(s) [\kappa_p + \mathbf{M}_a H(s)] = 0. \quad (15)$$

It defines the mathematical meaning of perfect cancellation of the primary paths.

The analysis is based on the modal theory that does not take pure propagation delays into explicit account. Instead, a modal model uses infinite modes to describe a reverberant field collectively. While mode truncation introduces inevitable errors, the theory is reasonably accurate at predicting path transfer functions in low-frequency ranges when mode density is sparse.

One may be concerned with the sharp low-pass filter as it introduces a phase shift to Eq. (5). This is not a problem when the filters are only applied to the error sensors. The filter transfer function  $F(s)$  need not be included in the model since it is a common multiplier to all paths and  $F(s) \neq 0$  in the frequency range of interest. Equation (5) uses  $|F(s)|$  to justify mode truncation in the passband. The truncated model is inaccurate in the transition and stop bands that are not of interest here. The reference signal  $r(s)$  does not pass the sharp low-pass filter and hence avoids phase distortion.

## II. CONDITIONS FOR PERFECT CANCELLATION

This section shows the existence of a polynomial solution  $H(s)$  for Eq. (15) under a relaxed condition. For better focus, the reference signal is assumed available. Otherwise, it will involve more problems to be addressed in Sec. IV.

### A. System assumptions

If there are  $m$  properly placed actuators to ensure a square and full rank  $\mathbf{M}_a$ , then there exists a vector  $H$  such that  $\kappa_p + \mathbf{M}_a H = 0$  to achieve perfect cancellation. This, however, is not the case to be studied here. A more general problem is to design an ANC when the number of actuators is smaller than the number of modes. Hence, the following assumption:

*Assumption 1:* Rank  $(\mathbf{M}_s^T) = n$  and rank  $(\mathbf{M}_a) = n + r^* < m$  for an integer  $r^* > 0$ .

Generally,  $\kappa_p + \mathbf{M}_a H = 0$  is not possible when assumption 1 holds. An ANC transfer function should be a polynomial vector  $H(s)$  of degree  $n_0$ . Equation (15) becomes

$$P^*(s) = -\mathbf{G}^*(s)H(s), \quad (16)$$

where  $P^*(s) = \mathbf{M}_s^T \Theta(s) \kappa_p$  and  $\mathbf{G}^*(s) = \mathbf{M}_s^T \Theta(s) \mathbf{M}_a$  are numerator parts of transfer functions  $P(s)$  and  $\mathbf{G}(s)$ , respectively.

### B. Matrix form of polynomial products

To solve Eq. (16), a matrix form<sup>8</sup> of polynomial products is adopted here. It expresses a polynomial product  $C(s) = A(s)B(s)$  as

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n_1+n_2} \end{bmatrix} = \underbrace{\begin{bmatrix} a_0 & & & & \\ & a_1 & a_0 & & \\ & \vdots & a_1 & \ddots & \\ & a_{n_1} & \vdots & \ddots & a_0 \\ & & a_{n_1} & \ddots & a_1 \\ & & & \ddots & \vdots \\ & & & & a_{n_1} \end{bmatrix}}_{n_2+1 \text{ columns}} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n_2} \end{bmatrix},$$

where

$$A(s) = \sum_{i=0}^{n_1} a_i s^i$$

and

$$B(s) = \sum_{i=0}^{n_2} b_i s^i.$$

Coefficients of  $A(s)$  contribute to the above matrix that has a row size  $n_1 + n_2 + 1$  and a column size  $n_2 + 1$ . Coefficient vector of  $B(s)$  multiplies to coefficient matrix of  $A(s)$  to produce coefficient vector of  $C(s)$ .

Without excessive use of symbols, one may express the above matrix form as  $C = \mathbf{A}B$  with  $C$  denoting coefficient vector of  $C(s)$ ,  $\mathbf{A}$  representing coefficient matrix of  $A(s)$  and  $B$  coefficient vector of  $B(s)$ . By the same token, the matrix form of Eq. (16) is denoted as  $P^* = -\mathbf{G}^*H$ . Detailed expressions of  $P^*$  and  $\mathbf{G}^*$  read, respectively,

$$P^* = \begin{bmatrix} P_0^* \\ P_1^* \\ \vdots \\ P_{2m-2}^* \\ O \\ \vdots \\ O \end{bmatrix}$$

and

$$\mathbf{G}^* = \underbrace{\begin{bmatrix} \mathbf{G}_0^* & & & & \\ \mathbf{G}_1^* & \mathbf{G}_0^* & & & \\ \vdots & \mathbf{G}_1^* & \ddots & & \\ \mathbf{G}_{2m-2}^* & \vdots & \ddots & \mathbf{G}_0^* & \\ & \mathbf{G}_{2m-2}^* & \ddots & \mathbf{G}_1^* & \\ & & \ddots & \vdots & \\ & & & & \mathbf{G}_{2m-2}^* \end{bmatrix}}_{n_0+1 \text{ blocks}},$$



(17)

where  $\{P_i^* = \mathbf{M}_s^T \Theta_i \kappa_p\}_{i=0}^{2m-2}$ ,  $\{G_i^* = \mathbf{M}_s^T \Theta_i \mathbf{M}_a\}_{i=0}^{2m-2}$  and  $\Theta_i$  is the  $i$ th-order coefficient of  $\Theta(z)$ .

A subspace analysis is conducted in the Appendix to show that  $P^*$  belongs to the column space of  $\mathbf{G}^*$  when assumption 1 is true and  $n_0$  satisfies Eq. (A5).

### C. MINT and SCBN solutions

Both MINT and SCBN try to solve a finite-degree polynomial  $H(s)$  whose coefficient vector  $H$  satisfies  $P^* = -\mathbf{G}^*H$ . MINT depends on Eq. (1) to match an arbitrary transfer function. It requires a square and full rank  $\mathbf{G}^*$ . With  $n$  error sensors and  $(n+1)$  actuators,  $\mathbf{G}^*$  is square if  $n_0 = (2m-2)n-1$ . Its rank, however, depends on the Smith form of  $\mathbf{G}^*(s)$  (Ref. 7)—a condition derived for general polynomial matrices without any specific application background. The Appendix reveals that  $\mathbf{G}^*$  is not necessarily full rank in a 3D reverberant field. That means the MINT assumption is not necessarily true. Fortunately,  $P^*$  is not an arbitrary vector. It belongs to the column space of  $\mathbf{G}^*$  under a relaxed condition. Physically, this is understandable since all sources are subject to the same wave equation under the same boundary conditions when placed in the same enclosure. Therefore,  $P^* = -\mathbf{G}^*H$  has an exact solution even if  $\mathbf{G}^*$  is neither square nor full rank.

### D. A solution based on a discrete-time model

A continuous-time sound field becomes a hybrid system when subject to the interference of a discrete-time ANC. A well-accepted way<sup>11</sup> to obtain a discrete-time model for a hybrid system is mathematically given by

$$G(z)H(z) = (1-z^{-1})Z\left[\frac{G(s)}{s}\right]H(z),$$

where  $G(s)$  and  $H(z)$  denote, respectively, transfer functions of the continuous-time part and the discrete-time part of a hybrid system. The output of  $H(z)$  excites  $G(s)$  via a first-order hold circuit represented by  $(1-z^{-1})/s$ .

Comparing the two sides of the above equation, one can see  $G(z) = (1-z^{-1})Z[G(s)/s]$ . Applying this to Eq. (5), one obtains

$$\begin{aligned} g_{kj}(z) &= (1-z^{-1}) \sum_{i=1}^m Z\left[\frac{\phi_i(\mathbf{x}_k^s)\phi_i(\mathbf{x}_j^a)}{s(s^2 + 2\xi_i\omega_i s + \omega_i^2)}\right] \\ &= \sum_{i=1}^m \frac{\phi_i(\mathbf{x}_k^s)\phi_i(\mathbf{x}_j^a)(\eta_i z^{-1} + \rho_i z^{-2})}{[1 - 2z^{-1}e^{-T\sigma_i}\cos(\beta_i T) + e^{-2T\sigma_i}z^{-2}] \omega_i^2} \\ &= \frac{g_{kj}^*(z)}{D(z)}, \end{aligned} \quad (18)$$

where  $\sigma_i = \xi_i \omega_i$ ,  $\beta_i = \omega_i \sqrt{1 - \xi_i^2}$ ,  $\eta_i = 1 - e^{-T\sigma_i} \cos(\beta_i T) + e^{-T\sigma_i}(\sigma_i/\beta_i)\sin(\beta_i T)$ ,  $\rho_i = e^{-2T\sigma_i} + e^{-T\sigma_i}(\sigma_i/\beta_i)\sin(\beta_i T) - e^{-T\sigma_i} \cos(\beta_i T)$  and  $T$  is the sampling interval. The characteristic polynomial is given by

$$D(z) = \prod_{i=1}^m [1 - 2z^{-1}e^{-T\sigma_i}\cos(\beta_i T) + e^{-2T\sigma_i}z^{-2}].$$

Similar to analog systems, the numerator of Eq. (18) has a concise form  $g_{kj}^*(z) = \Phi^T(\mathbf{x}_k^s)\Theta(z)\Phi(\mathbf{x}_j^a)$ . It uses exactly the same  $\Phi^T(\mathbf{x}_k^s)$  and  $\Phi(\mathbf{x}_j^a)$  as Eq. (10), but a different  $\Theta(z)$ . The  $i$ th diagonal element of  $\Theta(z)$  is a scalar polynomial

$$\begin{aligned} \theta_i(z) &= \frac{(\eta_i z^{-1} + \rho_i z^{-2})}{\omega_i^2} \prod_{l \neq i} [1 - 2z^{-1}e^{-T\sigma_l}\cos(\beta_l T) \\ &\quad + e^{-2T\sigma_l}z^{-2}], \end{aligned}$$

whose degree increases to  $2m$  due to  $(\eta_i z^{-1} + \rho_i z^{-2})/\omega_i^2$  in the above expression. The secondary path matrix and the primary path vector become, respectively,

$$\mathbf{G}(s) = \frac{1}{D(z)} \mathbf{M}_s^T \Theta(z) \mathbf{M}_a \quad (19)$$

and

$$P(s) = \frac{1}{D(z)} \mathbf{M}_s^T \Theta(z) \kappa_p.$$

The primary paths and secondary paths still share the same denominator  $D(z)$ . The numerators differ from their analog counterparts in terms of polynomial matrix  $\Theta(z)$ .

Perfect cancellation means

$$\mathbf{M}_s^T \Theta(z) [\kappa_p + \mathbf{M}_a H(z)] = 0, \quad (20)$$

which is very similar to Eq. (15), only with a different  $\Theta(z)$ . Such a similarity implies a matrix form of Eq. (20) as  $P^* = -\mathbf{G}^*H$ , where

$$P^* = \begin{bmatrix} P_1^* \\ P_2^* \\ \vdots \\ P_{2m}^* \\ O \\ \vdots \\ O \end{bmatrix} \quad \text{and} \quad \mathbf{G}^* = \underbrace{\begin{bmatrix} \mathbf{G}_1^* & & & & \\ \mathbf{G}_2^* & \mathbf{G}_1^* & & & \\ \vdots & \mathbf{G}_2^* & \ddots & & \\ \mathbf{G}_{2m}^* & \vdots & \ddots & \mathbf{G}_1^* & \\ & \mathbf{G}_{2m}^* & \ddots & \mathbf{G}_2^* & \\ & & \ddots & \vdots & \\ & & & & \mathbf{G}_{2m}^* \end{bmatrix}}_{n_0+1 \text{ blocks}}. \quad (21)$$

It can be shown, similar to the Appendix, that  $P^*$  belongs to the column space of  $\mathbf{G}^*$  when assumption 1 is true and  $n_0$  satisfies

$$n_0 = \left\lceil \frac{2m + \min(2n, m) - 3n - r^*}{r^*} \right\rceil \quad (22)$$

which denotes the smallest positive integer larger than or equal to  $[2m + \min(2n, m) - 3n - r^*]/r^*$ . The rank of  $\mathbf{G}^*$  is bounded by

$$\text{rank}(\mathbf{G}^*) \leq 2m + \min(2n, m) + n(n_0 - 2). \quad (23)$$

In the presence of 0th mode, Eqs. (22) and (23) become

$$\begin{aligned} n_0 &= \left\lceil \frac{2m + \min(3n, m) - 3n - r^*}{r^*} \right\rceil \\ \text{rank}(\mathbf{G}^*) &\leq 2m + \min(3n, m) + n(n_0 - 2). \end{aligned} \quad (24)$$

Under these conditions,  $P^* = -\mathbf{G}^*H$  has an exact solution as coefficients of an FIR filter vector.

### E. Adaptive implementation

Like all other adaptive ANC schemes, SCBN depends on a pre-estimated model  $\mathbf{G}(z) = 1/[D(z)]\mathbf{G}^*(z)$  that includes system characteristic polynomial  $D(z)$  and a numerator polynomial matrix  $\mathbf{G}^*(z)$ . The error sensors measure a signal vector  $Y(z)$  to get  $\epsilon(z) = D(z)Y(z)$ . It can be shown that

$$\epsilon(z) = Y_p(z) + \mathbf{G}^*(z)H(z)r(z), \quad (25)$$

where the primary source contributes  $Y_p(z) = P^*(z)r(z)$  and the ANC contributes  $\mathbf{G}^*(z)H(z)r(z)$ . The filtered residue  $\epsilon(z)$  may be used by a proper adaptation law to adjust  $H(z)$ .

An excellent study<sup>9</sup> has been conducted to address the application of LMS to multichannel ANC systems, with a thorough convergence analysis. All left to do here is to show that Eq. (25) fits into the framework studied in Ref. 9.

Let  $G_j^*(z)$  denote the  $j$ th column of matrix  $\mathbf{G}^*(z)$ ; then, Eq. (25) may be expressed as

$$\begin{aligned} \epsilon(z) &= Y_p(z) + \sum_{j=1}^{n+r^*} G_j^*(z)h_j(z)r(z) \\ &= Y_p(z) + \sum_{j=1}^{n+r^*} h_j(z)\Psi_j(z) = Y_p(z) + \mathbf{\Psi}(z)H(z), \end{aligned} \quad (26)$$

where  $\Psi_j(z) = G_j^*(z)r(z)$  represents the  $j$ th column of filtered signals and matrix

$$\mathbf{\Psi}(z) = [\Psi_1(z) \ \Psi_2(z) \ \cdots \ \Psi_{n+r^*}(z)]$$

contains the  $Z$  transforms of all filtered signals. A regressive form is now derived in  $Z$ -transform domain as Eq. (26). It has a time-domain expression of

$$\epsilon(t) = Y_p(t) + [\mathbf{\Psi}(t) \ \mathbf{\Psi}(t-1) \ \cdots \ \mathbf{\Psi}(t-n_0)] \begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_{n_0} \end{bmatrix},$$

where  $Y_p(t)$  and  $\mathbf{\Psi}(t)$  are the time-domain versions of  $Y_p(z)$  and  $\mathbf{\Psi}(z)$ , respectively. Adaptive implementation of SCBN becomes straightforward thereafter.

## III. NUMERICAL VERIFICATION

This section presents a numerical example to verify the theoretical results obtained by Sec. II. All path transfer functions are available analytically by the modal theory.

### A. Model selection

The sound field is enclosed in a rectangle cavity of  $260 \times 64 \times 60$  in.<sup>3</sup> with modes and resonant frequencies available from Ref. 10. The numerical example excludes the 0th mode and uses the next 14 modes. The ANC consists of three sensors and five actuators with  $\text{rank}(\mathbf{M}_s^T) = 3$  and  $\text{rank}(\mathbf{M}_a) = 5$ , respectively. Figure 1 illustrates the sensor and

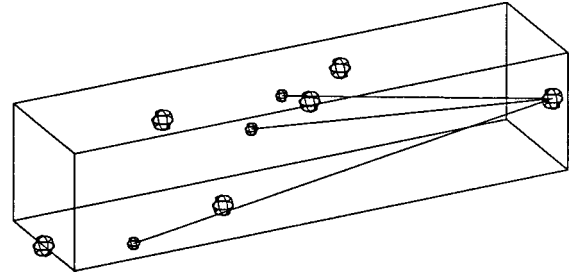


FIG. 1. Locations of sensors and actuators.

actuator locations, where the sensors are represented by small spheres and the actuators are represented by larger ones. Three lines connect the primary source to the sensors to distinguish the primary paths.

Since the SCBN condition is  $\text{rank}(\mathbf{M}_s^T) < \text{rank}(\mathbf{M}_a)$ , one may worry about the accessibility of this criterion when the mode functions are not accessible. Usually, locations of error sensors may be specified by a quiet zone. The rank of  $\mathbf{M}_s^T$  may not be changed even if the mode functions are accessible. However, one always knows  $\text{rank}(\mathbf{M}_s^T) \leq n_s$  without the exact form of  $\mathbf{M}_s^T$ . The problem becomes how to make  $n_s < \text{rank}(\mathbf{M}_a)$ . Generally, an ANC designer has the flexibility to determine the number and locations of actuators. It is not difficult to use a conservatively large number of actuators to meet the requirement, as long as the actuators do not collocate substantially.

In the present example, it is not difficult to get  $\text{rank}(\mathbf{M}_a) = 5$  by placing actuators sufficiently away from each other. The actuator placement is disorganized as shown in Fig. 1. Available information of mode shapes could suggest a more organized actuator placement. This demonstrates the ease to satisfy assumption 1. In Fig. 1, the actuators are closer to the error sensors than the primary source. Therefore, the secondary fields reach the sensors ahead of the primary field. This is an important principle of actuator placement. It should not be difficult to implement in practice.

### B. Cancellation results

Substituting  $m = 14$ ,  $n = 3$ , and  $r^* = 2$  into Eq. (22), the degree of  $\mathbf{H}(z)$  is estimated to be  $n_0 = 12$ . It suggests that  $P^*$  is a  $120 \times 1$  vector and  $\mathbf{G}^*$  is a  $120 \times 65$  matrix. The rank of  $\mathbf{G}^*$  is estimated by Eq. (23) to be 64, which matches exactly what is returned by a MATLAB command “rank( $\mathbf{G}^*$ ).” Both Eq. (22) and Eq. (23) prove very accurate when different numbers of modes are included in the model, with different numbers of sensors and actuators tested at different locations.

A MATLAB command “ $H = -\text{pinv}(\mathbf{G}^*)P^*$ ” solves the coefficient vector of  $H(z)$ . Cancellation effect is evaluated by a residue transfer function vector  $R(z) = P(z) + \mathbf{G}(z)H(z)$ . The primary paths  $P(z)$  are compared with  $R(z)$  sensor by sensor in Figs. 2, 3, and 4, respectively. Residue paths obtained by a standard multichannel filtered- $x$  algorithm are also compared in these figures. The two kinds of residues are labeled as “standard residue” and “SCBN residue,” respectively. The SCBN residues demonstrate reasonably good cancellation effects over a wild frequency

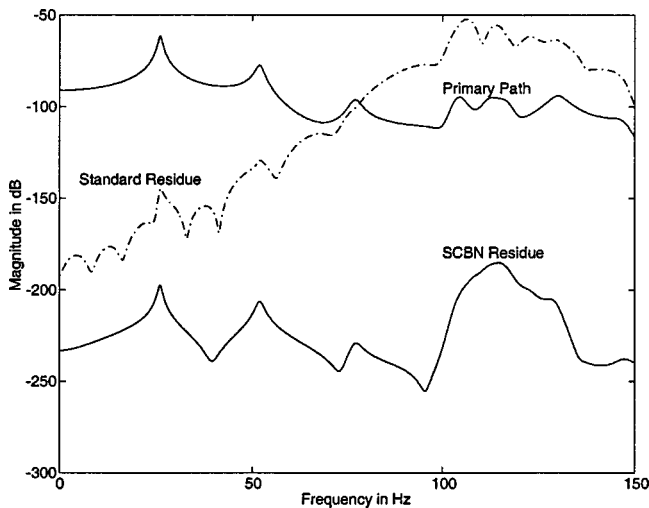


FIG. 2. Primary and residue paths to sensor 1.

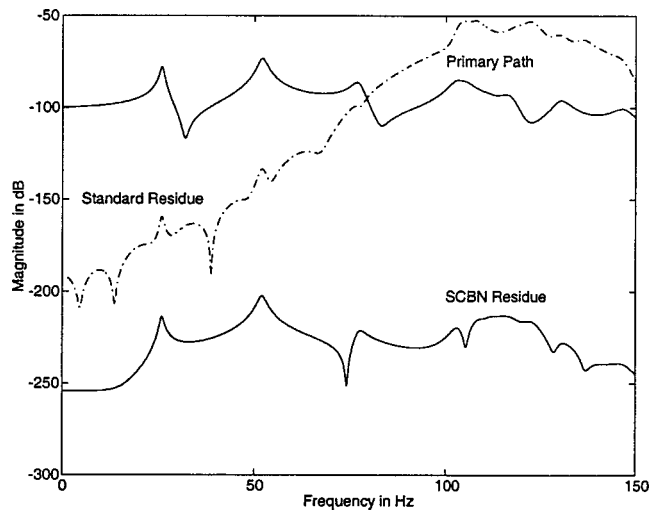


FIG. 4. Primary and residue paths to sensor 3.

range. The standard residues, however, attenuate noise in low-frequency range but enhance noise in high-frequency range.

In this example,  $\mathbf{G}^*$  is a  $120 \times 65$  matrix with rank 64—neither square nor full rank. Therefore, the MINT condition cannot be satisfied under the present situation. It is not analytically clear how the MINT condition can be met in a 3D reverberant sound field. On the contrary, SCBN manages to achieve perfect cancellation under a relaxed condition.

#### IV. RECOVERING THE REFERENCE SIGNAL

In many applications, the reference signal has to be measured by a reference sensor placed near the primary source. The sensor measures signals from the primary source and secondary actuators. It is important to estimate the paths from the actuators to the sensor and cancel the acoustic feedback by a software subtraction.<sup>5,12</sup> This point has been addressed<sup>5,12</sup> thoroughly and hence is not repeated here.

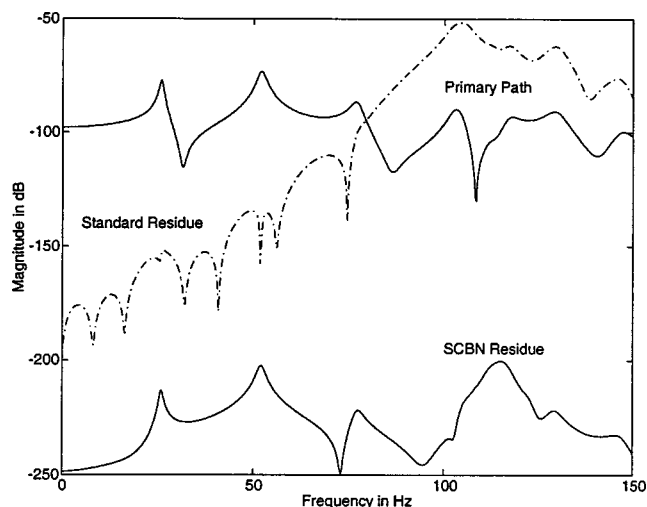


FIG. 3. Primary and residue paths to sensor 2.

#### A. Another problem

There is another problem with the reference signal that also affects the performance of an ANC system. For better focus, one may assume perfect cancellation of acoustic feedback by software subtraction. The sensed signal is then given by  $y_s(z) = [N(z)]/[D(z)]r(z)$ , where  $N(z)$  is the numerator polynomial and  $D(z)$  is the system characteristic polynomial. Perfect recovery of  $r(z)$  calls for an inverse filtering  $r(z) = [D(z)]/[N(z)]y_s(z)$  that may be unstable if  $N(z)$  is nonminimum phase.

#### B. Suggested solution

A possible solution is to place two reference sensors at separate spots near the primary source. The sensed signals are expressed as

$$\begin{aligned} s_1(z) &= D(z)y_{s1}(z) = N_1(z)r(z) \\ s_2(z) &= D(z)y_{s2}(z) = N_2(z)r(z), \end{aligned} \quad (27)$$

where  $N_1(z)$  and  $N_2(z)$  are, respectively, numerators of the transfer functions from the primary source to the reference sensors. The reference signal can be recovered perfectly using FIR filters  $F_1(z)$  and  $F_2(z)$  via  $r(z) = F_1(z)s_1(z) + F_2(z)s_2(z)$ . It is equivalent to

$$1 = N_1(z)F_1(z) + N_2(z)F_2(z), \quad (28)$$

after substituting Eq. (27). This is a special case of Eq. (1) with  $n=1$ . It is also known as the Bezout equation, solvable if  $N_1(z)$  and  $N_2(z)$  are relatively coprime.<sup>13</sup>

#### C. SCBN and the Bezout equation

It is interesting that SCBN also becomes the Bezout equation when  $n=r^*=1$ . In the presence of 0th mode, the degree of  $F_1(z)$  and  $F_2(z)$  is estimated by Eq. (24) to be  $n_0 = 2m - 1$ . Let

$$N_1(z) = \sum_{i=0}^{2m} n_{i,1}z^{-i}$$

and

$$N_2(z) = \sum_{i=0}^{2m} n_{i,2} z^{-i}.$$

A matrix form of

$$\mathbf{G}^*(z) = [N_1(z) \ N_2(z)] = \sum_{i=0}^{2m} [n_{i,1} \ n_{i,2}] z^{-i}$$

can be constructed as

$$\mathbf{G}^* = \underbrace{\begin{bmatrix} \mathbf{G}_0^* & & & \\ \mathbf{G}_1^* & \mathbf{G}_0^* & & \\ \vdots & \mathbf{G}_1^* & \ddots & \\ \mathbf{G}_{2m}^* & \vdots & \ddots & \mathbf{G}_0^* \\ & \mathbf{G}_{2m}^* & \ddots & \mathbf{G}_1^* \\ & & \ddots & \vdots \\ & & & \mathbf{G}_{2m}^* \end{bmatrix}}_{2m \text{ blocks}}$$

$$= \begin{bmatrix} n_{0,1} & n_{0,2} & & & & & & & \\ n_{1,1} & n_{1,2} & n_{0,1} & n_{0,2} & & & & & \\ \vdots & \vdots & n_{1,1} & n_{1,2} & \ddots & & & & \\ n_{2m,1} & n_{2m,2} & \vdots & \vdots & \ddots & n_{0,1} & n_{0,2} & & \\ & & n_{2m,1} & n_{2m,2} & \ddots & n_{1,1} & n_{1,2} & \vdots & \\ & & & & \ddots & \vdots & \vdots & \vdots & \\ & & & & & n_{2m,1} & n_{2m,2} & & \end{bmatrix}.$$

By a proper column permutation, this matrix is equivalent to

$$\mathbf{N} = \begin{bmatrix} n_{0,1} & & & & n_{0,2} & & & & & & \\ n_{1,1} & n_{0,1} & & & n_{1,2} & n_{0,2} & & & & & \\ \vdots & n_{1,1} & \ddots & & \vdots & n_{1,2} & \ddots & & & & \\ n_{2m,1} & \vdots & \ddots & n_{0,1} & n_{2m,2} & \vdots & \ddots & & n_{0,2} & & \\ & n_{2m,1} & \ddots & n_{1,1} & & n_{2m,2} & \ddots & & n_{1,2} & & \\ & & \ddots & \vdots & & & \ddots & & \vdots & & \\ & & & n_{2m,1} & & & & & n_{2m,2} & & \end{bmatrix}.$$

It turns out to be the Sylvester resultant matrix of Bezout equation Eq. (28). Let

$$F_1(z) = \sum_{i=0}^{2m-1} f_{i,1} z^{-i}$$

and

$$F_2(z) = \sum_{i=0}^{2m-1} f_{i,2} z^{-i},$$

then Eq. (28) can be expressed as

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} n_{0,1} & & & & n_{0,2} & & & & \\ n_{1,1} & n_{0,1} & & & n_{1,2} & n_{0,2} & & & \\ \vdots & n_{1,1} & \ddots & & \vdots & n_{1,2} & \ddots & & \\ n_{2m,1} & \vdots & \ddots & n_{0,1} & n_{2m,2} & \vdots & \ddots & & n_{0,2} \\ & n_{2m,1} & \ddots & n_{1,1} & & n_{2m,2} & \ddots & & n_{1,2} \\ & & \ddots & \vdots & & & \ddots & & \vdots \\ & & & n_{2m,1} & & & & & n_{2m,2} \end{bmatrix} \begin{bmatrix} f_{0,1} \\ f_{1,1} \\ \vdots \\ f_{2m-1,1} \\ f_{0,2} \\ f_{1,2} \\ \vdots \\ f_{2m-1,2} \end{bmatrix}$$

$$= \begin{bmatrix} n_{0,1} & n_{0,2} & & & & & & & \\ n_{1,1} & n_{1,2} & n_{0,1} & n_{0,2} & & & & & \\ \vdots & \vdots & n_{1,1} & n_{1,2} & \ddots & & & & \\ n_{2m,1} & n_{2m,2} & \vdots & \vdots & \ddots & n_{0,1} & n_{0,2} & & \\ & & n_{2m,1} & n_{2m,2} & \ddots & n_{1,1} & n_{1,2} & \vdots & \\ & & & & \ddots & \vdots & \vdots & \vdots & \\ & & & & & n_{2m,1} & n_{2m,2} & & \end{bmatrix} \begin{bmatrix} f_{0,1} \\ f_{0,2} \\ f_{1,1} \\ f_{1,2} \\ \vdots \\ f_{2m-1,1} \\ f_{2m-1,2} \end{bmatrix}. \tag{29}$$

The above matrix is nonsingular if  $N_1(z)$  and  $N_2(z)$  are relatively coprime (the MINT assumption). The product is able to match any  $4m$ -dimensional vectors that correspond to polynomials of degree less than or equal to  $4m - 1$ .

The objective of SCBN is matching the right side of Eq. (29) to the numerator of a path transfer function in the same enclosure, but not an arbitrary polynomial. It does not necessarily need a full rank  $\mathbf{G}^*$  as explained in Sec. II; hence, a relaxed condition. By coincident,  $N_1(z)$  and  $N_2(z)$  are found relatively coprime when the two sensors locate sufficiently away from each other. Therefore, both MINT and SCBN merge to the Bezout equation in a special case when  $n=r^*=1$ .

## V. CONCLUSIONS

This study presents SCBN to relax the condition of MINT for perfect noise cancellation in  $n_s$  sensor locations. The SCBN assumption depends on an actuation matrix  $\mathbf{M}_a$  and a sensor matrix  $\mathbf{M}_s^T$ . Perfect cancellation is possible if  $\text{rank}(\mathbf{M}_a) > \text{rank}(\mathbf{M}_s^T)$ .

The principle of SCBN is verified numerically with an exemplary ANC in a 3D rectangular enclosure. It is not difficult to satisfy the SCBN assumption by placing actuators sufficiently away from each other. In a special case when an ANC system consists of two actuators and a single error sensor, both MINT and SCBN are equivalent to a Bezout equation. The principle can be applied to recover the reference signal with two reference sensors.

In a sound field, all acoustical paths are subject to the same physical constraints. The primary paths are therefore acoustically correlated to the secondary paths to a certain degree. SCBN uses a truncated modal model to show complete correlation between the primary and secondary paths. The uncorrelated parts are due to the truncated modes. The method works well in a lightly damped sound field with sparse modes in the frequency range of interest. In highly damped enclosures, the direct field becomes dominant and a separate study is needed to reach a conclusion. A proper placement of actuators, however, will enable secondary fields to reach the sensors ahead of the primary field. Therefore, SCBN is likely to work well in that case.

## ACKNOWLEDGMENTS

The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project No. PolyU 5159/99E). Penetrating remarks by three anonymous reviewers prompted a much improved version of the manuscript.

## APPENDIX

Equation (17) allows one to express  $P^*$  as

$$P^* = \begin{bmatrix} \mathbf{B} \\ O \\ \vdots \\ O \end{bmatrix} k_p \quad \text{where} \quad \mathbf{B} = \begin{bmatrix} \mathbf{M}_s^T \Theta_0 \\ \mathbf{M}_s^T \Theta_1 \\ \vdots \\ \mathbf{M}_s^T \Theta_{2m-2} \end{bmatrix}. \quad (\text{A1})$$

Similarly,  $\mathbf{G}^*$  can be expressed as a product of matrices  $\mathbf{R}$  and  $\mathbf{S}$  as follows:

$$\mathbf{G}^* = \underbrace{\begin{bmatrix} \mathbf{M}_s^T \Theta_0 & & & & \\ \mathbf{M}_s^T \Theta_1 & \mathbf{M}_s^T \Theta_0 & & & \\ \vdots & \mathbf{M}_s^T \Theta_1 & \ddots & & \\ \mathbf{M}_s^T \Theta_{2m-2} & \vdots & \ddots & \mathbf{M}_s^T \Theta_0 & \\ & \mathbf{M}_s^T \Theta_{2m-2} & \ddots & \mathbf{M}_s^T \Theta_1 & \\ & & \ddots & \vdots & \\ & & & & \mathbf{M}_s^T \Theta_{2m-2} \end{bmatrix}}_{\mathbf{R} \text{ with } n_0+1 \text{ shifted } \mathbf{B}\text{-blocks}} \times \underbrace{\begin{bmatrix} \mathbf{M}_a & & & & \\ & \mathbf{M}_a & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \mathbf{M}_a \end{bmatrix}}_{\mathbf{S} \text{ with } n_0+1 \text{ diagonal } \mathbf{M}_a\text{-blocks}}. \quad (\text{A2})$$

A subspace analysis of  $P^*$  and  $\mathbf{G}^*$  leads to a series of lemmas and properties:

*Lemma 1:* Let  $\theta_i$  denote the coefficient vector of polynomial  $\theta_i(s)$ , then  $\{\theta_i\}_{i=1}^m$  are linearly independent when the resonance frequencies  $\{\omega_i\}_{i=1}^m$  are distinct.

*Proof:* Equation (7) suggests  $\theta_i(s) = D(s)/(s^2 + 2\xi_i\omega_i s + \omega_i^2)$  for  $1 \leq i \leq m$ . Let  $p_i = -\xi_i\omega_i - j\omega_i\sqrt{1-\xi_i^2}$ , then  $1/(s^2 + 2\xi_i\omega_i s + \omega_i^2) = -1/(\omega_i\sqrt{1-\xi_i^2})\text{Im}(1/(s-p_i))$ . When the resonance frequencies are distinct,  $\{p_i\}_{i=1}^m$  are distinct as well. As a result, Vandermonde matrix

$$\mathbf{V} = \begin{bmatrix} 1 & p_1 & p_1^2 & \cdots & p_1^{m-1} \\ 1 & p_2 & p_2^2 & \cdots & p_2^{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & p_m & p_m^2 & \cdots & p_m^{m-1} \end{bmatrix}$$

is nonsingular. The  $i$ th row of the above matrix consists of the first  $m$  coefficients of polynomial

$$\frac{1}{(s-p_i)} = \sum_{k=0}^{\infty} p_i^k s^k.$$

Therefore, coefficient vectors of  $\{1/(s-p_i)\}_{i=1}^m$  are linearly independent. The same applies to the coefficient vectors of  $\{\theta_i(s) = -D(s)/(\omega_i\sqrt{1-\xi_i^2})\text{Im}(1/(s-p_i))\}_{i=1}^m$ . Q.E.D.

*Lemma 2:*  $\text{rank}(\mathbf{B}) = m$  [matrix  $\mathbf{B}$  is defined by Eq. (A1)].

*Proof:* Since  $\Theta_l$  is a diagonal matrix,  $\mathbf{M}_s^T \Theta_l$  has  $m$  columns, with the  $i$ th column of  $\mathbf{M}_s^T$  multiplied by the  $i$ th diagonal element of  $\Theta_l$ . Let  $\mathbf{m}_i^s$  denote the  $i$ th column vector of  $\mathbf{M}_s^T$ , then the  $i$ th column of  $\mathbf{B}$  forms a coefficient vector of a polynomial vector

$$\{\mathbf{m}_i^s \theta_i(s)\}_{i=1}^m.$$

Recalling from Eq. (12), one can see that

$$\mathbf{m}_i^{sT} = [\phi_i(\mathbf{x}_1^s) \phi_i(\mathbf{x}_2^s) \cdots \phi_i(\mathbf{x}_n^s)].$$

If  $\|\mathbf{m}_i^s\|$  were zero, then all  $n_s$  sensors would locate at the node planes or node lines of the  $i$ th mode. That mode would not have been included in the truncated model in the first place. On the contrary, a mode is included in a truncated model only if it is measurable by at least one sensor. Therefore,  $\|\mathbf{m}_i^s\| \neq 0$  for all  $1 \leq i \leq m$ ; and the rank of  $\mathbf{B}$  is the same as the rank of polynomials

$$\{\theta_i(s)\}_{i=1}^m.$$

Quoting from Lemma 1, one can see that  $\text{rank}(\mathbf{B}) = m$ . Q.E.D.

A close look at Eq. (A2) indicates that  $\mathbf{R}$  consists of  $n_0 + 1$  blocks of  $\mathbf{B}$ . Counting from left to right, each column block of  $\mathbf{B}$  is downshifted by the row size of  $\mathbf{M}_s^T \Theta_i$ . Therefore, the row size of  $\mathbf{R}$  is the row size of  $\mathbf{B}$  plus  $n_0$  times the row size of  $\mathbf{M}_s^T \Theta_i$ , which is  $n_s(2m - 1 + n_0)$ . The column size of  $\mathbf{R}$  is  $n_0 + 1$  times the column size of  $\mathbf{B}$ ; that is,  $n_0 m + m$ . Hence,  $\mathbf{R}$  is an  $[n_s(2m - 1 + n_0)] \times (n_0 m + m)$  matrix. Evidently, the rank of  $\mathbf{R}$  depends on the integer value of  $n_0$ . When  $n_0 = 1$ ,  $\mathbf{R}$  consists of two blocks of  $\mathbf{B}$ . The second block is downshifted by  $n_s$  rows. These two blocks could be linearly independent, implying

*Property 1:*  $\text{rank}(\mathbf{R}) = \gamma \leq 2m$  when  $n_0 = 1$ .

When  $n_0$  increases to 2,  $\mathbf{R}$  contains one more block of  $\mathbf{B}$  downshifted by  $n_s$  rows. Its rank, however, cannot reach  $3m$ . In order to prove this, one has to recall from Eq. (7) that  $\theta_i(s)(s^2 + 2\xi_i \omega_i s + \omega_i^2) = D(s)$  for  $1 \leq i \leq m$ . This is equivalent to a very useful equation

$$\Theta(s)(s^2 \mathbf{I}_m + \Xi_s + \Omega) = D(s) \mathbf{I}_m, \quad (\text{A3})$$

where  $\mathbf{I}_m$  is an  $m \times m$  identity

$$\Xi(s) = \begin{bmatrix} 2\xi_1 \omega_1 & & & \\ & 2\xi_2 \omega_2 & & \\ & & \ddots & \\ & & & 2\xi_m \omega_m \end{bmatrix}$$

and

$$\Omega(s) = \begin{bmatrix} \omega_1^2 & & & \\ & \omega_2^2 & & \\ & & \ddots & \\ & & & \omega_m^2 \end{bmatrix}.$$

A matrix form is easily available for Eq. (A3) as

$$\begin{bmatrix} d_0 \mathbf{I}_m \\ d_1 \mathbf{I}_m \\ \vdots \\ d_{2m} \mathbf{I}_m \end{bmatrix} = \begin{bmatrix} \Theta_0 & & & \\ \Theta_1 & \Theta_0 & & \\ \vdots & \Theta_1 & \Theta_0 & \\ \Theta_{2m-2} & \vdots & \Theta_1 & \\ & \Theta_{2m-2} & \vdots & \\ & & \Theta_{2m-2} & \end{bmatrix} \begin{bmatrix} \Omega \\ \Xi \\ \mathbf{I}_m \end{bmatrix} \quad (\text{A4})$$

where  $\{d_i\}_{i=0}^{2m}$  are the coefficients of system characteristic polynomial  $D(s)$ . This equation is useful in the proof of

*Lemma 3:*  $\text{Rank}(\mathbf{R}) = k_r = \gamma + n(n_0 - 1) \leq 2m + n(n_0 - 1)$  when  $n_0 \geq 2$ .

*Proof:* To check  $\text{rank}(\mathbf{R})$  for  $n_0 \geq 1$ , one may examine the product of the following two matrices:

$$\mathbf{R}|_{n_0=2} = \begin{bmatrix} \mathbf{M}_s^T \Theta_0 & & & \\ \mathbf{M}_s^T \Theta_1 & \mathbf{M}_s^T \Theta_0 & & \\ \vdots & \mathbf{M}_s^T \Theta_1 & \mathbf{M}_s^T \Theta_0 & \\ \mathbf{M}_s^T \Theta_{2m-2} & \vdots & \mathbf{M}_s^T \Theta_1 & \\ & \mathbf{M}_s^T \Theta_{2m-2} & \vdots & \\ & & \mathbf{M}_s^T \Theta_{2m-2} & \end{bmatrix}$$

and

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_m & \Omega \\ & \mathbf{I}_m & \Xi \\ & & \mathbf{I}_m \end{bmatrix}.$$

Here,  $\mathbf{R}|_{n_0=2}$  represents matrix  $\mathbf{R}$  when  $n_0 = 2$ . It contains three blocks of  $\mathbf{B}$ .  $\mathbf{T}$  is a triangle matrix with nonzero off-diagonal elements only in the last  $m$  columns. This implies  $\text{rank}(\mathbf{R}|_{n_0=2}) = \text{rank}(\mathbf{R}|_{n_0=2} \mathbf{T})$ . Using Eq. (A4), one obtains

$$\mathbf{R}|_{n_0=2} \mathbf{T} = \begin{bmatrix} \mathbf{M}_s^T \Theta_0 & & \mathbf{M}_s^T d_0 \\ \mathbf{M}_s^T \Theta_1 & \mathbf{M}_s^T \Theta_0 & \mathbf{M}_s^T d_1 \\ \vdots & \mathbf{M}_s^T \Theta_1 & \mathbf{M}_s^T d_2 \\ \mathbf{M}_s^T \Theta_{2m-2} & \vdots & \mathbf{M}_s^T d_3 \\ & \mathbf{M}_s^T \Theta_{2m-2} & \vdots \\ & & \mathbf{M}_s^T d_{2m} \end{bmatrix}.$$

The next step is the use of a linear transformation

$$\Gamma_0 = \begin{bmatrix} \mathbf{I}_m & -\Theta_0^{-1} d_0 \\ & \mathbf{I}_m & \\ & & \mathbf{I}_m \end{bmatrix},$$

to get

$$\mathbf{R}|_{n_0=2} \mathbf{T} \Gamma_0 = \begin{bmatrix} \mathbf{M}_s^T \Theta_0 & & & \\ \mathbf{M}_s^T \Theta_1 & \mathbf{M}_s^T \Theta_0 & \mathbf{M}_s^T (d_1 \mathbf{I}_m - \Theta_1 \Theta_0^{-1} d_0) & \\ \vdots & \mathbf{M}_s^T \Theta_1 & \mathbf{M}_s^T (d_2 \mathbf{I}_m - \Theta_2 \Theta_0^{-1} d_0) & \\ \mathbf{M}_s^T \Theta_{2m-2} & \vdots & \mathbf{M}_s^T (d_3 \mathbf{I}_m - \Theta_3 \Theta_0^{-1} d_0) & \\ & \mathbf{M}_s^T \Theta_{2m-2} & \vdots & \\ & & \mathbf{M}_s^T d_{2m} & \end{bmatrix}.$$

It suggests the existence of a series of linear transformations  $\Gamma_1 \cdots \Gamma_{m-1}$ , such that

$$\mathbf{R}|_{n_0=2} = \mathbf{T}\Gamma_0\Gamma_1\cdots\Gamma_{m-1}$$

$$= \begin{bmatrix} \mathbf{M}_s^T \Theta_0 & & & & & & \\ & \mathbf{M}_s^T \Theta_1 & & \mathbf{M}_s^T \Theta_0 & & & \\ & \vdots & & \mathbf{M}_s^T \Theta_1 & & & \\ \mathbf{M}_s^T \Theta_{2m-2} & & & \vdots & & & \\ & & & \mathbf{M}_s^T \Theta_{2m-2} & & & \\ & & & & & & \mathbf{M}_s^T d_{2m} \end{bmatrix}.$$

The above expression suggests  $\text{rank}(\mathbf{R}|_{n_0=2}) = \gamma + n$  by assumption 1 and Property 1. One can apply Eq. (A4) and construct matrices similar to  $\mathbf{T}$  for consecutive increments of  $n_0$ . The result is similar: each increment of  $n_0$  increases  $\text{rank}(\mathbf{R})$  by  $n$ . That proves the lemma. Q.E.D.

*Property 2:*  $P^*$  belongs to the column space of  $\mathbf{R}$ .

This is rather obvious because  $P^*$  is a linear combination of the first  $m$  columns of  $\mathbf{R}$ , as defined by Eq. (A1).

*Property 3:*  $\text{Rank}(\mathbf{S}) = (n + r^*)(n_0 + 1)$ .

From Eq. (A2), one can see that  $\mathbf{S}$  has  $n_0 + 1$  diagonal blocks of  $\mathbf{M}_a$ . Hence,  $\text{rank}(\mathbf{S}) = (n_0 + 1)\text{rank}(\mathbf{M}_a) = (n + r^*)(n_0 + 1)$  by assumption 1.

Since  $\mathbf{G}^* = \mathbf{RS}$ , both  $\text{rank}(\mathbf{R})$  and  $\text{rank}(\mathbf{S})$  affect  $\text{rank}(\mathbf{G}^*)$ . According to Lemma 2 and Property 3,  $\text{rank}(\mathbf{R}) > \text{rank}(\mathbf{S})$  for  $n_0 = 0$ . As  $n_0$  increases,  $\text{rank}(\mathbf{R})$  increases by  $n$  for each increment of  $n_0$  (Lemma 3); whereas  $\text{rank}(\mathbf{S})$  increases by  $n + r^*$  for each increment of  $n_0$  (Property 3). The role of integer  $r^* > 0$  now becomes clear: to make  $\text{rank}(\mathbf{S})$  grow faster than  $\text{rank}(\mathbf{R})$  does. There must exist an integer value of  $n_0$  such that  $\text{rank}(\mathbf{R}) \leq \text{rank}(\mathbf{S})$  for the first time. It is solved by  $n_0 = \lceil (\gamma - 2n - r^*) / r^* \rceil$  that is the small-

est positive integer larger than or equal to  $(\gamma - 2n - r^*) / r^*$ . Substituting  $\gamma \leq 2m$  (Property 1), one obtains an estimate of  $n_0$  as

$$n_0 = \left\lceil \frac{2m - 2n - r^*}{r^*} \right\rceil. \quad (\text{A5})$$

Generally, any integer values of  $n_0$ , if larger than Eq. (A5), will cause  $\text{rank}(\mathbf{R}) \leq \text{rank}(\mathbf{S})$ . Equation (A5) estimates the smallest integer to meet the requirement. It implies

*Lemma 4:* The column space of  $\mathbf{G}^*$  is equivalent to the column space of  $\mathbf{R}$  when assumption 1 is valid and  $n_0$  satisfies Eq. (A5).

Combining Property 2 and Lemma 4, one concludes that  $P^*$  belongs to the column space of  $\mathbf{G}^*$  when assumption 1 is true and  $n_0$  satisfies Eq. (A5).

<sup>1</sup>P. A. Nelson and S. J. Elliot, *Active Control of Sound* (Academic, London, 1992).

<sup>2</sup>M. Bai and H. Chen, *J. Sound Vib.* **198**, 81–94 (1996).

<sup>3</sup>L. J. Eriksson, *J. Acoust. Soc. Am.* **89**, 257–265 (1991).

<sup>4</sup>S. J. Elliot and P. A. Nelson, *IEEE Signal Process. Mag.* **10**, 12–35 (1993).

<sup>5</sup>D. H. Crawford and R. W. Stewart, *J. Acoust. Soc. Am.* **101**, 2097–2103 (1997).

<sup>6</sup>M. Miyoshi and Y. Kaneda, *IEEE Trans. Acoust., Speech, Signal Process.* **36**, 145–152 (1988).

<sup>7</sup>S. Barnett, *IEEE Trans. Autom. Control* **18**, 1–9 (1973).

<sup>8</sup>M. Miyoshi and Y. Kaneda, *Noise Control Eng. J.* **36**, 85–90 (1991).

<sup>9</sup>S. D. Snyder and C. H. Hansen, *J. Sound Vib.* **152**, 157–174 (1992).

<sup>10</sup>R. L. Clark and G. P. Gibbs, *J. Acoust. Soc. Am.* **105**, 2277–2286 (1999).

<sup>11</sup>N. Nise, *Control Systems Engineering* (Benjamin/Cumming, New York, 1995).

<sup>12</sup>M. R. Bai and T. Wu, *J. Acoust. Soc. Am.* **102**, 1004–1012 (1997).

<sup>13</sup>G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control* (Prentice-Hall, Englewood Cliffs, NJ, 1984).

# Community reaction to aircraft noise: Time-of-day penalty and tradeoff between levels of overflights

Henk M. E. Miedema,<sup>a)</sup> Henk Vos, and Ronald G. de Jong  
TNO Prevention and Health, P.O. Box 2215, 2301 CE Leiden, The Netherlands

(Received 20 October 1999; accepted for publication 15 March 2000)

A decrease in the level of sound events can compensate for an increase in the level of other events, but noise metrics assume different tradeoffs. Noise metrics also differ in the penalty applied to noise in the evening and to noise in the night, and in the definition of these periods. These two aspects of noise metrics, i.e., the tradeoff and the penalty for the nighttime (23–7h), are investigated. A general model of the relation between SELs of sound events (aircraft overflights) and noise annoyance is presented which allows for a wide range of tradeoffs and time-of-day penalties. The (tradeoff and time-of-day penalty) parameters of the model are fitted to the data from an aircraft noise study conducted around Amsterdam Airport Schiphol, which is especially suited for investigating the tradeoff and time-of-day penalties. It was found that in this study the tradeoff between the levels of events in metrics based on  $L_{Aeq}$ 's, such as  $L_{Aeq}(24\text{ h})$ , DNL, and DENL, is approximately correct for the prediction of noise annoyance. Furthermore, it was found that the strongest correlation with annoyance is obtained with a nighttime penalty of circa 10 dB. No suitable data were available for further tests of the tradeoff. The result with respect to the nighttime penalty was weakly further supported by the outcome of analyses of the original data from four other aircraft noise surveys (one survey conducted around British airports, and three coordinated surveys carried out around Paris Orly, Amsterdam Schiphol, and Glasgow Abbotsinch). © 2000 Acoustical Society of America.

[S0001-4966(00)05506-5]

PACS numbers: 43.50.Sr, 43.50.Qp [MRS]

## INTRODUCTION

Many different noise metrics have been proposed for the prediction of noise annoyance in communities. Schultz (1982) has given a good overview. Important aspects of noise metrics are their description of the tradeoff between the levels of sound events, and the weighting of noise at different times of the day. This paper investigates the optimal description of the tradeoff and the optimal time-of-day weights for the prediction of noise annoyance.

Noise metrics differ in the reduction of the level of events required to compensate for the increase in the level of other events. The special case of this tradeoff issue concerning the compensation required for the addition of events (i.e., a change from zero level to a nonzero level) is known as the ‘‘noise and number’’ issue. By using  $L_{Aeq}$ -based metrics for the prediction of noise annoyance, it is implicitly assumed that the effect of doubling the number events on annoyance can be exactly offset by a 3-dB reduction of the (SEL) level of events.

Another important aspect of noise metrics is the effect of noise at different times of the day on the metric. Noise metrics differ in the weights (penalties) applied to noise in the evening and to noise in the night, and in the definition of these periods. Examples of  $L_{Aeq}$ -based metrics with different weightings are  $L_{Aeq}(24\text{ h})$  (no weights), DNL (10-dB penalty for noise in the period 22–7h), CNEL (5-dB penalty for 19–22h, and 10-dB penalty for 22–7h), and DENL (5-dB penalty for 19–23h, and 10-dB penalty for 23–7h). A metric

such as DENL, with a 5-dB penalty for a 4-h period and a 10-dB penalty for an 8-h period, has been proposed as the new uniform noise metric for the European Union (EU/DG Environment, 2000).

In this article we analyze data from a large aircraft noise effects study (TNO-PG and RIVM, 1998; code NET-371 in a recent, unpublished update of Fields (1994) catalogue), which was conducted in 1996 around Amsterdam Airport Schiphol. The data are very suitable for the present purposes because of the following reasons.

It is essential for investigating the tradeoff between levels of events that the distribution of the levels of the *individual* sound events is known for *each subject* and can be related to the annoyance reported by this subject. This must be a distribution of *SEL* if the tradeoff assumptions in  $L_{Aeq}$ -based metrics are investigated. In earlier studies only the peak noise levels in PNdB or A-weighted dB averaged over a time and averaged per area was available, instead of a distribution of the (SEL) levels of the individual sound events for each subject. The averaging over time means that an important assumption regarding the tradeoff already had been made and could not be tested (see Sec. II). Consequently, earlier data only have limited relevance for the tradeoff issue. Spatial aggregation of the noise data made it more difficult to find the optimal tradeoff because it makes the estimation of the sound level distributions for individuals less accurate. In contrast, individual distributions of SEL values are determined per respondent for (subperiods of) the periods 7–23h and 23–7h in study NET-371 which is analyzed in this article. A distribution for a period was deter-

<sup>a)</sup>Electronic mail: HME.Miedema@pg.tno.nl



mined on the basis of all overflights in that period during one year.

An important condition for finding reliable estimates of time-of-day penalties is that there is sufficient variation in the time pattern of the noise level. Consider, for example, the extreme case in which the difference between the daytime and the nighttime  $L_{Aeq}$  is constant. Then changing the 10-dB penalty in DNL is equivalent to adding a constant to DNL and, consequently, does not affect the linear correlation with annoyance. Hence, it is not possible to find an optimal penalty that gives a highest correlation with annoyance. If the difference between the daytime and nighttime  $L_{Aeq}$  varies, but only a little, then an optimal penalty can be found, but only with large uncertainty. The data from study NET-371 are suitable for estimating the time-of-day penalty for the nighttime (23–7h) relative to the daytime (7–23h), because the standard deviation of the difference between the levels in these periods is relatively high (the standard deviation of the difference between the  $L_{Aeq}$ 's in these periods is equal to 4.8 dB). The relatively high standard deviation is a consequence of the way Schiphol is operated (nighttime activity concentrated in a part of the area where there is daytime activity) and a sample of 11 142 respondents covering a large area with all 24-h exposure patterns. The results regarding the time-of-day penalties obtained with NET-371 are further tested on the basis of the data from four additional aircraft noise studies.

In addition to the above-mentioned strengths, the study NET-371 has two important specific weaknesses, namely, a low response rate (usual for a mail survey) and uncertainties regarding the calculated SEL (mainly due to uncertainty regarding aircraft noise emissions). However, the impact of the low response rate and possible inaccuracies in SEL on tradeoff and time-of-day penalty estimates were investigated. The possible impact of inaccuracies in the calculated SEL is explored by using the results from a study in which measurements were compared with results from noise levels calculated with the same model and software implementation used in NET-371. The possible impact of the low response rate is explored by using results from a nonresponse study in which 271 persons who did not respond in the main study answered a limited set of questions over the telephone. These additional analyses give an indication of the uncertainty in the estimates of the tradeoff and time-of-day parameters which is related to uncertainty regarding selective response and SEL.

## I. STATE OF THE ART

Fields (1984, 1986a, 1986b, 1986c) conducted secondary analyses concerning both the tradeoff and the time-of-day penalties, which still represent the state of the art on these issues. He used all the relevant original data from the studies available to him at that time, but could not reach firm conclusions on the basis of those data.

Fields (1984) used in the analyses concerning the tradeoff the following generalization of the noise and number index (NNI: see MIL, 1971, and Schultz, 1982):

$$\text{noise index} = L + k \cdot \lg N + C,$$

and estimated the value of the parameter  $k$  that gives the strongest linear relation between the noise index and annoyance measures. Here  $L$  is the “logarithmic average of the peak noise levels” in PNdB or the A-weighted dB, and  $N$  is the number of events. He found estimates for *tradeoff* parameter  $k$  that varied between studies and that contained large standard errors. His main conclusion was that the implicit description of the tradeoff in  $L_{Aeq}$ -based metrics, i.e., the equal energy principle, could not be rejected. Examples of such metrics are  $L_{Aeq}(24\text{ h})$ , DNL, and DENL. The best estimate was that the relative impact of the number of events on annoyance is less than implicitly assumed by  $L_{Aeq}$ -based metrics, but the uncertainty was large. Possible causes of the large uncertainty discussed by Fields are imprecise or incomplete determination of noise exposure, and confounding factors related to study areas.

In addition to this uncertainty, a problem for the interpretation of Fields' (1984) results is that the above class of noise indices is not optimal for investigating whether  $L_{Aeq}$  or a noise index with another tradeoff is the best predictor of annoyance, because of the following reasons. The simplest reason is that  $L_{Aeq}$  is *not* a special case of the above equation when  $L$  is the “logarithmic average of the peak noise levels” in PNdB or the A-weighted dB. It is a special case, obtained with  $k=10$ , only if SEL is used as the descriptor of noise events instead of the “logarithmic average of the peak noise levels.” But then the indices *presuppose* that the tradeoff implicit in  $L_{Aeq}$ , i.e., the equal energy principle, is correct as long as the number of events remains constant. That is, all indices assume that it is not important for the annoyance how a certain amount of sound energy is divided over a fixed number of events. For example, let a situation with 110 events with equal SEL be changed in the following way: the level of 100 events is reduced by 10 dB and the level of 10 events is increased by 10 dB. This change does not affect the value of the noise index: not if  $k$  is equal to 10 (noise index =  $L_{Aeq}$ ), nor if  $k$  has another value. This means that the effect of equal versus unequal distribution of sound energy on annoyance cannot be investigated with the above equation.

Fields (1986b) used in the analyses concerning the relative effect of noise at different times of the day the following generalization of DNL:

$$\text{noise index} = 10 \lg [(T_d/T) \cdot x_d + (T_n/T) \cdot w_n \cdot x_n],$$

where  $x_d$  and  $x_n$  are  $10^{L_{Aeq}/10}$  for the daytime and nighttime, respectively, and  $T_d$ ,  $T_n$ , and  $T$  are the duration of the daytime, the nighttime, and the total period, respectively. He estimated the value of the *nighttime weight*  $w_n$  that gives the strongest linear relation between the noise index and annoyance measures. The studies involved had different definitions of the nighttime and for some studies other metrics were used as approximations to the  $L_{Aeq}$ 's needed.

The analyses had a disappointing outcome. Fields (1986b) concluded: “Available social surveys do not provide valid reliable estimates of the relative weighting for noise levels during different times of the day which will best predict overall (24-h) noise annoyance.” The nearly equal time patterns of the noise level within studies is probably an im-

portant cause of this outcome (see the Introduction). Other probable causes of the disappointing outcome are differences between studies in the definition of the nighttime period, approximation of  $L_{Aeq}$ 's by other metrics and, especially in the case of surface transportation (road traffic, railways), the unknown exposure level of the sleeping room of respondents (see discussion in Sec. VIII).

## II. A MODEL FOR COMBINING SEL OF SOUND EVENTS

In this article we investigate which description of the tradeoff between levels of noise events and which time-of-day penalties are optimal for predicting noise annoyance. The first step in the analysis is the presentation of a general model for combining SEL values of individual sound events with parameter  $\alpha$  for the tradeoff and parameters  $w_k$  for time-of-day weights. This model will be fitted to the data, and the estimates of the parameters give insight in the tradeoff and the time-of-day penalties. The model is a simple tool for investigating both tradeoff and time-of-day issues in the same analysis.

The model of the relation between SEL of individual sound events (e.g., aircraft overflights) and noise annoyance consists of three equations. Two equations define the noise metric  $L_{\alpha,w}$ , and the following linear equation describes the relation between this metric and annoyance:

$$A = p(L_{\alpha,w} - q), \quad (1)$$

where  $A$  is a noise annoyance score (see Sec. III), and the rate of increase of the annoyance  $p$  and annoyance threshold  $q$  are (positive) parameters. A simple linear relationship has been found earlier to give an adequate description if annoyance scores ( $\neq$  percentage highly annoyed) are used with  $L_{Aeq}$  (24 h) or DNL as noise metrics (Miedema, 1992). These latter metrics are special cases of  $L_{\alpha,w}$  as will be shown below.

The composition of metric  $L_{\alpha,w}$  will be discussed in two steps. The first step concerns the combination of SEL's of individual events into a measure,  $L_{Aeq,\alpha}$ , for noise during a period of the day. The combination rule is the same for different periods of the day. The second step concerns the combination of the  $L_{Aeq,\alpha}$ 's for the periods of the day into the overall metric,  $L_{\alpha,w}$ .

A general rule for combining SEL values into a measure of long-term noise exposure  $L_{Aeq,\alpha}$  (which is equal to the common  $L_{Aeq}$  if  $\alpha = 1$ ) is

$$L_{Aeq,\alpha} = 10 \lg [(1/T) \sum_i x_i^\alpha], \quad (2)$$

where  $x_i$  is  $10^{SEL/10}$  of event  $i$  (dB),  $\alpha$  is the positive tradeoff parameter, and  $T$  is the duration of the period (s). The right-hand side of Eq. (2) is a special case of a (logarithmic transformation of a) so-called power sum. Using the approach and results from measurement theory (Krantz *et al.*, 1971; Narens, 1985; Luce *et al.*, 1990), Miedema (1996) discusses basic qualitative properties of power sums that can be empirically tested.

Compared to  $L_{Aeq}$  (i.e.,  $\alpha = 1$ ) the events with the highest SEL have a larger effect on  $L_{Aeq,\alpha}$  if  $\alpha > 1$  and a smaller effect if  $\alpha < 1$ . In other words, the effect on  $L_{Aeq,\alpha}$  of events with the lower SEL is smaller if  $\alpha$  is higher.

The importance of the value for tradeoff parameter  $\alpha$  can be illustrated with a "noise and number" example. Let there be  $N$  aircraft overflights all having equal SEL so that  $x_i = x$ . Then it follows from Eq. (2) that

$$L_{Aeq,\alpha} = 10 \lg (N \cdot x^\alpha) / T = \alpha \cdot 10 \lg x + 10 \lg N - 10 \lg T.$$

This means that  $(10/\alpha) \cdot \lg k$  must be subtracted from the SEL ( $= 10 \lg x$ ) to exactly compensate for a multiplication of the number of overflights by  $k$ . Suppose that the number of overflights becomes twice as high. If  $\alpha = 1$ , as in  $L_{Aeq}$ , then the SEL of overflights must be reduced by  $(10 \lg 2 =) 3$  dB. However, if  $\alpha > 1$ , then the number of overflights has a smaller effect on  $L_{Aeq,\alpha}$  and a smaller reduction of SEL is sufficient. On the other hand, if  $\alpha < 1$ , then the effect of the number of overflights on  $L_{Aeq,\alpha}$  is larger, and a reduction of SEL greater than 3 dB is required in order to compensate that the number of overflights becomes twice as high. Related to this,  $\alpha$  can be conceived of as the parameter that determines the tradeoff between the number of overflights and the SEL value per overflight.

The following rule for combining the  $L_{Aeq,\alpha}$ 's from various periods of the day into the noise measure  $L_{\alpha,w}$  for the whole day is also a special case of the above-mentioned power sum:

$$L_{\omega,w} = 10 \lg [\sum_k (T_k/T) \cdot w_k x_k], \quad (3)$$

where  $x_k$  is  $10^{L_{Aeq,\alpha}/10}$  of period  $k$  (dB),  $w_k$  is a positive time-of-day weight,  $T_k$  is the duration of period  $k$  (s), and  $T$  is the duration of the total period concerned (s).

The time-of-day weights  $w_k$  in Eq. (3) are weights relative to the weight for the daytime, e.g., the latter is equal to 1. Various common metrics can be obtained by setting  $\alpha = 1$  and using the appropriate distinction in periods and values for weights  $w_k$  for these periods. For example, when the 24-h period is divided in the periods 7–22h and 22–7h, and  $w_k$  is set equal to 1 and 10, respectively, for these periods, then  $L_{\alpha,w}$  is equal to DNL. When the 24-h period is divided in the three periods 7–19h, 19–23h, and 23–7h, and  $w_k$  is set equal to 1 ( $= 10^{0/10}$ ), 3.16 ( $= 10^{5/10}$ ), and 10 ( $= 10^{10/10}$ ), respectively, then  $L_{\alpha,w}$  is equal to DENL.

## III. AMSTERDAM AIRPORT SCHIPHOL STUDY, 1996 (NET-371)

The data from a survey conducted in November/December 1996 around Amsterdam Airport Schiphol (TNO-PG and RIVM, 1998) are used for estimating the optimal values of the parameters in the above model [Eqs. (1)–(3)]. Schiphol is the Netherlands' international airport. It is a hub for international and intercontinental transport. Schiphol has five runways, bearing in different directions. In 1997 Schiphol accommodated 368 000 takeoffs and landings, with 31.6 million passengers and nearly 1.2 million tons of air-cargo. To handle this, the airport accommodates around 550 businesses, which employ a total of over 46 000 people.

A sample was drawn from all dwellings (in 1991) within a circle around Schiphol with a radius of 25 km. The sample was stratified according to noise load and distance to the airport. Within the strata the sample of addresses to which a questionnaire concerning annoyance and health was sent by

TABLE I. The number of events in a year per SEL class (calculated values). For three periods of the day the average, minimum, and maximum over 11 142 respondents in NET-371. The total numbers of flights at Schiphol was 319 845.

SEL	19–23h			7–19h			23–7h		
	Minimum	Maximum	Mean	Minimum	Maximum	Mean	Minimum	Maximum	Mean
66–71	406	11070	3122	1075	45871	12299	87	3264	754
71–76	132	9907	2541	395	35951	10752	38	3378	635
76–81	85	11483	2166	253	43313	9502	24	3486	608
81–86	26	9764	1453	113	40536	7004	8	4135	476
86–91	2	11684	512	17	42257	2426	0	5776	137
91–96	0	6310	152	0	23754	632	0	2626	19
96–101	0	2892	36	0	11686	181	0	1164	1
101–106	0	878	4	0	2602	14	0	518	0
106–111	0	931	1	0	2029	2	0	59	0
111–116	0	206	0	0	698	0	0	18	0
116–121	0	0	0	0	1	0	0	0	0

mail was drawn at random. At each address the person of 18 years or older who was the first to have his or her anniversary was requested in an accompanying letter to fill out the questionnaire. After circa a month a letter was sent to those who had not responded to remind them of the questionnaire. A total of 11 812 respondents (response rate: 39%) returned the questionnaire. The dwellings of 670 respondents in the highest exposure zone were heavily insulated against aircraft noise in a special program funded by the government. These respondents have been excluded from the analyses presented here because of this important change in their situation.

Respondents rated, among other aspects, noise annoyance. First they answered the question “How often do you hear the following noise sources at home?” (“Hoe vaak hoort u thuis de volgende geluidbronnen?”). One of the sources to be rated was aircraft. The response categories were “never,” “at least once in the preceding year,” “at least once a month,” “at least once a week,” and “daily.” Except when the response for a source was “never,” the next question “How annoying or not annoying is the noise of the following sources according to you at home?” (“Hoe hinderlijk of niet hinderlijk vindt u bij u thuis het geluid van de volgende bronnen?”) had to be answered for that source. There were 11 numbered response categories with label “not at all annoying” (“helemaal niet hinderlijk”) at category 0 and label “very annoying” (“heel erg hinderlijk”) at category 10.

The annoyance score is obtained by assigning the numbers 4.5, 13.6, ..., 86.4, 95.5 to annoyance categories. The general rule applied is  $\text{score}_{\text{category } i} = 100(i - \frac{1}{2})/m$ , where  $m$  is the number of categories and  $i = 1, \dots, m$  is the rank number of the category. Respondents who never heard aircraft noise were assumed to be not at all annoyed by aircraft noise.

For each of the respondents various noise metrics were calculated with the method that is legally prescribed in the Netherlands for aircraft noise (Rijksluchtvaartdienst, 1996). The following brief description of this method is adapted from Salomons and De Roo (1999). The calculation of the geometrical attenuation and atmospheric absorption is relatively straightforward and more or less standard. The sum of the effects of atmospheric refraction and ground reflection is referred to as the “lateral attenuation” (LAT). The LAT is calculated from two parameters, the distance between the aircraft and the receiver, and the elevation angle of the air-

craft as seen by the receiver. An elevation angle of 0 degrees corresponds with horizontal sound propagation, which means that the aircraft is on the ground. The model predicts a LAT of 18 dB for zero elevation angle and distances larger than 2 km. With increasing elevation angle, the LAT decreases. The LAT is 10 dB for an aircraft at 10-km distance and an elevation angle of 6 degrees, corresponding with an aircraft height of 1 km. For elevation angles of 20 degrees and more, the model predicts a value of zero for the LAT. The LAT in the prescribed method also contains a contribution for screening of sound by aircraft parts.

Input for the calculations were the actual flight data (time, takeoff or landing, type of aircraft, flight path recorded by the flight tracking system) for each individual flight obtained from the airport for the year preceding the survey. In addition to overall noise metrics frequency distributions of SEL were also determined for each dwelling of a respondent for each of the following time periods: 7–19h, 19–22h, 22–23h, 23–6h, and 6–7h. Table I gives the average, minimum, and maximum number of events per SEL class for the periods 7–19h, 19–23h, and 23–7h.

#### IV. WEIGHTS FOR CORRECTION OF SELECTIVE RESPONSE

The percentage highly annoyed persons (i.e., with an annoyance response that exceeds 72 after translation to a 0–100 scale; see Miedema and Vos, 1998) found at a given noise exposure level was on the order of four times higher than the percentage found in other studies conducted around airports (TNO-PG and RIVM, 1998). This is probably partly due to an overrepresentation of annoyed persons in the sample that returned the self-administrative mail questionnaire. Although in principle this selectivity in the response is a problem, it is not likely that an overrepresentation of annoyed persons also affects the relationships investigated in this article. First this hypothesis is investigated by comparing the structure of relationships between variables in NET-371 with the structure of the same variables in other studies.

In order to get an indication whether the relational structure between variables in NET-371 is unusual, the effect of demographic and attitudinal factors on annoyance has been investigated. These variables are used because there are good reference values for their effects. Table II shows that the

TABLE II. Extra noise annoyance (DNL equivalent) associated to categories of demographic and attitudinal factors. The estimates for the Schiphol study NET-371 can be compared with the estimates based on the original data from a number of other studies from various countries (Miedema and Vos, 1999). For NET-371 the same analysis has been used as described in the publication mentioned. DNL was calculated on the basis of the corrected SEL.

Categories		Estimated extra annoyance (DNL equivalent)	
		NET-371	Miedema and Vos (1999)
Sex	Male	0	0
	Female	-0.4	0
Age	15-20	-5.1	-4.0
	20-30 <sup>a</sup>	-2.8	-1.1
	30-40	0	0
	40-50	1.1	0.3
	50-60	-0.1	-0.5
	60-70	-1.2	-1.5
	70+	-3.1	-3.0
Education	1st	-1.5	-0.5
	2nd	0	0
	High	1.8	0.4
	University	1.7	1.4
Size household	1	-1.4	-1.9
	2	0	0
	3	0.9	-0.2
	4	2.1	-0.9
	≥5	0.6	-1.0
Home ownership	Rent	0	0
	Own	2.1	1.6
Dependency	Depend not	0	0
	Depend	-3.6	-2.6
Use	Not	0	0
	Low		-0.7
	Use	-0.9	-1.0
	High		-0.7
Sensitivity	Low	0	0
	Medium	6.4	4.6
	High	8.6	11.2
Fear	Low	0	0
	Medium		11.4 <sup>b</sup>
	High	7.5 <sup>a</sup>	19.5 <sup>b</sup>

<sup>a</sup>Startle or fear in one question.

<sup>b</sup>Upper boundaries.

relations between these variables and annoyance are roughly the same as found in other studies. The most important differences are the different direction of the small effect of a large size of the household, and the smaller (but still large) effect found in the Schiphol study for noise sensitivity and fear. The general agreement indicates that the effect of an overrepresentation of annoyed persons on relational structures in the sample is limited, although of course the evidence in Table II pertains only directly to the relation between annoyance (DNL) and the variables in that table.

A more direct, quantitative exploration of the impact of an overrepresentation of annoyed persons is possible with weights that correct for selective responses. Such weights have been derived from a nonresponse study. Estimates of

the tradeoff and time-of-day parameters will be made with and without these weights. The derivation of weights is discussed in the remainder of this section.

In the nonresponse study a sample of 500 addresses was drawn from the 17 800 addresses with telephone from which no response was received on 31 January 1997. At 21% of these addresses nobody could be reached and at 25% of these addresses cooperation was refused (again). With 271 persons (54%) a brief interview was successfully completed through the telephone.

A weight was calculated for each person in the sample of the main study ( $N=11\,812$ ) as follows:  $\text{weight}=1/p$ , where  $p$  is the estimated proportion of persons with the same characteristics in the total sample of 11 812 respondents and 18 404 nonrespondents. This proportion has been estimated by using the data for the respondents from the main study, and the data for the nonrespondents from the nonresponse study. It was assumed that the 271 persons in the latter sample represent the total group of nonrespondents. Then different logistic regression models that predict the "probability" of being a respondent on the basis of different predictor variables were fitted to these combined data for respondents and nonrespondents (see TNO-PG and RIVM, 1998, for the details). Of the tested models the one with the following variables gave the best prediction of the probability of response: high education or not, born in the Netherlands or not, older than 75 years or not, highly annoyed by aircraft noise or not. This model was applied to the respondents in the main study to estimate the proportion  $p$  of persons with the same characteristics in the total sample of 11 812 respondents and 18 404 nonrespondents. Then  $\text{weight}=1/p$  was calculated for each respondent. Finally, the weights to correct for selective response were obtained in a last step by dividing the weights by the average of all weights.

## V. CORRECTION OF THE CALCULATED NOISE EXPOSURE DATA

Discrepancies between measurements in the vicinity of an airport and measurements in standardized conditions on which aircraft noise emission values are based have already been recognized for a long time (see, e.g., Lane, 1972, 1974). Such discrepancies lead to differences between measured and calculated levels of aircraft noise found, e.g., by Jonkhart (1997).

Jonkhart (1997) found a systematic deviation of calculated SEL from measured SEL. She compared  $L_{Amax}$  measured at four of the permanent measuring points in the vicinity of Schiphol with  $L_{Amax}$  calculated with the same model and software implementation used in the above study. She concluded that measured and calculated  $L_{Amax}$  were equal around 80 dB, and that at lower levels the measured  $L_{Amax}$  was higher: for landings 6 dB and for takeoffs 3 dB if the calculated  $L_{Amax}$  is equal to 65 dB. The relation of the difference with the calculated level is approximately linear for landings as well as for takeoffs. The figures presented by Jonkhart (1997, Fig. 4) indicate that the linear relationships extend above 80 dB, i.e., that above 80 dB the measured level is lower than the calculated level. Jonkhart indicated

that the aircraft noise emission values used in the calculation of the sound levels are the most important cause of the differences with the measured levels. Machet (1999) recently found similar relations between calculated and measured  $L_{Amax}$  near Paris-Orly.

To enable exploration of the impact of inaccuracies in the calculated values, corrections were derived on the basis of the above results from Jonkhart. Because landings and takeoffs could not be distinguished in the database of NET-371 a single linear relation was used for the difference between measured and calculated values which has value 0 at 80 dB and 4.5 at 65 dB, i.e.,  $L_{Amax} - L'_{Amax} = -0.3(L_{Amax} - 80)$ . Here  $L_{Amax}$  is the measured value and  $L'_{Amax}$  the calculated value. This means that the difference as a function of the measured  $L_{Amax}$  is  $L_{Amax} - L'_{Amax} = -(3/7) \cdot (L_{Amax} - 80)$ .

In order to estimate the difference for SEL, this latter equation was combined with the empirical equation for the relationship between (measured) SEL and (measured)  $L_{Amax}$  found by Ollerhead (1992) around airports in the UK:  $L_{Amax} = (SEL - 25)/0.8$ . Furthermore, it was assumed that for an overflight the difference between measured and calculated SEL, and the difference between measured and calculated  $L_{Amax}$  are equal. A basis for this latter assumption is the above-mentioned indication, given by Jonkhart (1997), that the aircraft noise emission values used in the calculations are an important cause of the differences between calculated and measured levels. In many cases, differences in these emission values are expected to have the same effect on  $L_{Amax}$  and SEL. Then it is found that

$$SEL = 0.6512 SEL' + 31.05, \quad (4)$$

where SEL is the corrected value, and SEL' the calculated value. This means that, e.g.,  $SEL' = 68$  dB is adjusted so that  $SEL = 75$  dB, and  $SEL' = 113$  dB is adjusted so that  $SEL = 105$  dB.

The sensitivity of estimates of parameter values in the model specified by Eqs. (1)–(3) to linear transformations of the SEL as in Eq. (4) can be found simply by substituting SEL in Eq. (2) by  $c \cdot SEL + d$ . The parameter  $p$  in Eq. (1) and the time-of-day weights  $w_k$  in Eq. (3) are not affected by such a transformation. The effect of transformation  $SEL = c \cdot SEL' + d$  on the two other parameters is as follows:

$$\alpha' = c \cdot \alpha, \quad (5)$$

$$q' = q - d \cdot \alpha. \quad (6)$$

Equations (5) and (6) give the estimates  $\alpha'$  [Eq. (3)] and  $q'$  [Eq. (1)] for the original, calculated noise data as a function of the parameters  $\alpha$  and  $q$  found with the corrected SEL.

## VI. ESTIMATES OF THE TRADEOFF AND TIME-OF-DAY PARAMETERS

The objective of the analyses is to find optimal values for the tradeoff parameter  $\alpha$  and the time-of-day weights  $w_k$  in the model [Eqs. (1)–(3)] for the prediction of annoyance caused by aircraft noise. Estimates are obtained by fitting the model to the data from NET-371, in which annoyance score [A in Eq. (1)] and (distribution of) SEL values of overflights

TABLE III. The standard deviations of the differences (in dB) between the  $L_{Aeq}$  at various periods of the day of 11 142 respondents in NET-371. A cell value is the standard deviation of the difference between the row and column  $L_{Aeq}$ . The  $L_{Aeq}$ 's are based on the corrected SEL.

	$L_{Aeq}$ (6–7h)	$L_{Aeq}$ (7–19h)	$L_{Aeq}$ (19–22h)	$L_{Aeq}$ (22–23h)	$L_{Aeq}$ (23–6h)
$L_{Aeq}$ (6–7h)		4.8	5.3	4.4	1.1
$L_{Aeq}$ (7–19h)			0.9	1.6	4.8
$L_{Aeq}$ (19–22h)				2.0	5.1
$L_{Aeq}$ (22–23h)					4.4

[ $x_i$  in Eq. (2)] were determined. The midpoints of the SEL classes (see Table I) were used as the values of  $x_i$  in Eq. (2). Before estimates of the parameters could be obtained, a division of the 24-h period into subperiods had to be chosen. This point is discussed first.

Estimates of the  $w_k$  are not stable if the 24-h period is divided into periods with nearly constant differences in sound levels (see the Introduction). Table III gives the standard deviations of the differences between the  $L_{Aeq}$ 's at different periods of the day. It shows that there are two clusters of  $L_{Aeq}$ 's that vary only a little relative to each other, namely, on the one hand,  $L_{Aeq}(7-19h)$ ,  $L_{Aeq}(19-22h)$ , and  $L_{Aeq}(22-23h)$ , and on the other hand  $L_{Aeq}(23-6)$  and  $L_{Aeq}(6-7h)$ . Therefore, subdivisions of the periods 7–23h and 23–7h could not be used, and the 24-h period was divided into these two periods.

The parameters  $(p, q)$ ,  $\alpha$ , and  $w_k$  cannot be estimated with the standard regression technique that finds the optimal weights in a linear combination, because Eqs. (2) and (3) are nonlinear. Instead the parameters are estimated with the iterative Marquardt procedure (cf. Draper, 1981). Like the analytical regression procedure for linear combinations this seeks the values of the parameters that minimize the sum of squared deviations of the predicted values from the observed values. The results in Table IV are obtained with the daytime weight  $w_{7-23h}$  set equal to 1. Results are presented for all four combinations of no SEL correction applied versus SEL correction applied (the ‘‘columns’’ within the cells of Table IV) with no weights applied versus weights for selective response applied (the ‘‘rows’’ within the cells of Table IV).

If tradeoff parameter  $\alpha$  and nighttime weight  $w_{23-07h}$  (and  $p$  and  $q$ ) are optimized (first main row of Table IV), then applying the SEL correction has no impact on the fit of the optimal solution to the data, i.e., the correlation coefficients are not affected by the SEL correction. The SEL correction has no effect on parameter  $p$  and the time-of-day weight  $w_{23-07h}$ . The effect on parameter  $q$  and tradeoff parameter  $\alpha$  is described by Eqs. (4)–(6).

If tradeoff parameter  $\alpha$  and nighttime weight  $w_{23-07h}$  (and  $p$  and  $q$ ) are optimized (first main row of Table IV), then the correlation coefficient is equal to 0.319 if no weights are applied and 0.296 if weights are applied. The range of the four estimates of  $\alpha$  is roughly 0.8–1.4. The influence of the correction of SEL on the estimate is larger than application of weights for selective response. The correlation coefficient hardly decreases ( $\leq 0.002$ ) if tradeoff parameter  $\alpha$  is set equal to 1 and only the nighttime weight  $w_{23-07h}$  is optimized

TABLE IV. Optimal values for the parameters  $p$  and  $q$ , the trade-off parameter  $\alpha$  and the nighttime penalty  $W_{23-07h} = 10 \lg w_{23-07h}$ , when  $W_{07-23h} = 0$  [see Eqs. (1)–(3)], together with their 95% confidence intervals in the first row. In the second row  $\alpha$  is set equal to 1, and in the third row also  $W_{23-07h}$  is fixed at 0. The resulting correlation coefficient  $r$  is given. Each cell gives the results of four analyses, namely, upper left: no weights and no SEL correction; upper right: no weights and SEL correction; lower left: weights and no SEL correction; lower right: weights and SEL correction. The results are based on 10 505 respondents in NET-371 with a valid annoyance score.

Restrictions	$W_{23-07h}$		$\alpha$		$p$		$q$		$r$	
None	9.6±1.5	9.6±1.5	0.82±0.13	1.25±0.20	2.81±0.29	2.81±0.29	17.4±9.6	56.4±16.8	0.319	0.319
	10.8±1.7	10.8±1.7	0.90±0.14	1.39±0.23	2.40±0.26	2.40±0.26	25.3±11.1	68.4±18.0	0.296	0.296
$\alpha=1$	11.3±1.0	8.1±1.1	1		2.50±0.14	3.13±0.18	30.7±1.4	36.3±1.2	0.318	0.318
	11.7±1.0	8.6±1.2			2.27±0.14	2.81±0.18	32.6±1.4	37.6±1.1	0.296	0.294
$L_{Aeq}(24h)$ : $\alpha=1$ & $W_{23-07h}=0$	0		1		2.11±0.13	2.88±0.17	24.6±1.9	33.7±1.3	0.290	0.305
					1.90±0.13	2.57±0.17	26.5±1.9	34.9±1.3	0.267	0.280

(second main row). This indicates that values different from  $\alpha=1$  do not improve the accuracy of the prediction substantially. Consequently, the tradeoff between the levels of events assumed by metrics based on  $L_{Aeq}$ 's, such as  $L_{Aeq}(24h)$ , DNL, and DENL, which assume that  $\alpha=1$ , is approximately correct for the prediction of annoyance from aircraft noise in this study.

If  $\alpha=1$ , then the range of the penalty for the night,  $W_{23-07h} (= 10 \lg w_{23-07h})$ , is roughly 8–12 dB, which encompasses the 10-dB penalty conventionally applied to nighttime noise, e.g., in DNL and DNEL. If no nighttime penalty is applied (third main row) so that  $L_{Aeq}(24h)$  is calculated, the correlation decreases 0.013 to 0.029, depending on SEL correction and weighting. These results indicate that a metric based on the  $L_{Aeq}$ 's for the periods 7–23h and 23–7h gives the best prediction of aircraft noise annoyance if a penalty of circa 10 dB is applied to the nighttime noise.

## VII. FURTHER TESTS OF THE PENALTY FOR NIGHTTIME NOISE

The parameter estimates given in the previous section are the best, unbiased estimates for the population and situation around Schiphol, and their confidence intervals describe the uncertainty in these estimates for the population and situation around Schiphol. In principle, conditions that affect the estimates may be different around other airports. More certainty about the parameters for (large) airports in general can be obtained if data suitable for estimating the parameters are also available for other airports. We are not aware of the existence of other data suitable for estimating the tradeoff parameter  $\alpha$ .

Aircraft noise surveys suitable for further testing of the penalty for nighttime noise were selected from the TNO archive of noise annoyance surveys (see Miedema and Vos, 1998). A survey was considered to be suitable if  $L_{Aeq}(7-23h)$  and  $L_{Aeq}(23-7h)$  were determined in the original study. Studies for which  $L_{Aeq}(7-23h)$  and  $L_{Aeq}(23-7h)$  had to be estimated on the basis of indirect information (e.g., PNL and number of overflights) available for the periods 7–23h and 23–7h were considered to be less suitable but still usable, while studies with exposure data only for other periods were considered to be not suited for the present purpose. Four aircraft noise surveys come into the first category of the most suitable studies: a survey (Brooker and Richmond,

1985) conducted around five British airports, namely Heathrow, Gatwick, Luton, Manchester, and Aberdeen (UKD-242,  $N=1993$ ), and three coordinated surveys (Diamond and Walker, 1986) carried out around Glasgow Abbotsinch, Paris Orly, and Amsterdam Schiphol (UKD-238,  $N=598$ ; FRA-239,  $N=565$ ; and NET-240,  $N=573$ , respectively). The data from UKD-238 for Aberdeen were not used and are not counted above in  $N$  because the situation was exceptional due to the relatively intensive helicopter traffic at that airport. No studies come into the second category. Consequently, it was further tested with the four studies in the first category whether the optimal penalty for nighttime noise is 10 dB.

Because now different studies are used to estimate the nighttime penalty, Eq. (1) is replaced by

$$A = (p + \sum_i p_i S_i) [L_{\alpha,w} - (q + \sum_i q_i S_i)], \quad (1')$$

in which  $S_i$  is an indicator variable which is 1 for cases from dataset  $i$  and otherwise 0. Because there are four studies, three indicator variables  $S_i$  are used in the above equation. These indicators are all equal to 0 for the remaining dataset. Consequently, the parameters  $p$  and  $q$  (intercept and slope) are determined by the cases in the study for which there is no indicator with value 1. The parameters  $p_i$  and  $q_i$  are the additive ‘‘corrections’’ to  $p$  and  $q$ , respectively, which give the intercept and the slope for study  $i$ . Thus, separate linear relationships between noise metric  $L_{\alpha,w}$  and annoyance are fitted for individual datasets.

The nighttime weight  $w_{2307}$  in Eq. (3) [and  $p$ ,  $p_i$ ,  $q$ , and  $q_i$  in Eq. (1)] again was estimated with the iterative Marquardt procedure (parameter  $\alpha=1$  since  $L_{Aeq}$ 's are used). The standard deviations of the differences between  $L_{Aeq}(7-23h)$  and  $L_{Aeq}(23-7h)$  for the four studies are 5.6 dB for UKD-242; 1.6 dB for UKD-238; 8.4 dB for FRA-239; and 0.9 dB for NET-240. This and the above-mentioned number of cases in the studies mean that the following outcome depends strongly on UKD-242 and to a lesser extent on FRA-239. The penalty found is  $W_{23-07h} (= 10 \lg w_{23-07h}) = 10.6$  dB, and the 95% confidence interval is 8.0–13.2 dB. This penalty is close to the ones found with the data from NET-371, and the confidence interval contains the range of value (8–12 dB) previously found. The correlation is equal to 0.468. If no nighttime penalty is applied so that  $L_{Aeq}(24h)$  is calculated, then the correlation decreases slightly to 0.463.

These results further indicate that a metric based on the  $L_{Aeq}$ 's for the periods 7–23h and 23–7h gives the best prediction of aircraft noise annoyance if a penalty of circa 10 dB is applied to the nighttime noise. However, the additional evidence is weak because the difference between the two correlations is small.

## VIII. CONCLUSION AND DISCUSSION

In this article we present results regarding the metrics that give the best prediction of annoyance. Reduction of the value of such a metric in practice will not only reduce the number annoyed persons, but also the number of people that suffer from other noise-induced (health) effects such as sleep disturbance, interference with communication, or hypertension. However, for an optimal assessment and reduction of such other specific effects the use of additional noise metrics may be necessary. In its draft directive on environmental noise, EU/DG Environment (2000) proposes  $L_{Aeq}$  (8-h night) as the metric to be used for the reduction of effects related to noise exposure during the night (quality of the sleep, awakenings, problems to fall asleep, closing of bedroom windows, etc.).

This article focuses on annoyance. It was found that the tradeoff between the levels of events assumed by metrics based on  $L_{Aeq}$ 's is approximately correct for the prediction of annoyance caused by aircraft noise in a large study ( $N = 10495$ ) conducted around Schiphol. Examples of such metrics are  $L_{Aeq}(24h)$ , DNL (similar to  $L_{Aeq}(24h)$ , but with a 10-dB penalty for noise in the period 22–7h), and DENL (with a 5-dB penalty for 7–23h, and a 10-dB penalty for 23–7h). Furthermore, it was found that a metric based on the  $L_{Aeq}$ 's for the periods 7–23h and 23–7h gives the best prediction of aircraft noise annoyance if a penalty of circa 10 dB is applied to the nighttime noise. DNL uses the same 10-dB penalty, but for a longer night period (night period is 22–7 h in DNL). This conclusion regarding the penalty for nighttime noise is based on the Schiphol study, and four additional aircraft noise studies with 3729 cases in total.

The above conclusions are relevant for discussions about the proper metric for aircraft noise and other sources. The outcome supports the use of metrics based on  $L_{Aeq}$ 's and the application of a 10-dB penalty to nighttime noise. *A priori* it appears to be unlikely that the tradeoff and nighttime penalty for surface transportation (road traffic, railways) is different.

To our knowledge there are no existing data that are suitable for testing the assumption that the tradeoff is the same for surface transportation. Testing the assumption that the nighttime penalty is the same for surface transportation is complicated due to the following reason. A significant difference in the exposure levels at different parts of the dwelling is more common for surface transportation than for aircraft noise. Consequently, the noise level caused by surface transportation at the sleeping room of a respondent may be substantially lower than the level at the most exposed facade. Therefore it appears to be necessary to take both the level at the most exposed facade and the level at the sleeping room of the respondent into account when the nighttime penalty is investigated for surface transportation. There are few existing datasets which contain this information and can be used

to investigate the nighttime penalty for surface transportation noise. This is unsatisfactory because the issue of the nighttime penalty for surface transportation is becoming more important due to a change towards a 24-h economy in many societies and the increasing intensity of traffic at night. It is recommended that suitable existing datasets are reanalyzed, and that research is carried out which is specifically designed to test whether 10 dB is the correct nighttime penalty for surface transportation noise. It is advised that in these studies information is collected on the difference in noise level between the most and the least exposed facade, and information on the sound insulation (of the sleeping room) and the usage of windows. Of course, aircraft noise studies with this information would improve the basis for estimating the nighttime penalty for aircraft noise that was available for the analyses in this paper.

## ACKNOWLEDGMENTS

We thank our colleagues James M. Fields, currently at the Kumamoto University in Japan, and Willy Passchier-Vermeer from our institute, whose comments on drafts of this paper contributed to significant improvements.

- Brooker, P., and Richmond, C. (1985). "The United Kingdom Aircraft Noise Index Study: Part I—Main Results," in Proceedings of the Institute of Acoustics 1985 (Spring, York), pp. 323–329.
- Diamond, I. D., and Walker, J. G. (1986). "CEC Joint Research Project: Community Reactions to Aircraft Noise: Final Report," Southampton, University of Southampton/ISVR.
- Draper, N. (1981). *Applied Regression Analysis* (Wiley, New York).
- EU/DG Environment (2000). "Proposal for a European Parliament and Council Directive on the approximation of the laws of the Member States relating to the Assessment and Reduction of Environmental Noise," EU/DG Environment, Brussels, DG ENV Working Draft.
- Fields, J. M. (1984). "The effect of number of noise events on people's reactions to noise: An analysis of existing survey data," *J. Acoust. Soc. Am.* **75**, 447–467.
- Fields, J. M. (1986a). "An evaluation of study designs for estimating a time-of-day noise weighting," Hampton, VA, NASA Report CR-178062.
- Fields, J. M. (1986b). "The relative effect of noise at different times of the day—an analysis of existing survey data," Hampton, VA, NASA Report CR-3965.
- Fields, J. M. (1986c). "Cumulative airport noise exposure metrics: an assessment of evidence for time-of-day weightings," Washington DC, Federal Aviation Administration, Report FAA-EE-86-10, revised November 1989.
- Fields, J. M. (1994). "An updated catalog of 360 social surveys of residents reactions to environmental noise (1943–1993)," Georgia Institute of Technology, Atlanta, GA (later updates are unpublished).
- Jonkhart, S. (1997). "Vergelijking van gemeten en berekende geluidsniveaus," NLR, Amsterdam, The Netherlands, Report CR 97263 L.
- Kranz, D. H., Luce, R. D., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement* (Academic, New York), Vol. I.
- Lane, S. R. (1972). "Measured noise levels from jet aircraft compared to data published by the FAA and Air Transport Association," Los Angeles, UCLA/School of Engineering and Applied Science, UCLA paper ENG-0772.
- Lane, S. R. (1974). "Measured jet noise compared to California noise codes and health limits," *J. Acoust. Soc. Am.* **56**, 1805–1810.
- Luce, R. D., Kranz, D. H., Suppes, P., and Tversky, A. (1990). *Foundations of Measurement: Representation, Axiomatization and Invariance, Vol. III* (Academic, San Diego).
- Machet, J-M. (1999). "La surveillance du bruit des avions autour des aéroports de Paris," *Acoustique & Technique* **18**, 31–38.
- Miedema, H. M. E. (1992). "Response functions for environmental noise in residential areas," TNO-PG, Leiden, The Netherlands, Report 92.021.
- Miedema, H. M. E. (1996). "Quantification of annoyance caused by environmental noise and odour," TNO-PG, Leiden, The Netherlands.

- Miedema, H. M. E., and Vos, H. (1998). "Exposure-response relationships for transportation noise," *J. Acoust. Soc. Am.* **104**, 3432–3445.
- Miedema, H. M. E., and Vos, H. (1999). "Demographic and attitudinal factors that modify annoyance from transportation noise," *J. Acoust. Soc. Am.* **105**, 3336–3344.
- MIL (Market Investigations Ltd.) (1971). "Second Survey of Aircraft Noise Annoyance Around London (Heathrow) Airport," Her Majesty's Stationary Office, London.
- Narens, L. (1985). *Abstract Measurement Theory* (MIT, Cambridge, MA).
- Ollerhead, J. B., Jones, C. J., Cadoux, R. E. *et al.* (1992). "Report of a field study of aircraft noise and sleep disturbance," Civil Aviation Authority, London.
- Rijksluchtvaardienst. (1996). "Voorschrift voor de berekening van de LAeq-geluidsbelasting in dB(A) ten gevolge van structureel uitgevoerd nachtelijk vliegverkeer," Rijksluchtvaardienst, Den Haag, RLD/BV-02.
- Salomons, E. M., and de Roo, F. (1999). "Modeling of aircraft noise propagation," *Journal van Nederlands Acoustisch Genootschap* **147**, 31–44.
- Schultz, Th. J. (1982). *Community Noise Rating*, 2nd ed. (Applied Science, London, UK).
- TNO-PG and RIVM (1998). "Hinder, slaapverstoring, gezondheids en belevingsaspecten in de regio Schiphol, resultaten van een vragenlijstonderzoek," TNO-PG, Leiden, The Netherlands, Report 98.039/RIVM, Bilthoven, The Netherlands, Report 441520010.



# Spatial-temporal correlation of a diffuse sound field

Boaz Rafaely<sup>a)</sup>

*Institute of Sound and Vibration Research, University of Southampton,  
Southampton SO17 1BJ, United Kingdom*

(Received 21 July 1999; revised 26 January 2000; accepted 17 February 2000)

The spatial correlation function of the sound in a diffuse field is a quantity widely used in many reverberant room acoustic applications. Although results for the spatial and temporal correlation for pure-tone and band-limited diffuse fields have already been developed, these have not been generalized for other signal types. This work presents a generalized derivation of the diffuse field spatial-temporal correlation which can be used for stationary random signals with given power spectral density. It is shown that the spatial-temporal correlation depends entirely on the temporal correlation of the signal exiting the diffuse field, or alternatively on its power spectral density. A simulation using the plane wave model is presented for tonal and broadband diffuse sound fields.  
© 2000 Acoustical Society of America. [S0001-4966(00)00606-8]

PACS numbers: 43.55.Br, 43.55.Cs [JDQ]

## INTRODUCTION

The diffuse sound field model is widely used in the analysis of sound in enclosures. In particular, the spatial and temporal correlation of the sound in a diffuse field is a useful measure that finds applications in areas from reverberation room measurements<sup>1,2</sup> to active control of sound.<sup>3</sup> The spatial and temporal correlation function for a diffuse pure-tone sound field was first derived and tested experimentally by Cook *et al.*<sup>4</sup> A plane wave model was assumed for the sound field, with waves arriving from all directions, and a spatial-averaging calculation was used for the correlation function. It was shown that the spatial correlation behaves as a  $\sin(kr)/kr$  function for waves arriving from three dimensions, with  $k$  and  $r$  denoting the wave number and distance, respectively. In the same work the spatial correlation for an acoustic signal made from a band of frequencies was also computed as the average over frequency of the single-frequency spatial correlation function, and was shown to contain high-order terms of  $kr$ . A more detailed calculation of the correlation function for tonal and band-limited fields can be found in Refs. 2 and 5. Jacobsen<sup>5</sup> presented a comprehensive analysis of the diffuse field and a computation of the correlation function for both spatial and temporal variations.

Although previous work provided useful results for pure-tone and band-limited acoustic signals, no result for the spatial correlation of more general acoustic signals appears to exist. This could be useful in cases where the diffuse field is excited by a broadband random noise signal, for example, with a given power spectral density. This paper presents a derivation of the diffuse field spatial-temporal correlation function for such signals, and shows that the spatial-temporal correlation is an explicit function of the temporal autocorrelation of the signal exiting the diffuse field. The correlation functions for pure-tone and band-limited sound fields are then shown to be special cases of the more general formulation. The paper is concluded with simulations of the spatial correlation of tonal and broadband diffuse sound fields.

## I. GENERALIZED SPATIAL-TEMPORAL CORRELATION

### A. The plane wave model

The plane wave model of a diffuse field assumes an infinite number of plane waves, arriving uniformly from all directions, with random amplitudes and phases.<sup>1,5</sup> The plane wave arriving from the  $n$ -th direction, denoted by  $p_n$ , undergoes pressure fluctuations as a function of both position and time, and is denoted using spherical coordinates as in Fig. 1, as  $p_n(\mathbf{r}, t)$ , where  $\mathbf{r} = (r', \theta, \phi)$ . A simplification is introduced, where only positions along the  $z$ -axis, simply denoted by  $\mathbf{r} = r$ , are considered. This reduces the complexity of the correlation function derivation, as introduced below, and was previously used in Refs. 4 and 5. The correlation results obtained for positions along the  $z$ -axis can be generalized to the entire space assuming the diffuse field is spatially stationary.

The time-domain pressure fluctuations of the  $n$ -th plane wave, when measured at the origin, are denoted by

$$p_n(0, t) = x_n(t). \quad (1)$$

The pressure signal at location  $r$  along the  $z$ -axis can now be written as a function of the pressure signal at the origin, by converting spatial intervals to temporal intervals, as illustrated in Fig. 2. That is, the pressure fluctuations at position  $r$  are equal to the pressure fluctuations at the origin but at a prior time, i.e.,

$$p_n(r, t) = p_n\left(0, t - \frac{r}{c} \cos \theta_n\right) = x_n\left(t - \frac{r}{c} \cos \theta_n\right), \quad (2)$$

where  $c$  is the speed of sound and  $\theta_n$  is the arrival angle of the  $n$ -th plane wave.

The pressure in the diffuse field along the  $z$ -axis can now be represented as an infinite summation of the plane waves arriving from all directions,<sup>1</sup> which is written as<sup>5</sup>

$$p(r, t) = \lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \sum_{n=1}^N p_n(r, t). \quad (3)$$

<sup>a)</sup>Electronic mail: br@isvr.soton.ac.uk

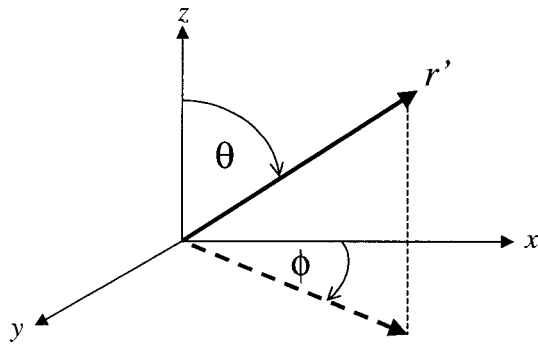


FIG. 1. The spherical coordinate system.

Substituting Eq. (2) in Eq. (3), the pressure in the diffuse field along the  $z$ -axis is written as

$$p(r, t) = \lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \sum_{n=1}^N x_n \left( t - \frac{r}{c} \cos \theta_n \right). \quad (4)$$

This equation is used below to derive the correlation function in the diffuse field. The signal  $x_n(t)$  can be tonal, band-limited, or broadband random, and is assumed to be stationary with an autocorrelation function  $R_0(\tau)$  and power spectral density  $S_0(\omega)$ . It is also assumed that the signal  $x_n(t)$  has the same autocorrelation function for all  $n$ , although signals corresponding to plane waves from different directions are assumed uncorrelated. The assumption that all plane waves carry a signal with the same autocorrelation can be justified by the fact that all plane waves in the diffuse field are excited by the same source(s), while the assumption that plane waves arriving from different directions are uncorrelated can be justified by the spatial behavior of the diffuse field, and has been used previously in similar calculations.<sup>2,4,5</sup>

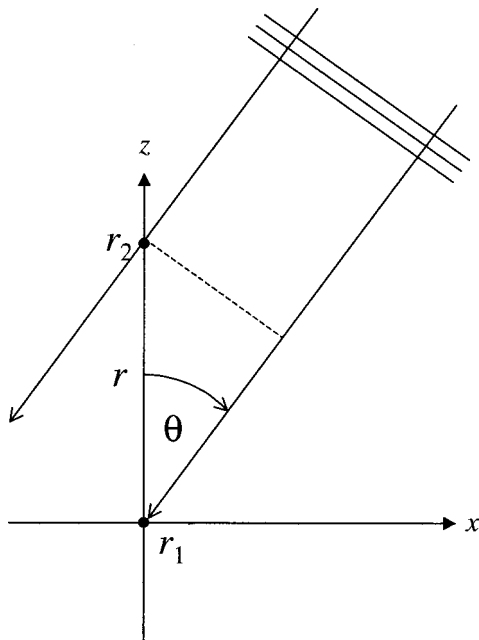


FIG. 2. Incident plane wave arriving at positions  $\mathbf{r}_1$  and  $\mathbf{r}_2$  distance  $r$  apart.

## B. Derivation of the correlation function

The spatial-temporal correlation function of the pressure in a diffuse field is calculated in this section using the plane wave model described above. The correlation function is computed as the expectation of the product of the pressure fluctuations at two positions and times.<sup>1</sup> The expectation operation denotes averaging over many samples of diffuse fields. The spatial-temporal correlation function, denoted by  $R$ , for a diffuse field that is assumed stationary over both time and space, is therefore given by

$$R(r, \tau) = E\{p(r_1, t_1)p(r_2, t_2)\}, \quad (5)$$

where  $r = |r_2 - r_1|$  is a distance along the  $z$ -axis as illustrated in Fig. 2 and  $\tau = |t_2 - t_1|$ . It should be noted that to compute the correlation coefficient,  $\rho(r, \tau) \in [-1, 1]$ , the correlation function is normalized by the power of the signal, which is given by  $R(0, 0)$ . The correlation function can now be written by substituting Eq. (3) in Eq. (5), as

$$R(r, \tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \sum_{m=1}^N E\{p_n(r_1, t_1)p_m(r_2, t_2)\}. \quad (6)$$

Since plane waves from different directions are assumed uncorrelated, only terms in which  $n = m$  contribute, and the correlation function becomes

$$R(r, \tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N E\{p_n(r_1, t_1)p_n(r_2, t_2)\}. \quad (7)$$

Substituting  $r_1 = 0$  and  $r_2 = r$  as illustrated in Fig. 2, together with Eq. (2) in Eq. (7), the correlation function can be written as

$$\begin{aligned} R(r, \tau) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N E\left\{x_n(t_1)x_n\left(t_2 - \frac{r}{c} \cos \theta_n\right)\right\} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N R_0\left(\tau - \frac{r}{c} \cos \theta_n\right), \end{aligned} \quad (8)$$

where  $R_0$  is the autocorrelation of signal  $x_n$ , and is equal for all  $n$ , as discussed above. The infinite summation, in the limit, converges to a double integral over  $\theta$  and  $\phi$ , covering all plane waves' arrival directions. The area element for a sphere with a unit radius in this case is  $\sin \theta d\theta d\phi$ ,<sup>8</sup> which replaces the area element of  $4\pi/N$  in the summation. The correlation function can therefore be written using an integral equation as

$$R(r, \tau) = \frac{1}{4\pi} \int_0^\pi \int_0^{2\pi} R_0\left(\tau - \frac{r}{c} \cos \theta\right) \sin \theta \cdot d\phi \cdot d\theta. \quad (9)$$

Equation (9) suggests that the correlation of the diffuse field can be calculated as the spatial average over  $\phi$  and  $\theta$  of the autocorrelation function of the signal carried by the plane waves. A similar approach was used previously for the calculation of the correlation of a pure-tone diffuse field.<sup>2,4</sup> The result in Eq. (9), however, is more general and is used below for signal types other than pure tone and band limited. The integral in Eq. (9) with respect to  $\phi$  reduces to a constant of  $2\pi$ , while the parameter  $t$  is introduced to replace  $\tau$

$-(r/c)\cos\theta$  and is substituted in the integral with  $dt = r/c \sin\theta d\theta$  to produce the following equation:

$$R(r, \tau) = \frac{c}{2r} \int_{\tau-r/c}^{\tau+r/c} R_0(t) dt. \quad (10)$$

This equation shows that the spatial-temporal correlation of the pressure in a diffuse field is entirely dependent on the temporal autocorrelation of the signal carried by the plane waves composing the diffuse field.

It should be noted that this expression is not well defined for  $r=0$ , so in practice the integral must first be calculated before the value of the correlation at  $r=0$  can be evaluated. A special case of this equation is the correlation dependent on space only, with  $\tau=0$ , which is written as

$$R(r, 0) = \frac{c}{2r} \int_{-r/c}^{r/c} R_0(t) dt. \quad (11)$$

### C. Frequency domain formulation

The temporal autocorrelation of the signal carried by the plane waves can be written in terms of the inverse Fourier transform of the power spectral density of the signal<sup>6</sup> as

$$R_0(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) e^{j\omega\tau} d\omega. \quad (12)$$

Substituting this equation for the autocorrelation in Eq. (10), the spatial-temporal correlation in a diffuse field can be written as

$$\begin{aligned} R(r, \tau) &= \frac{c}{4\pi r} \int_{\tau-r/c}^{\tau+r/c} \int_{-\infty}^{\infty} S_0(\omega) e^{j\omega t} d\omega dt \\ &= \frac{c}{4\pi r} \int_{-\infty}^{\infty} S_0(\omega) \int_{\tau-r/c}^{\tau+r/c} e^{j\omega t} dt d\omega \\ &= \frac{c}{2\pi r} \int_{-\infty}^{\infty} S_0(\omega) \frac{1}{\omega} e^{j\omega\tau} \sin\left(\omega \frac{r}{c}\right) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) \operatorname{sinc}\left(\frac{\omega r}{c}\right) e^{j\omega\tau} d\omega. \end{aligned} \quad (13)$$

This shows that the spatial-temporal correlation is the inverse Fourier transform of the power spectral density of the signal exiting the plane waves when weighted by the spatial sinc function.

An interesting result is derived for the special case of  $r=0$  in this equation, so that the temporal autocorrelation of the acoustic signal in the diffuse sound field becomes

$$R(0, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) e^{j\omega\tau} d\omega = R_0(\tau). \quad (14)$$

This result clearly shows that the temporal correlation of the signal in the diffuse sound field is identical to the temporal correlation of the signal carried by the individual plane waves composing the diffuse field. This result suggests that  $R_0(\tau)$  in Eq. (10) can be replaced by  $R(0, \tau)$ , which is the temporal autocorrelation of the acoustic signal in the diffuse field.

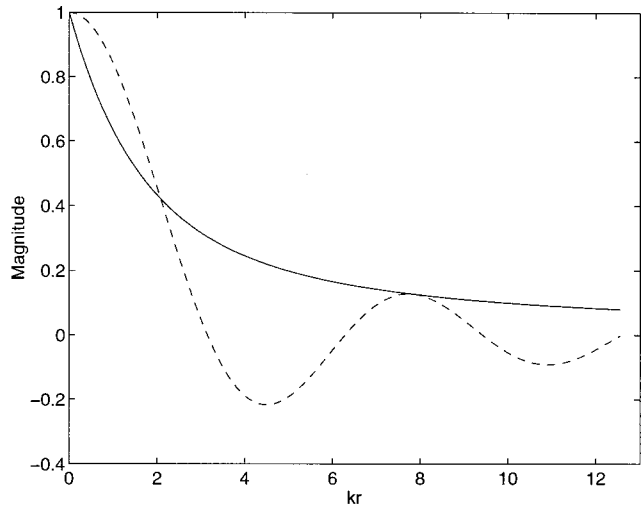


FIG. 3. The spatial correlation of the diffuse field composed of plane waves carrying a low-pass filtered random noise (solid curve), and plane waves carrying a tonal signal at the same frequency as the low-pass filter cutoff frequency (dashed curve).

## II. EXAMPLES

### A. Pure-tone diffuse field

Some special cases can be calculated from the integral equation for the spatial correlation as derived above. For example, if the pressure signal of the plane waves composing the diffuse field is a pure tone at frequency  $\omega$ , with a normalized temporal autocorrelation function given by  $R_0(\tau) = \cos(\omega\tau)$ , the spatial-temporal correlation of the sound in the diffuse field is calculated using Eq. (10) as

$$\begin{aligned} R(r, \tau) &= \frac{c}{2r} \int_{\tau-r/c}^{\tau+r/c} \cos(\omega t) dt \\ &= \frac{c}{\omega r} \sin\left(\omega \frac{r}{c}\right) \cos(\omega\tau) \\ &= \frac{\sin(kr)}{kr} \cos(\omega\tau), \end{aligned} \quad (15)$$

where  $k = \omega/c$ . This result is widely known and has been previously derived for a pure-tone diffuse field.<sup>1,4,5</sup> A plot of this correlation function for  $\tau=0$  as a function of  $kr$  is presented in Fig. 3 (dashed curve).

### B. Band-limited diffuse field

A band-limited random noise is considered next, with a power spectral density of  $2\pi/(\omega_2 - \omega_1)$  at a frequency range between  $\omega_1$  and  $\omega_2$ . For simplicity,  $\tau$  is set to zero, and the spatial correlation can be computed using Eq. (13) as

$$\begin{aligned} R(r, 0) &= \frac{1}{4\pi} \int_{-\infty}^{\infty} S_0(\omega) \operatorname{sinc}(kr) d\omega \\ &= \frac{1}{\omega_2 - \omega_1} \int_{\omega_1}^{\omega_2} \operatorname{sinc}(kr) d\omega. \end{aligned} \quad (16)$$

A similar result and some suggestions for solutions were reported in Refs. 2 and 4. Extending this result to include

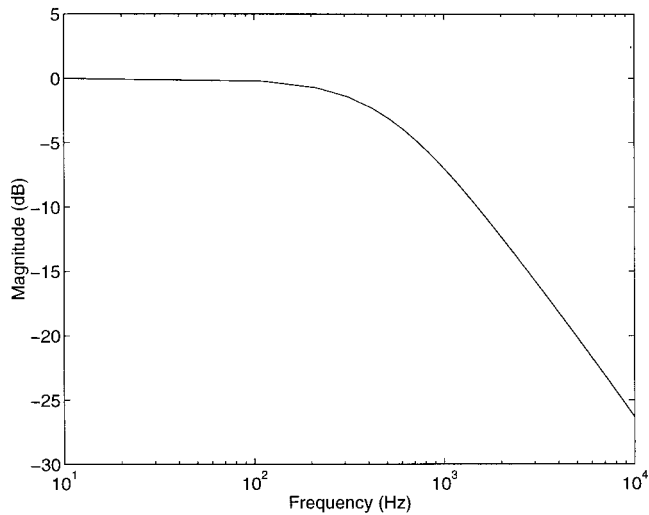


FIG. 4. The magnitude of the frequency response of the low-pass filter used to generate the signal carried by the plane waves.

temporal variations, the corresponding spatial-temporal correlation function can be expressed using Eq. (13) as

$$R(r, \tau) = \frac{1}{\omega_2 - \omega_1} \int_{\omega_1}^{\omega_2} \text{sinc}(kr) e^{j\omega\tau} d\omega. \quad (17)$$

### C. Broadband diffuse field

The formulation derived in this paper provides the means to calculate the spatial-temporal correlation for other pressure signals than pure-tone or band-limited noise. Consider a diffuse field composed of plane waves carrying a signal with a power spectral density of a white noise, which is filtered by a low-pass filter with a single pole at  $\omega_0$  and hence has a frequency response of  $\omega_0/(\omega_0 + j\omega)$ . The power spectral density of the signal when normalized to a unit power is then given by  $2\omega_0/(\omega_0^2 + \omega^2)$ .<sup>6</sup> The autocorrelation, which is the inverse Fourier transform of the power spectral density, is given by  $R_0(\tau) = e^{-\omega_0|\tau|}$  in this case.<sup>6</sup> For simplicity, only the spatial dependence of the correlation function is considered, which is

$$\begin{aligned} R(r, 0) &= \frac{c}{2r} \int_{-r/c}^{r/c} e^{-\omega_0|t|} dt \\ &= \frac{c}{r} \int_0^{r/c} e^{-\omega_0 t} dt \\ &= \frac{(1 - e^{-\omega_0 r/c})}{\omega_0 r/c} = \frac{(1 - e^{-k_0 r})}{k_0 r}, \end{aligned} \quad (18)$$

where  $k_0 = \omega_0/c$ . Since a normalized signal was used, the value of  $R(0,0)$ , found using L'Hopital's rule,<sup>8</sup> is unity in this case. Figure 4 shows the magnitude of the frequency response of the low-pass filter, and Fig. 3 shows the spatial correlation of the diffuse field in this case (solid curve). This is also compared to the correlation of a pure tone diffuse field at a frequency  $\omega_0$  (dashed curve).

The correlation of the diffuse field for other types of signals can be examined using the analysis of the low-pass

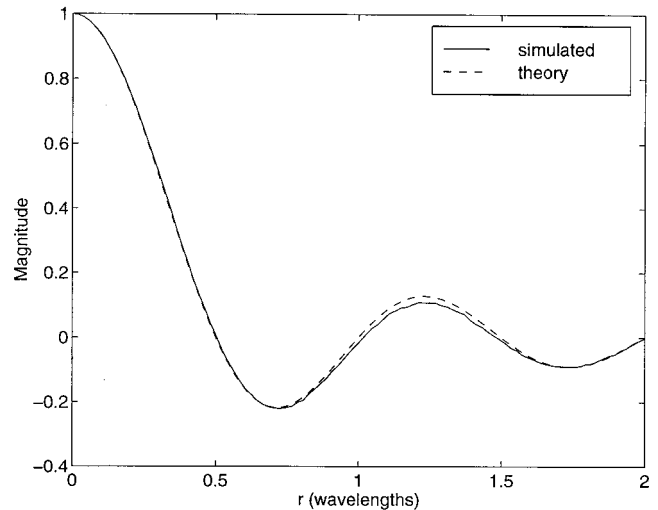


FIG. 5. The spatial correlation of the diffuse field for pure-tone excitation, theoretical result (dashed curve) and simulated result (solid curve).

filtered signal. For example, increasing the cutoff frequency  $\omega_0$  to infinity can approximate a white-noise signal. In this case  $k_0 r$  will be infinitely large for any value of  $r > 0$ , with a corresponding infinitesimally small value for the correlation. The correlation is therefore unity for  $r=0$  and zero elsewhere, making the field spatially uncorrelated for any  $r > 0$ .

### III. SIMULATIONS

Diffuse field simulations were performed in MATLAB to calculate the spatial correlation in tonal and broadband diffuse sound fields. The simulations provide, in addition to numerical verification of the analytic results, an example of a more practical situation where a finite number of plane waves compose the diffuse field. A diffuse field model was used as in Eq. (3), where 1145 spatially uniformly distributed plane waves were superimposed to simulate the diffuse field, with plane waves arriving from different directions designed to be uncorrelated, as detailed below. Spherical coordinates were used to generate the diffuse field, as detailed in the Appendix. Each plane wave carried a signal that was constructed in the time domain as a sampled sequence, using a sampling frequency of 100 kHz to allow for accurate time resolution. The propagation delay of the plane waves was then simulated by shifting the signals an appropriate number of samples. As in the theoretical analysis, the spatial correlation along the  $z$ -axis was chosen for simplicity. The spatial correlation was calculated using Eq. (5), for  $\tau=0$ , normalized by the power of the signal, where the expectation operation was approximated by averaging 10 000 samples of different diffuse fields. In the first simulation a pure-tone diffuse field was constructed, with the individual plane waves having pure tone waveforms at 500 Hz, each with a complex amplitude generated by two independent normally distributed random variables corresponding to the real and imaginary parts of the complex amplitude.<sup>5,7</sup>

Figure 5 shows the theoretical and simulated spatial correlation as a function of the distance in wavelengths. The

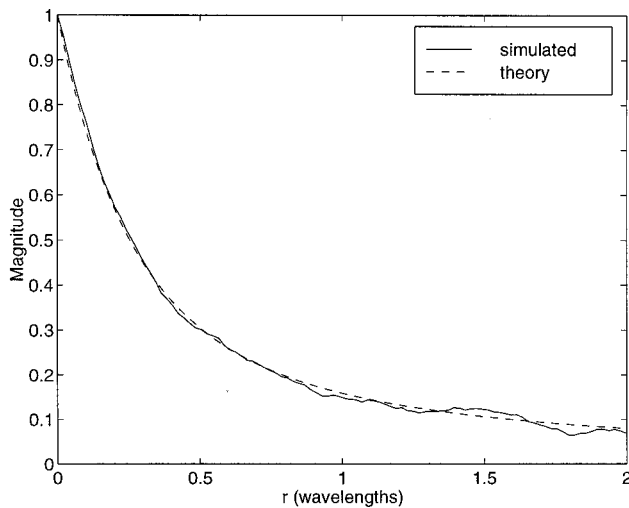


FIG. 6. The spatial correlation of the diffuse field for low-pass filtered random excitation, theoretical result (dashed curve) and simulated result (solid curve).

simulation results follows well the theoretical sinc( $kr$ ) function.

In the second simulation each plane wave carried a time-domain signal generated by filtering normally distributed white noise with a low-pass filter having a single pole at  $\omega_0 = 2\pi 500$  rad/sec, and a frequency response function  $\omega_0/(\omega_0 + j\omega)$ . Since the signal for each plane wave was generated using a different random variable, the signals were all uncorrelated. The autocorrelation of the signals, however, is decaying exponentially with time, with  $R_0(\tau) = e^{-\omega_0|\tau|}$ , as suggested above. Figure 6 shows the spatial correlation calculated theoretically, as in Eq. (18) (dashed curve), and the correlation obtained using the simulation (solid curve). The simulation results follow the theoretical  $(1 - e^{-k_0 r})/k_0 r$  function accurately, which supports the theoretical derivation presented above.

#### IV. CONCLUSIONS

A derivation for the spatial-temporal correlation of a sound signal with arbitrary power spectral density in a diffuse field was developed in this paper. This is calculated as the integral over the temporal correlation or the power spectral density of the signal exiting the diffuse field. The formulation is therefore general and can be used for sound signals with any given spectrum.

Experimental verification of the theoretical results for broadband diffuse fields, and the use of this formulation in applications involving diffuse sound fields such as reverberant room measurements and active sound control, are suggested for future study. In these cases, the assumptions made in this paper concerning the diffuse field, e.g., infinite number of plane waves, all uncorrelated and with the same spectrum, will probably not hold exactly. However, in some cases the assumptions might not be very far from the reality.

The simulations in this paper already showed that even with only 1145 plane waves the correlation is similar to that predicted analytically. Also, the plane waves in highly reverberant rooms with small absorption are expected to have similar spectrum if all originated from the same source. Furthermore, these plane waves will be poorly correlated if separated in time by delays larger than their correlation time, which is small for wide-band signals.

#### APPENDIX

The diffuse field in this study was simulated by  $N$  planes waves arriving from uniformly distributed directions in space. Spherical coordinates were used, as in Ref. 5 (see Fig. 1), to represent the diffuse field as a double summation over the angles  $\theta_i$  and  $\phi_j$ , as follows:

$$p(r, t) = \lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \sum_{i=1}^I \sum_{j=1}^J p_{i,j}(r, t), \quad (\text{A1})$$

where  $I$  is the integer value of  $\sqrt{\pi N/4}$ , and  $J$  is the integer value of  $2I \sin \theta_i$ , such that the plane wave  $p_{i,j}$  arrives from the direction

$$(\theta_i, \phi_j) = \left( \frac{i\pi}{I}, \frac{2j\pi}{J} \right). \quad (\text{A2})$$

Using this formulation, the plane waves are spatially uniformly distributed. The semicircle over the angle  $\theta$  will have  $I$  waves, while the full circle over  $\phi$  for  $\theta=0$  will have  $2I$  waves, and a number decreasing by  $\sin \theta$  to account for the smaller diameter of the circles around  $\phi$  for smaller values of  $\theta$ . In this case a total number of  $N$  waves compose the diffuse field. It should be noted that the formulation with  $I$  being the integer value of  $\sqrt{N}$ , and  $J$  the integer value of  $(\pi^2 I/8) \sin \theta_i$ , as suggested in Ref. 5, will result in a slightly higher density of waves in the  $\theta$  direction compared to the  $\phi$  direction. The resulting waves will therefore not be completely uniformly distributed in space. This, however, did not affect the final spatial correlation result as presented in Ref. 5. Equations (A1) and (A2) were used to generate the diffuse fields in the simulations above, with  $I=30$ , so that a total of  $N=1145$  plane waves composed the diffuse field.

<sup>1</sup>A. D. Pierce, *Acoustics—An Introduction to its Physical Principles and Application* (McGraw-Hill, New York, 1981).

<sup>2</sup>H. Nelisse and J. Nicolas, "Characterization of a diffuse field in a reverberant room," *J. Acoust. Soc. Am.* **101**(6), 3517–3524 (1997).

<sup>3</sup>P. A. Nelson and S. J. Elliott, *Active Control of Sound* (Academic, London, 1992).

<sup>4</sup>R. K. Cook, R. V. Waterhouse, R. D. Berendt, E. Seymour, and M. C. Thompson, "Measurement of correlation coefficients in reverberant sound fields," *J. Acoust. Soc. Am.* **27**(6), 1072–1077 (1955).

<sup>5</sup>F. Jacobsen, "The diffuse sound field," Report No. 27, The Acoustic Laboratory, Technical University of Denmark, 1979.

<sup>6</sup>A. Papoulis, *Signal Analysis* (McGraw-Hill, New York, 1984).

<sup>7</sup>J. Garcia-Bonito, "Local active control in pure tone diffracted diffuse sound field," Ph.D. thesis, ISVR, University of Southampton, 1996.

<sup>8</sup>K. Weltner, J. Grosjean, P. Schuster, and W. J. Weber, *Mathematics for Engineers and Scientists* (Stanley Thornes, U.K., 1986).

# Ultrasound field estimation method using a secondary source-array numerically constructed from a limited number of pressure measurements

Xiaobing Fan

*Department of Radiology, University of Chicago, Chicago, Illinois 60637*

Eduardo G. Moros<sup>a)</sup> and William L. Straube

*Radiation Oncology Center, Washington University, St. Louis, Missouri 63108*

(Received 3 September 1999; revised 8 December 1999; accepted 8 March 2000)

A new and faster method for the accurate estimation of acoustic fields of underwater ultrasonic transducers was developed, tested experimentally, and compared to previously reported methods. Using a limited number of pressure measurements close to the transducer's face, the method numerically constructs a virtual secondary source-array whose acoustic field is similar to the field generated by the actual transducer (primary source). The measured data are used to obtain the normal particle velocity on the surface of the virtual secondary source-array, which in turn permits the calculation of the forward propagating field using the Rayleigh–Sommerfeld diffraction integral. The method is novel in that it constructs a virtual secondary source-array, thus eliminating the problems associated with obtaining the excitation source of a real transducer; and it is faster because it uses finite differences instead of a matrix inversion to obtain the excitation source. Results showed that predicted ultrasound fields agreed quantitatively and qualitatively with measured fields for three commonly used transducer types: two planar radiators (one circular, 0.5 MHz, 1.9-cm diam.; and one square, 1 MHz, 1.2 cm on a side), and a sharply focused radiator (1.5 MHz, 10-cm diam., 10-cm radius of curvature). The agreements suggest that the secondary source-array method (SSAM) is applicable to a wide range of radiator sizes, shapes, and operating frequencies. The SSAM was also compared to these authors' previous equivalent phased array methods (EPAM) [J. Acoust. Soc. Am. **102**, 2734–2741 (1997); and Concentric ring equivalent phased array method (CREPAM), UFFC **46**, 830–841 (1999)] which require matrix inversions. The SSAM proved to be much faster and equally or more nearly accurate than the previous methods. © 2000 Acoustical Society of America. [S0001-4966(00)03006-X]

PACS numbers: 43.58.Ta, 43.35.Yb [SLE]

## INTRODUCTION

Theoretically, the pressure field generated by an ultrasonic source vibrating in an infinite rigid baffle is described by the Rayleigh–Sommerfeld diffraction integral (RSDI). In the last decades various numerical methods have been developed to evaluate this integral to obtain pressure fields for circular planar transducers (e.g., Zemanek, 1971; Hutchins *et al.*, 1986), rectangular planar transducers (e.g., Freedman, 1960; Ocheltree and Frizzell, 1989), and focused transducers (e.g., O'Neil, 1949; Madsen *et al.*, 1981; Swindell *et al.*, 1982). In these investigations the transducers were usually assumed to vibrate with normal particle velocities having uniform amplitude and phase distributions across their front faces. However, under these assumptions the theoretical results generated with the RSDI have been shown not to agree well with measurements for circular planar ultrasound transducers (e.g., Hutchins *et al.*, 1986) and focused ultrasound transducers (e.g., Moros and Hynynen, 1992), especially for piezoelectric transducer (PZT) crystals. In practice, the particle velocity distribution (both amplitude and phase distributions) across the surface of a transducer is not uniform;

this appears to be the main reason for the discrepancies between experimental and RSDI generated results (Hutchins *et al.*, 1986; Moros and Hynynen, 1992; Schafer and Lewin, 1989; Vecchio *et al.*, 1994).

In recent years, measured one- or two-dimensional pressure distributions (Vecchio *et al.*, 1994; Zhou *et al.*, 1996; Fan *et al.*, 1997, 1999) have been combined with existing mathematical models to develop new techniques to predict the entire field of an arbitrary ultrasound transducer. Among the most popular models used thus far are the angular spectrum method (ASM) (Stepanishen and Benjamin, 1982; Medeiros *et al.*, 1984; Waag *et al.*, 1985), the superposition of Gaussian beams (SGB) technique (Wen and Breazeale, 1988; Zhou *et al.*, 1996), and methods based on ultrasound phased array theory (Ebbini and Cain, 1989; Fan *et al.*, 1997, 1999). Despite the successful predicting ability of these recently developed field calculation methods, they all require computational intensive numerical algorithms and/or software packages to calculate Fourier transforms, solve nonlinear parameter estimation problems, or compute the pseudo-inverse of a matrix. Furthermore, some of these methods cannot always be applied to an arbitrary transducer because, for example, they may require the source to be symmetric, or they may require exceedingly large computational resources

<sup>a)</sup>Electronic mail: moros@castor.wustl.edu

for large and/or high-frequency transducers (Fan *et al.*, 1997, 1999). Therefore, the development of a mathematically simpler method of ultrasound field calculation that requires less computer memory and that is computationally faster is warranted, especially for the accurate prediction of pressure fields from large non-axisymmetric ultrasonic transducers.

In this paper a new method is presented that meets the characteristics of being mathematically simpler and computationally faster. The new method, as do our previous ones, requires measured pressure data (generated by the original transducer), but it uses them differently. Here, the complex pressure (amplitude and phase) data are used to compute the particle velocity distribution across a virtual secondary source-array from finite pressure differences, thus the new method has been named the secondary source-array method or SSAM. The calculation of the particle velocity from finite (measured) pressure differences has been extensively applied to determine the acoustic intensity distributions close to vibrating sources for frequencies less than 2 kHz in air (e.g., Kristiansen, 1981; Pavic, 1977; Pettersen, 1979), whereas the concept of a virtual secondary source was recently introduced and applied to compute ultrasonic fields in heterogeneous media (Fan and Hynynen, 1994; Moros *et al.*, 1999). Therefore, the SSAM is novel in that it combines these techniques to estimate acoustic fields more nearly accurately and faster without incurring the problems associated with obtaining the excitation source of a real transducer (primary source) or of an equivalent phased array transducer using inverse methods (Fan *et al.*, 1997, 1999). In the following sections the mathematical formulation of the SSAM is presented first. Second, the experimental setup and procedures for measuring the pressure fields are described. Third, measured and predicted pressure and phase distributions are compared. Fourth, the SSAM is compared to our previous methods, namely, the equivalent phased array method or EPAM (Fan *et al.*, 1997) and the concentric ring equivalent phased array method or CREPAM (Fan *et al.*, 1999). Last, the results are discussed and conclusions are drawn.

For a detailed description and evaluation of the EPAM and the CREPAM (Fan *et al.*, 1997, 1999), the reader is referred to the original papers. For the purposes of this paper it suffices to say that both methods combined phased array theory (Ebbini and Cain, 1989) with the RSDI and measured acoustic data to predict the field generated by a single-source ultrasonic transducer. This is performed by mathematically treating the transducer as a phased array composed of many small elements. The methods find the excitation source for the (equivalent) phased array that produces an acoustic field similar to that of the single-source transducer. The most important mathematical step in both of these methods is the computation of the pseudo-inverse of a large matrix. The CREPAM is limited to axisymmetric radiators.

## I. METHODS

### A. Pressure field and particle velocity calculations

For a planar or focused ultrasound transducer the velocity potential ( $\psi$ ) at a field point can be evaluated using the Rayleigh–Sommerfeld diffraction integral (RSDI),

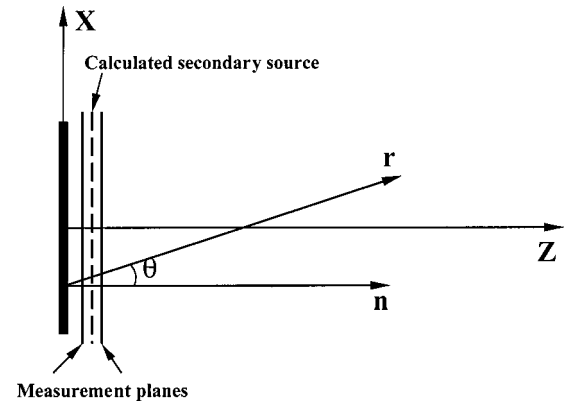


FIG. 1. Diagram showing the relative locations of the measurements planes and the secondary source with respect to the transducer. The coordinate system used in the simulations is also shown. The y-axis is pointed out from the paper.

$$\psi(x, y, z) = \frac{1}{2\pi} \int_s u \frac{e^{-jkr}}{r} ds, \quad (1)$$

where  $j = \sqrt{-1}$ ,  $u = u_0 e^{j\omega t}$  is the particle velocity normal to the surface of the transducer with amplitude  $u_0$  and angular frequency  $\omega$ ,  $t$  is time,  $k = \omega/c$  is the real wave number,  $c$  is the speed of sound in the medium,  $r = \sqrt{(x-x_s)^2 + (y-y_s)^2 + (z-z_s)^2}$  is the distance from a point  $(x_s, y_s, z_s)$  on the surface of the transducer to the field point of interest  $(x, y, z)$ , and  $s$  is the surface area of the transducer (Fig. 1). The particle velocity along the radial direction ( $v_r$ ) and the complex acoustic pressure ( $p$ ) at a field point are given by definition as

$$v_r = -\frac{\partial \psi}{\partial r} \quad (2)$$

and

$$p = \rho \frac{\partial \psi}{\partial t} = j\omega\rho\psi = j\rho ck\psi, \quad (3)$$

where  $\rho$  is the density of the medium.

### B. Secondary source estimations

The new method of pressure field calculation developed in this article is based on the numerical construction of a secondary source (array) using measured pressure fields. In order to use the constructed secondary source to calculate the entire pressure field for a given transducer, the normal particle velocity on the secondary source surface must be determined. Without loss of generality, we assume the secondary source is a plane parallel and close to the face of the transducer. Since plane wave assumptions do not hold in the very near-field, the use of the impedance relationship [i.e.,  $v = p/(\rho c)$ ] to calculate the particle velocity could introduce large errors (Mair *et al.*, 1987). Here we apply a different technique that avoids these potential errors. According to Eqs. (2) and (3), the particle velocity along the radial direction ( $v_r$ ) is given by

$$v_r = -\frac{1}{j\rho ck} \frac{\partial p}{\partial r} = \frac{j}{\rho ck} \frac{\partial p}{\partial r}, \quad (4)$$

hence, the normal particle velocity ( $v$ ) can be determined by projecting the radial particle velocity ( $v_r$ ) onto the normal to the plane of the secondary source (Fig. 1), thus,

$$v = \frac{j}{\rho c k} \frac{\partial p}{\partial r} \cos(\theta) = \frac{j}{\rho c k} \frac{\partial p}{\partial z}, \quad (5)$$

where  $\theta$  is the angle between the radial vector and the normal to the secondary source, and  $\cos(\theta) = z/r$ . Equation (5) says that the particle velocity of the secondary source can be determined from the partial derivative of the pressure with respect to the spatial variable  $z$ . This latter quantity can be approximated numerically (e.g., using finite differences and measured pressures), i.e.,

$$\frac{\partial p}{\partial z} \approx \frac{p(x, y, z + \Delta z/2) - p(x, y, z - \Delta z/2)}{\Delta z}, \quad (6)$$

where  $\Delta z$  is a small spatial interval in the  $Z$ -direction. Therefore, the normal particle velocity on the secondary source plane can be calculated approximately from two, closely spaced (separated by  $\Delta z$ ), measured pressure planes; the simplifying assumption being that the particle velocity is proportional to a finite pressure difference. This assumption has been previously applied successfully in the ‘‘two microphones’’ technique developed for the measurement of acoustic intensity in air generated by a vibrating membrane (frequency  $< 2$  kHz) (e.g., Kristiansen, 1981; Pavic, 1977; Pettersen, 1979). The choice of  $\Delta z$  is important. It should be chosen such that  $\Delta z \ll \lambda$ , where  $\lambda$  is the wavelength in the medium. This is because for a small change in location  $\Delta z$ , the phase change is given by  $k\Delta z$  (where  $k = \omega/c = 2\pi/\lambda$ ). Since the maximum phase change is  $2\pi$ , it follows that  $\Delta z$  should be small enough so that a phase change is much less than  $2\pi$ , i.e.,  $k\Delta z = (2\pi/\lambda)\Delta z \ll 2\pi$ , and thus  $\Delta z \ll \lambda$  or  $\Delta z/\lambda \ll 1$ . Notice that the choice of  $\Delta z$  depends on the operating frequency, therefore, for a broadband signal  $\Delta z$  should be chosen so that the above inequality is satisfied for the smallest significant wavelength component in the signal.

Once the normal particle velocity is calculated, the pressure field produced by the secondary source-array, assumed to be located in between the measurement planes, can be calculated using Eq. (3) as

$$p = \frac{j\rho c k}{2\pi} \sum_{n=1}^N \int_{s_n} v_n \frac{e^{-jkr_n}}{r_n} ds_n, \quad (7)$$

where  $N$  is the total number of elements on the secondary source plane,  $v_n$  is the normal particle velocity for the  $n$ th element,  $r_n$  is the distance from a point on the  $n$ th element to the field point of interest, and  $s_n$  is the area of the  $n$ th element. Therefore, from a limited number of pressure measurements the complete field of an arbitrary ultrasound source can be accurately predicted by computing the field for the numerically constructed virtual secondary source-array. The integral in Eq. (7) was evaluated by dividing each element into smaller subelements, small enough that they could be treated as point sources. The size of these subelements was therefore wavelength ( $\lambda$ ) dependent (usually about  $\lambda/3$ ).

The above formulation can be applied to an arbitrary ultrasonic source. For axisymmetric ultrasound transducers,

TABLE I. The transducers used in the experiments and simulations to evaluate the SSAM.

Transducer	Shape	Size	Frequency (MHz)
T1	Planar disk	1.9-cm diameter	0.5
T2	Planar square	1.2-cm side length	1.0
T3	Spherically curved (focused)	10-cm diameter & 10-cm radius of curvature	1.5

however, the constructed secondary source can be a concentric-ring array. In the latter case the measurements are reduced to radial profiles [one dimensional (1D)]. Therefore, both measurement and calculation times can be significantly decreased if it is known that the real source produces an axisymmetric field (Fan *et al.*, 1999).

## II. EXPERIMENTS

### A. Transducers and hydrophone

The ultrasound transducers used in this study were air-backed single (nondiced) PZT elements (EBL#4, Staveley Sensors, Inc., East Hartford, CT) mounted on acrylic plastic holders (Plexiglas, Rohm & Haas Co., Philadelphia, PA), and were excited by a continuous radio frequency (rf) signal. Their sizes, shapes, and operating frequencies are given in Table I.

The pressure fields were detected in a degassed water tank using a preamplified 0.6-mm diameter polyvinylidene fluoride (PVDF) needle hydrophone (Model TNU001A, NTR Systems, Inc., Seattle, WA). According to the manufacturer's data and in-house evaluation, the hydrophone directional response is above  $-10$  dB for frequencies between 0.5 and 1.5 MHz for incident angles within  $\pm 60^\circ$  relative to normal incidence.

### B. Pressure field measurements

The continuous wave pressure field measurement technique used has been reported previously (Fan *et al.*, 1997, 1999). Since the fields produced by transducers T1 and T3 were axisymmetric, the measured pressure fields used to calculate the secondary source were two, parallel, 1D radial profiles as mentioned in Sec. I, while pressure distributions on two, parallel, 2D meshes were measured for transducer T2. The pressure measurements were taken as close as practically possible to the face of the transducers because the secondary source can only be used to predict forward propagation (Fan *et al.*, 1999). In this study, the pressures were measured in front of the transducers (i) along radial profiles at  $Z = 1.0$  and  $Z = 1.25$  mm for transducer T1, (ii) on planes at  $Z = 1.0$  and  $Z = 1.25$  mm for transducer T2, and (iii) along radial profiles at  $Z = 15.0$  and  $Z = 15.25$  mm for the focused radiator T3. The reason for choosing 15.0 mm as the closest practical distance to the face of the spherically curved transducer was that the height of the cap (depth of concavity) was 13.4 mm for T3. The length of the radial profiles, size of the planes, and the number of measurements points are given in Table II. Notice that in all cases  $\Delta z$  was equal to 0.25 mm, therefore, the requirement that  $\Delta z/\lambda \ll 1$  was always satisfied



TABLE II. The size of the measurement plane or length of the radial profile, number of measurement points, and the step size used for the calculation of the secondary source. The separation between planes divided by the wavelength ( $\Delta z/\lambda$ ) is also given.

Transducer	Radial profile length or size of plane (mm)	Grid spacing (mm)	Number of points	$\Delta z/\lambda$
T1	10.5	0.5	22	0.083
T2	15.0×15.0	0.5	31×31(961)	0.167
T3	50.0	0.5	101	0.25

(see Table II). The pressure fields used for comparison with the simulation results were consistently measured at other radial and axial planes. The schematic of the overall measurement setup and coordinate system used in this study was given in Fig. 1 of Fan *et al.*, 1997 for the general case, and in Fig. 1 of Fan *et al.*, 1999 for the axisymmetric transducer case.

### III. RESULTS

#### A. Comparison between SSAM-predicted and measured pressure distributions

Once the pressure measurements were made, Eq. (6) was applied and the particle velocities to construct the secondary sources were obtained per Eq. (5). Then the pressure fields were predicted by Eq. (7). First, the results for the planar disk T1 are presented in Fig. 2. This figure shows a comparison between measured and predicted relative pressure amplitude and phase distributions along the central axis of the beam. It can be seen that the measured and predicted

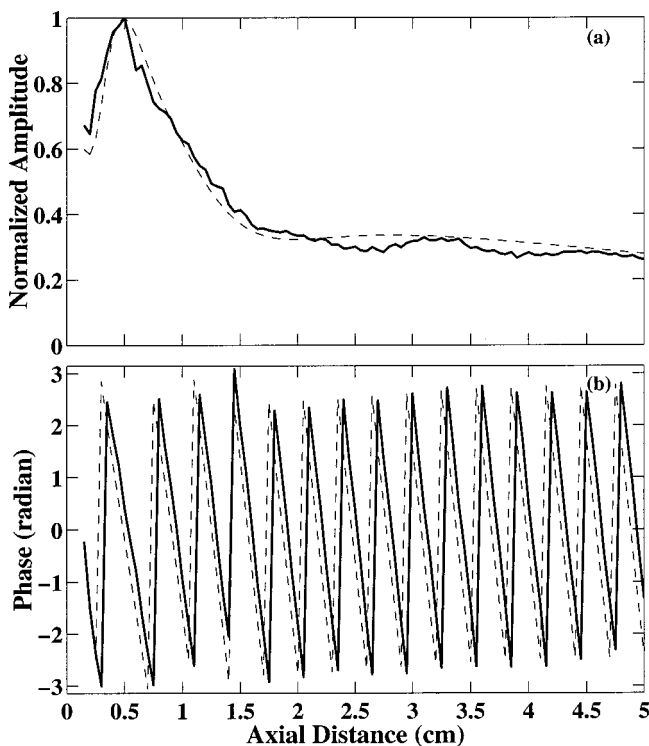


FIG. 2. Comparison between measured (solid line) and predicted (dashed line) relative pressure fields along the central axis of the 1.9-cm diameter planar disk transducer operating at 0.5 MHz (T1, Table I): (a) normalized amplitude, and (b) phase in radians.

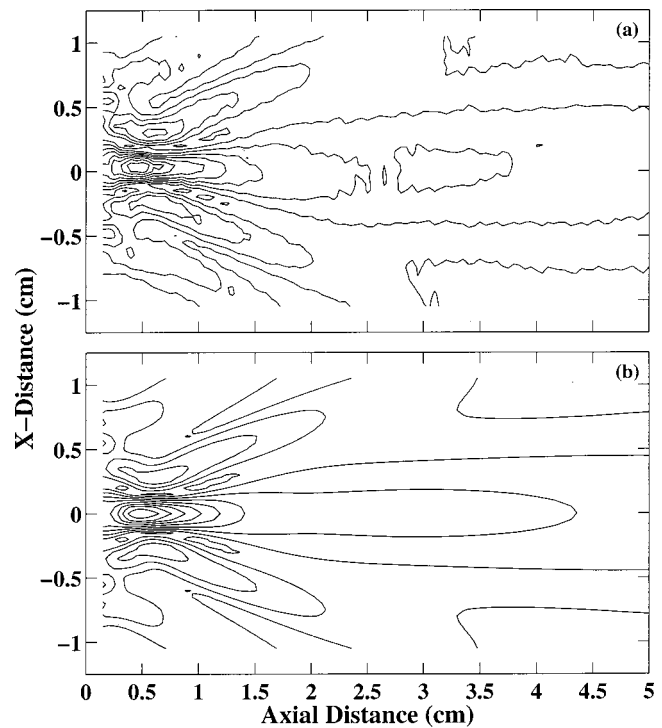


FIG. 3. Contour plots of measured (a) and predicted (b) relative amplitude pressure fields on an axial plane for transducer T1. Contours are plotted every 10% of the maximum value beginning with the 10% contour.

data agreed quantitatively for both amplitude and phase distributions. However, some shifting (0.25–0.5 mm) between experimental and simulated results can be observed. We believe the shifting can be attributed to errors resulting from the limited resolution and accuracy (0.15 and 0.25 mm, respectively) of the stepping/scanning system used to move the hydrophone. The measured relative pressure amplitude distributions on an axial plane were also very similar to the predicted distribution as shown in Fig. 3. As expected, the predicted field was symmetric about the central axis because the secondary source was a concentric-ring array constructed using two one-dimensional radial profiles of measured complex pressures. Please note that both measured and predicted pressure amplitudes were normalized to unity by dividing each by its own maximum. This normalization was also applied for the results of transducers T2 and T3.

The measured and predicted relative pressure amplitude and phase distributions along the central axis for the planar square transducer T2 are shown in Fig. 4. The peaks and valleys of both magnitude and phase distributions were predicted accurately. Again, some shifting ( $\sim 0.25$  to 0.5 mm) can be observed. As with T1 the amplitude distributions on axial planes showed remarkable agreement for both near and far from the transducer's face (Fig. 5).

Finally, for the spherically focused transducer T3 the same comparisons are shown in Figs. 6 and 7. The location, shape, and size of the focus predicted by the SSAM method were in good agreement with the experimental results, again except for some amount of shifting ( $\sim 0.25$  to 0.5 mm). For T3 as for T1, only radial profiles were measured and used in the construction of the secondary source, thus the predicted fields were symmetric about the central axis. The results for

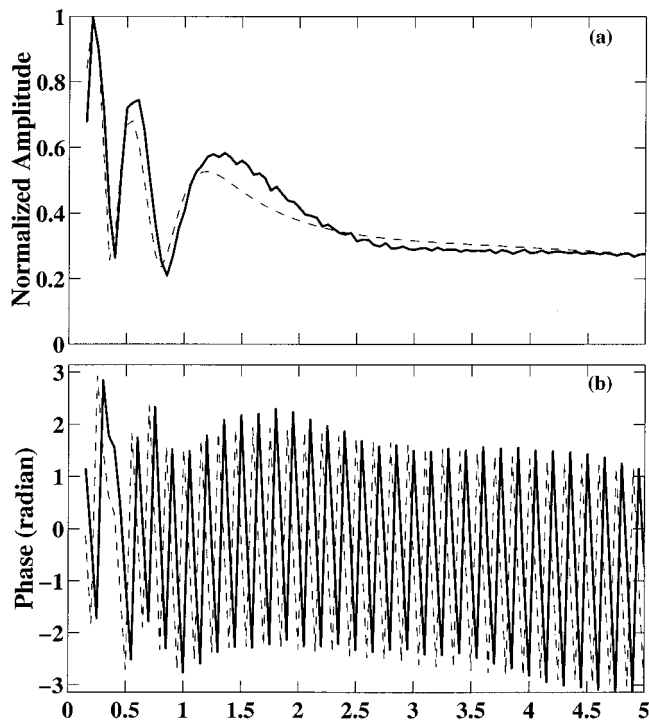


FIG. 4. Comparison between measured (solid line) and predicted (dashed line) relative pressure fields along the central axis for the  $1.2 \times 1.2$ -cm planar transducer operating at 1.0 MHz (T2, Table I): (a) normalized amplitude, and (b) phase in radians.

T3 demonstrate that a field from a focused beam can be accurately predicted by constructing a planar secondary source consisting of a concentric-ring phased array with only 101 measurement points of complex pressure (Table II).

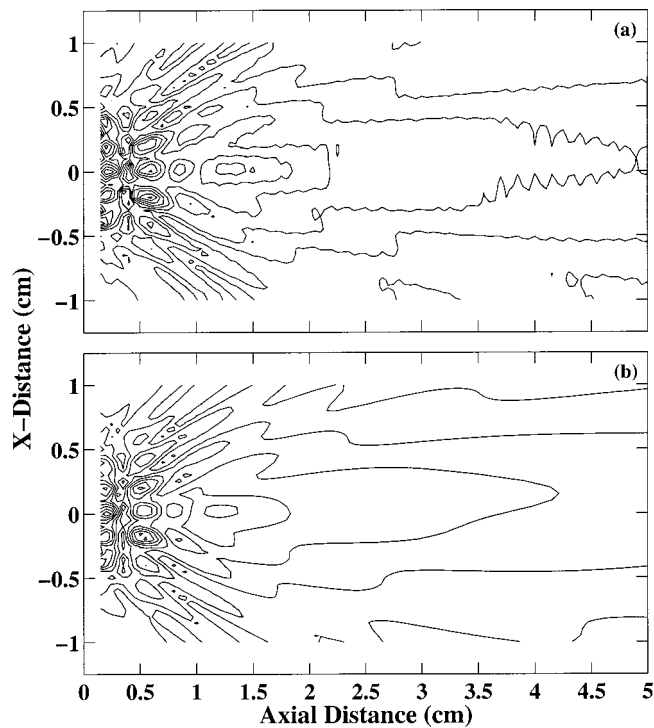


FIG. 5. Contour plots of measured (a) and predicted (b) relative amplitude pressure fields on the axial plane for transducer T2. Contours are plotted every 10% of the maximum value beginning with the 10% contour.

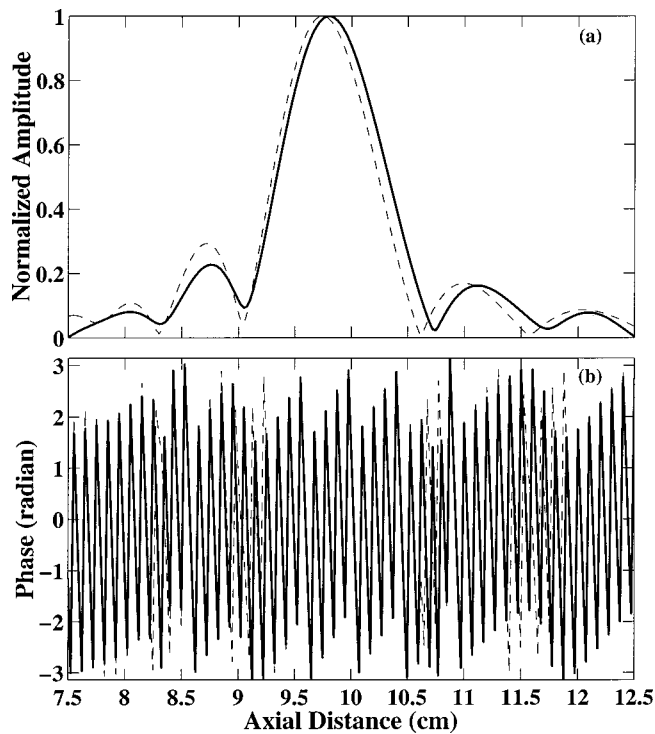


FIG. 6. Comparison between measured (solid line) and predicted (dashed line) relative pressure fields along the central axis for the 10-cm diameter and 10-cm radius of curvature spherically curved (focused) transducer operating at 1.5 MHz (T3, Table I): (a) normalized amplitude, and (b) phase in radians.

## B. Comparison between the SSAM, the EPAM, and the CREPAM

Additional simulations using the EPAM and the CREPAM (Fan *et al.*, 1997, 1999) for transducers T1, T2,

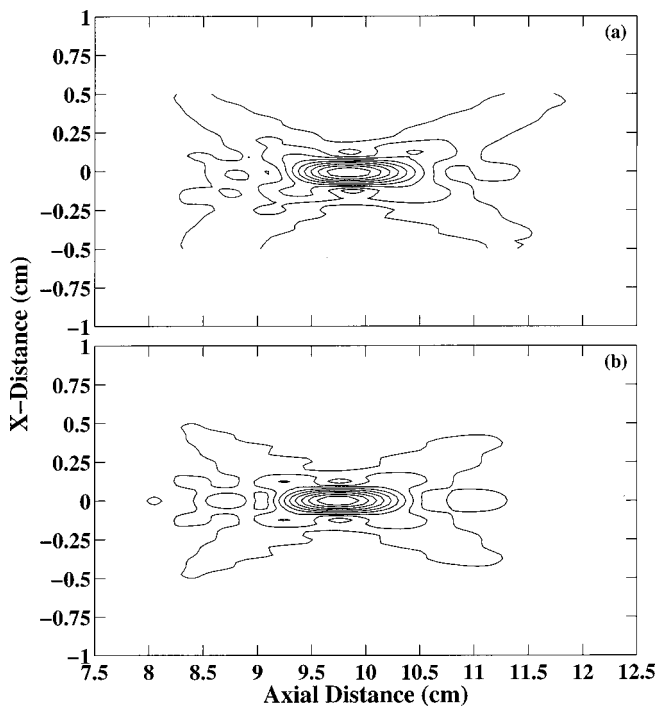


FIG. 7. Contour plots of measured (a) and predicted (b) relative amplitude pressure fields on the axial focal plane for transducer T3. Contours are plotted every 10% of the maximum value beginning with the 10% contour.

TABLE III. Comparison of the EPAM, the CREPAM, and the SSAM.

	EPAM			CREPAM			SSAM			
	T1	T2	T3	T1	T2	T3	T1	T2	T3	
Transducer	T1	T2	T3	T1	T2	T3	T1	T2	T3	
Symmetry	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	
Applicable	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	
ARE <sup>a</sup> (%)	6.0	5.0	...	8.0	...	<sup>d</sup>	7.0	8.0	<sup>d</sup>	
CPU times <sup>b</sup> (%)	79	100	...	0.3	...	9.5	0.6	0.2	3.4	
Memory <sup>c</sup> (%)	100	58	...	2.8	...	5.6	2.8	5.6	5.6	
Measured data	1 plane			1 radial line			2 lines		2 planes	2 lines
Sensitivity	high			very high			high		low	high
Math operation	matrix inversion			matrix inversion			finite differences			

<sup>a</sup>The average relative errors (ARE) were calculated between measured and predicted relative pressure fields along the central axis.

<sup>b</sup>The CPU times (computational times) to compute central axis distributions were normalized to a maximum of 100%.

<sup>c</sup>Computer memory requirements (RAM) were normalized to a maximum of 100%.

<sup>d</sup>The ARE were larger than 10% due to shifting effects.

and T3, where applicable, were performed in order to illustrate the advantages of the SSAM. The results are presented in Table III.

Both the EPAM and the SSAM can be applied to any transducer while the CREPAM can only be applied to transducers with axisymmetric fields. This explains the empty boxes for T2 under the CREPAM heading. Notice that empty boxes also appear for T3 under the EPAM heading although the method is applicable in principle. In this case the computer resources demanded by the EPAM exceeded those available. Although the average relative errors are all of the same order of magnitude, the computer resources (computing time and RAM memory) required by the EPAM are one to two orders of magnitude larger than for the CREPAM and the SSAM. The price for the advantages of the SSAM, namely, applicability to an arbitrary transducer, shorter computational times, and smaller memory requirements, is the need for more pressure measurements. Because the SSAM does not involve a matrix inversion operation and because it uses more measured pressure data, its accuracy is less sensitive to measurement error and/or noise than the EPAM's or the CREPAM's.

#### IV. DISCUSSION AND CONCLUSIONS

The measured and SSAM predicted ultrasound fields showed substantial agreement for all three transducers. The results indicate that the construction of a secondary source from a limited number of pressure measurements is an accurate numerical approach to predict ultrasound fields of monofrequency continuous wave transducers typically used in therapy applications. They also showed that the particle velocity can be accurately estimated from finite pressure differences. The advantages of the SSAM over other existing methods (e.g., Medeiros *et al.*, 1984; Zhou *et al.*, 1996; Fan *et al.*, 1997, 1999) can be summarized as follows: (1) lower error and noise amplification since it does not involve matrix inversion and it uses more measured data; (2) numerically straightforward and easy to implement; (3) lower computational requirements; (4) the measured pressure data (amplitude and phase) is used directly without preprocessing; and (5) it is applicable to any source given ample computational power.

The discrepancies between the experimental and simulation results were most likely the result of measurement error/

noise as we have discussed before for the EPAM and CREPAM (Fan *et al.*, 1997, 1999). However, since the SSAM uses more measured data than the EPAM or the CREPAM, the sensitivity of the results to sensor alignment and measurement error/noise is lower. The distance separating the two measurement radial lines or planes should be as small as possible (much smaller than the wavelength) to minimize errors during the calculation of the particle velocity of the secondary source from finite pressure differences. We used the simplest numerical approximation [the two-point rule given by Eq. (6)] to approximate the partial derivative in Eq. (5). Evidently it worked very well. However, more nearly accurate approximations can be applied to improve the accuracy such as central differences (a three-point rule). The required number of the radial profiles or planes would depend on the specific numerical technique chosen.

To demonstrate the predicting power and accuracy of the SSAM we applied it to three different transducers in this study (Table I). Given the mathematical foundation and formulations of the method and the results presented, we draw the inference that the SSAM is applicable to any ultrasound transducer given sufficient computational resources. In this sense the SSAM and the EPAM are equivalent. However, since the SSAM is computationally more efficient (by one to two orders of magnitude), the SSAM method can be used to predict ultrasonic forward propagation fields for much larger non-axisymmetric or irregularly shaped transducers.

All the results presented here were for transducers operating in continuous wave mode. However, we claim that the SSAM can be extended to predict fields resulting from other excitations (e.g., pulsed signals) as long as the pressure field's amplitudes and phases are accurately and properly measured.

In summary, ultrasound fields were accurately predicted by a new ultrasound field calculation method that is more computationally efficient than previously reported techniques. The method consists of constructing a secondary source-array from a limited number of complex pressure measurements; thus the name secondary source-array method or SSAM. The SSAM can be a powerful tool to accurately predict ultrasound fields in applications such as ultrasound hyperthermia and focused ultrasound surgery.

## ACKNOWLEDGMENTS

This work was supported by a DHHS/NCI Grant No. R29-CA63121. The authors wish to thank Dr. Gregory Karczmar for his kind support and helpful input.

- Ebbini, E., and Cain, C. A. (1989). "Multiple-focus ultrasound phased-array pattern synthesis: Optical driving-signal distributions for hyperthermia," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **36**, 540–548.
- Fan, X., and Hynynen, K. (1994). "The effects of curved tissue layers on power deposition patterns of therapeutic ultrasound beams," *Med. Phys.* **21**, 25–34.
- Fan, X., Moros, E. G., and Straube, W. L. (1997). "Acoustic field prediction for a single planar continuous-wave source using an equivalent phased array method," *J. Acoust. Soc. Am.* **102**, 2734–2741.
- Fan, X., Moros, E. G., and Straube, W. L. (1999). "A concentric-ring equivalent phased array method to model fields of larger axisymmetric ultrasound transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **46**, 830–841.
- Freedman, A. (1960). "Sound field of a rectangular piston," *J. Acoust. Soc. Am.* **32**, 197–209.
- Hutchins, D. A., Mair, H. D., Puhach, P. A., and Osei, A. J. (1986). "Continuous-wave pressure fields of ultrasonic transducers," *J. Acoust. Soc. Am.* **80**, 1–12.
- Kristiansen, U. R. (1981). "A numerical study of the acoustic intensity distribution close to a vibrating membrane," *J. Sound Vib.* **76**, 305–309.
- Madsen, E. L., Goodsitt, M. M., and Zagzebski, J. A. (1981). "Continuous wave generated by focused radiator," *J. Acoust. Soc. Am.* **70**, 1508–1517.
- Mair, H. D., Hutchins, D. A., and Puhach, P. A. (1987). "Intensity fields of continuous-wave axisymmetric transducers," *J. Acoust. Soc. Am.* **81**, 328–334.
- Medeiros, A. F., and Stepanishen, P. R. (1984). "The forward and backward projection of acoustic fields from axisymmetric ultrasonic radiators using impulse response and Hankel transform techniques," *J. Acoust. Soc. Am.* **75**, 1732–1740.
- Moros, E. G., Fan, X., and Straube, W. L. (1999). "An ultrasound power deposition model for the chest wall," *Ultrasound Med. Biol.* **25**, 1275–1287.
- Moros, E. G., and Hynynen, K. (1992). "A comparison of theoretical and experimental ultrasound field distributions in canine muscle tissue in vivo," *Ultrasound Med. Biol.* **18**, 81–95.
- Ocheltree, K. B., and Frizzell, L. A. (1989). "Sound field calculation for rectangular sources," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **36**, 242–248.
- O'Neil, H. T. (1949). "Theory of focusing radiators," *J. Acoust. Soc. Am.* **21**, 516–526.
- Pavic, G. (1977). "Measurement of sound intensity," *J. Sound Vib.* **51**, 533–545.
- Petersen, O. K. O. (1979). "A procedure for determining the sound intensity distribution close to a vibrating surface," *J. Sound Vib.* **66**, 626–629.
- Schafer, M. E., and Lewin, P. A. (1989). "Transducer characterization using the angular spectrum method," *J. Acoust. Soc. Am.* **85**, 2202–2214.
- Stepanishen, P. R., and Benjamin, K. C. (1982). "Forward and backward projection of acoustic fields using FFT methods," *J. Acoust. Soc. Am.* **71**, 803–812.
- Swindell, W., Roemer, R., and Clegg, R. (1982). "Temperature distributions caused by dynamic scanning of focused ultrasound transducers," *Proc. IEEE Ultrason. Symposium (IEEE No. 0090-5607)*, pp. 745–749.
- Vecchio, C. J., Schafer, M. E., and Lewin, P. A. (1994). "Prediction of ultrasonic field propagation through layered media using the extended angular spectrum method," *Ultrasound Med. Biol.* **20**, 611–622.
- Waag, J. W., Campbell, J. A., Ridder, J., and Mesdag, P. R. (1985). "Cross-sectional measurements and extrapolations of ultrasonic fields," *IEEE Trans. Sonics Ultrason.* **32**, 26–35.
- Wen, J. J., and Breazeale, M. A. (1988). "A diffraction beam field expressed as the superposition of Gaussian beams," *J. Acoust. Soc. Am.* **83**, 1752–1756.
- Zemanek, J. (1971). "Beam behavior within the nearfield of a vibrating piston," *J. Acoust. Soc. Am.* **49**, 181–191.
- Zhou, D., Peirlinckx, L., Lumori, M. L. D., and Van Biesen, L. (1996). "Parametric modeling and estimation of ultrasonic fields using a system identification technique," *J. Acoust. Soc. Am.* **99**, 1438–1445.

# Auditory perception of objects by blind persons, using a bioacoustic high resolution air sonar

Leslie Kay

*P.O. Box 124, Russell 0255, New Zealand*

(Received 26 January 1999; accepted for publication 9 February 2000)

A high-resolution octave band air sonar for spatial sensing and object imaging by blind persons is described. The system has wide-angle overlapping peripheral fields of view with a narrow central field superposed. It is noninvasively coupled to the auditory system for neural processing and spatial imaging. Blind persons learn to comprehend the auditory cortical multiple-object image that is created. The real-time synchronous relationship between hearing a change in the sensor sounds and the sensed motor actions causing the change seems to aid the learning process. Computer based testing of the sensor system is described so as to relate the physical system performance with the time-varying human auditory perception. This is so that the basic psychometric experiments studying the sensor bio-acoustic spatial resolution, resulting from the superposition of two wide-angle peripheral fields with one central narrow field, may be better understood. These tests confirm that the auditory ability of subjects to resolve close objects using the combined fields is significantly improved relative to using the peripheral fields alone. These measurements are supported by blind children learning to use the sensor system. © 2000 Acoustical Society of America. [S0001-4966(00)02605-9]

PACS numbers: 43.60.Lq, 43.66.Lj, 43.80.Ka [WA]

## INTRODUCTION

A number of totally blind persons have learned to perceive their nearby multiple-object environment using a wide-angle binaural CTFM (continuous transmission frequency modulation) air sonar sensor mounted in spectacle frames.<sup>1,2</sup> This took place following an extensive field evaluation exercise involving the first 100 totally blind persons and their 30 specialized "Orientation and Mobility" teachers.<sup>3</sup> Some blind persons have now used the sensor for over 25 years. The experimental sensor used, the Binaural Sensory Aid, later became known as the Sonicguide™. The basic system concept for imaging a multiple-object space is shown very schematically in Fig. 1, where the binaural output of the air sonar sensor is fed to the left and right ears via miniature earphones. (The term "object space" used in this study refers to the physical 3D layout of all objects and structural features in the immediate environment which produce some reflection or backscatter of ultrasonic waves. The concept of "imaging" relates to the cortical mapping of the audible signals into their cellular frequency and binaural domains.<sup>4</sup>) The distance to each object is proportionally mapped in the auditory frequency domain. Direction is mapped binaurally only by interaural intensity difference (IID). During the field evaluation, blind persons were trained to use the sensory information for locomotion control and guidance in simple peripatetic situations, increasing in complexity to travel in crowded downtown areas. All sensor users were experienced in the use of the long cane or guide dog. The evaluators considered that in such an evaluation, if the sensor was to be of significant use to a blind person, then these very experienced persons would find that their travel ability was improved relative to using either of their basic aids alone. All the blind persons appeared to respond similarly to the sounds of the sensor while under reasonably controlled basic train-

ing, were quickly able to locate objects in simple situations, and could control their locomotion relative to objects within a few hours of experience. After four weeks of training and some months of free use of the sensor as their own, a field trial questionnaire with 169 questions revealed that 70% of the respondents said that their mobility was "better" or "much better" even when using the sensory aid in complex situations. The question cross correlation's were good.

Of prime importance in studying human perception of object space through an acoustic spatial sensor is an understanding of the auditory experience being gained by a user. We know, for example, from long experience with the low angular spatial resolution of the binaural sensors, that a novice observer first wearing the sensor in a room is suddenly faced with a complex of sounds. Directing the sensor with a normal looking action of the head, a user is presented binaurally (approximately stereophonically) with a number of tone complexes which move laterally and change in frequency in synchrony with head motion. The user of the sensor quickly begins to understand that the multiple tone complexes represent the multiple objects "out there" in object space, each of which generates a tone complex representing the shape and viewing aspect of the object. This complex of sounds is startling to most novice users and is an entirely novel but generally interesting experience. To a blind person it is very stimulating; to many it is also motivating. The tone sounds may be anywhere within the frequency range 20–5000 Hz representing echoes from objects up to 5-m distance from within an arc of up to 80 degrees. (This is the sensing range and field view of the Sonicguide.)

Following the extensive work with adult blind persons, studies were initiated by several independent developmental psychologists to find out if blind children could also benefit

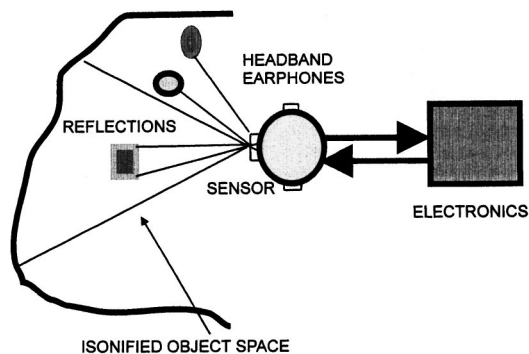


FIG. 1. Basic schematic diagram of the vision substitute, depicting a wide-angle radiation field and a wide-angle field of reception of multiple echoes of complex structure, operating from on the head of a user.

from the use of the sonar information, especially during their cognitive development.<sup>5-8</sup> Observation of the behavior of the children when fitted with an operating sensor was supplemented by the personal auditory experience of the experimenter monitoring the sounds. Sensor engineering development studies were also carried out in parallel using special experimental designs employing different arrays of sensory elements based on the binaural CTFM sonar concept. Exercises were designed to replicate controlled situations in which the blind children learning to use a sensor would be placed, so as to study the auditory information that was being made available to them. This led to the realization that the low angular resolution of the Binaural Aid type of sonar produced much confusing information in a cluttered home environment. This was not the reported experience of trained blind adults in dynamic outdoor mobility situations in busy pedestrian environments. The results of some of the studies with children were encouraging but were inconclusive about the usefulness of the information to a developing blind child. There was a noticeable difference in the use of the sensory information between those children who used the commercial model—the Sonicguide<sup>8</sup>—and those who used the specially designed Canterbury Child's Aid,<sup>9</sup> especially those children who exhibited dynamic behavior as they were becoming mobile. The Canterbury Child's Aid had better angular resolution. Under stationary conditions with only head motion, the spatial resolution in a multiple-object space was found to be limited to about the width of the field of view. (This term "field of view" is used in place of "beam width" as being more applicable to the human perception which was taking place.) It was evident, through personal use by the experimenters, that a better system was needed.

Simulation of dynamic conditions, and the resultant change in auditory information made available, lead to the discovery that under certain specified conditions of motion, binaural angular spatial resolution increased dramatically.<sup>10,11</sup> However, for very young developing blind children who are not accustomed to moving their head, a much higher intrinsic spatial resolution was seen to be essential while retaining a wide field of view for stimulating spatial exploration. Exploration by head motion is needed for development of cognition which encourages action.<sup>12</sup> It was evident from studies of the binaural cue to direction that if the two peripheral left and right fields of view were splayed

wider apart in angle, relative to the sensor axis, this should improve the perception of direction.<sup>13</sup> This would also emphasize the relative motion of objects so as to enable more rapid adaptation.<sup>14</sup> An increased angular spatial discrimination, modeling in an acoustic form the function of the foveal field of the eyes, was thought possible by adding a narrow central field of view to these two peripheral fields.<sup>15</sup> The wide-angle peripheral fields, like those of the eyes, would become more sensitive to acoustic flow (from Gibson's concept of optic flow?) but would still have poor spatial resolution. The reward of the narrow central field could perhaps encourage head scan—directed by the peripheral fields—to locate an object of interest and begin a kind of saccade on an object of interest using a fine angular neck scan. This would model, within the narrow ultrasonic field of view, fixation with the visual foveal field just as is done by the saccading foveal system. However, simple array theory indicates that when a central ultrasonic narrow field is superposed on the angled peripheral ultrasonic fields, serious phaselike interference between the three sensing channels would be caused, off the sensor axis, by the constantly changing wavelength of the CTFM generated ultrasound echoes. Nevertheless, experience with the binaural sonars led to the possibility that the seeming integration characteristics of the auditory system may well smooth out the cyclically time-varying phase interference. The left-right direction cue had been found to work, even though the instantaneous left and right signals varied with the angle off-axis.<sup>1</sup> The resultant high angular resolution sonar design which evolved, the testing of it as a system, and its use in psychophysical experiments with blind persons and blindfolded subjects are discussed here.

## I. ACOUSTIC SYSTEM DESIGN CONSIDERATIONS

An experimental study of the Binaural Sensory Aid showed that the optimum auditory direction cue was obtained by the left and right peripheral field of view having an azimuth angular response that is a Gaussian function<sup>13</sup> according to the relationship

$$I_R(\theta) = C_R \exp\{-(\theta - \alpha/2)^2/4K\alpha\}, \quad (1)$$

which is the right stimulus integrated over the sweep time:

$$I_L(\theta) = C_L \exp\{-(\theta + \alpha/2)^2/4K\alpha\}, \quad (2)$$

is the left stimulus integrated over the sweep time, where  $I_L$  and  $I_R$  represent the left and right signal intensities,  $\theta$  is the object or source angle,  $\alpha$  is the splay angle between the axes of the receiving fields as sensed by the user,  $K$  is a constant of magnification in the direction cue, and  $C$  is a loudness constant to be adjusted for hearing loss. This was because the Interaural Intensity Difference, produced by rotation of the fields past an object, generated a lateral shift which had a linear relationship with the angle of rotation. From this, direction of an object could be perceptually inferred. The Interaural Intensity Difference (IID) follows the relationship:

$$\theta_E = K \log_e(I_L/I_R), \quad (3)$$

where  $\theta_E$  is the estimated angle, the plot of which is called the Interaural Intensity Difference (IID) and is linear. This relationship was used in the binaural sensory aid as evaluated



FIG. 2. Photograph of the head-mounted vision substitute sensors that radiate and receive the continuous transmission frequency modulated octave band ultrasonic waves. The small oval sensor radiates the waves for echolocation. The remaining three sensors act as receivers of the reflected and scattered waves from the object space.

with an IID of approximately 0.9 dB/degree. In the manufacturing process of the Sonicguide, the relationship was reduced to 0.3 dB/degree.

### A. Transduction

This desired Gaussian field response cannot be obtained physically from a circular piston transducer. The small circular transducers as used in this new sensor act very much like pistons at a specific frequency showing the normal secondaries. Over an octave band the width of the field of view at any instantaneous frequency varies by a factor of 2. Even so, when listening for the variation in loudness, the auditory lateral shift of a fused signal appears to move in synchrony with the movement of a rod from left to right. The auditory cue, as sensed by the designers, appears to follow that of a linear lateral shift. It was assumed that the auditory system for binaural direction finding was integrating the varying left-right signals over at least one FM sweep time to produce fusion.

Key to the sensing process is the arrangement of the multielement ultrasonic sensing array. Increasing the splay angle of the peripheral fields required no change in transducer design. However, adding a large transducer that would produce an effectively narrow azimuth field of view required a new transducer design and a new physical layout of the sensor array. In Fig. 2 the array configuration eventually chosen is shown. A rectangular-shaped receiver was required to produce a narrow azimuthal field and a wider vertical field. A new design evolved and a magnified portion of the backplate of the large transducer is shown in Fig. 3. Both the radiation transducer and the larger receiver were made partly rectangular, having semicircular ends.

### B. Signal processing

The signal processing is essentially a form of demodulation converting the high frequency ultrasonic signals into audible signals in an optimum way. The basic CTFM system used in the Tri-sensor (named this after the use of three sen-



FIG. 3. A section of the backplate is shown of the large oval transducer seen in Fig. 2, showing the wide grooves and the treated rails that form the grooves.

sor elements) comprises a digitally generated range code signal having a sawtooth frequency modulation, driving the ultrasonic radiation from the small oval transducer. Each of the three channel receivers has a signal multiplier so as to provide the first stage of cross correlation in the frequency domain between the radiation signal and the echo signals. Suitable amplification is inserted in the channels.<sup>1</sup> The central channel signal is added to both the peripheral channels after multiplication. The system wave forms are depicted in Fig. 4 representing the case of two specular reflection objects in the field of view. The difference signal (or beat note) from the multipliers is a constant frequency. This is proportional to distance, with a duration almost as long as the sweep. Only the audible difference signal is amplified, frequency shaped to compensate for known losses in the propagation medium, and then fed to the earphone output. The digitally generated radiation sawtooth frequency sweep runs from 100 ( $f_1$ ) to 50 ( $f_2$ ) kHz in a time span of  $T$  s with instantaneous fly-back.  $T$  is selected to be 50 ms or 150 ms. The echo signals with a time delay of  $\tau$  and  $\tau + \Delta\tau$  are shown to be delayed relative to the radiation and produce constant difference frequencies  $f_A$  and  $f_B$ . These are proportional to the distance to the objects producing the reflections and are calculated from the following expression:<sup>16</sup>

$$f_a(t) \cong r(t)2m/c - f_r(t)2r'(t)/c, \quad (4)$$

where  $f_a(t)$  is the audible signal in generalized form,  $r$  is the range of the object causing the echo reflection,  $r'(t)$  is a derivative of  $r(t)$  when there is radial motion,  $m$  is the rate of change of frequency in the FM sweep, and  $c$  is the veloc-

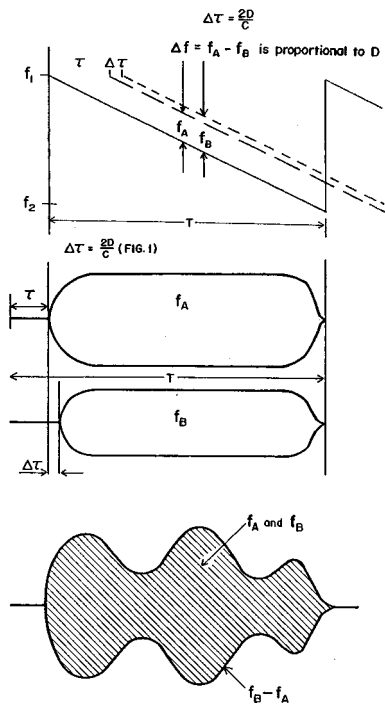


FIG. 4. CTFM time-varying ultrasonic radiation frequency and delayed echo replicas of the radiation which are cross correlated to produce two audible constant difference frequency tones of different pitch. The duration of the tones is  $T - \tau$  and  $T - \tau - \Delta\tau$ .

ity of sound in air. (Use of the central field only, could be selected by a readily accessible switch. An internal switch allowed selection of the peripheral fields only.)

The envelopes of the audible echo signals from on the axis of the central field, one lying behind the other and at a slightly different angle so that both are acoustically visible, are of the form shown idealized in Fig. 4. The sum of the two signals for the case when their radial separation is approximately 2.5 wavelengths is shown shaded. These are briefly interrupted tones representing the objects in the field of view. The interruptions are to eliminate any transient clicks in the output caused by the cyclical discontinuity in the waveform. A very large signal dynamic range is essential so that system noise is well below the small signal level of the minimum elemental reflections from complex-shaped objects at the maximum operational distance. The dynamic range of signal reflection from common objects and surfaces at the operating frequency is in the region of 50 dB. The demodulation process must therefore be very linear, i.e., that there be no intermodulation components.

The system transfer function between the radiation transducer electronic driver output and the receiver audio echo output, from a test object, was arranged to produce an echo with a shaped pulsed tone of constant amplitude, similar to the single signal envelopes shown in Fig. 4. The frequency response of the radiation transducer, the influence of the air propagation medium, that of the object acting as a reflector, the response of the central channel receiving transducer, and the electronic processing of the receiver acting together enabled this to be achieved. The test object was chosen to be a long circular 40-mm-diam tube on the sensor axis standing perpendicular to the sensor axis at a distance of

400 mm. The blind user of the system is expected to be able to recognize the presence of poles in the environment by categorization from the character of the tonal signal. This tonal signal becomes a reference for the categorization of other more complex object echoes. The constant amplitude single frequency tone may be referred to as a very simple object "signature."

### C. Bio-acoustic coupling

Blind persons rely greatly on their ability to interpret environmental sounds and echoes from sounds they themselves make. The coupling of sensory aid sounds to their ears must not interfere with this ability. This coupling itself has a frequency response which is part of the overall system frequency response that corrects for the losses over distance in the propagation medium.

## II. EXPERIMENTAL METHODOLOGY

### A. Measurement of the acoustic system performance

A means was developed to measure the new sonar performance to ascertain that this conformed with the predicted performance. The transducers were individually tested in a rig which comprised a CTFM driver with an octave frequency sweep and a variable repetition rate feeding a "standard" (reference) circular transducer. This was at a distance of 800 mm from the mounted transducer under test. The test transducer then fed a receiver having an auditory output, similar to those in the Tri-sensor. The output was monitored through earphones by the experimenter, and the signal envelope was visible on a triggered CRT. Additionally, the output was fed to a computer via a "field plot" test system which

- (1) filtered the audio output signal via a 300-Hz five pole band-pass filter centered on 2000 Hz;
- (2) rectified and integrated the filtered signal such that, at the rate of rotation of the transducer, a smoothed plot would be obtained;
- (3) amplified the smoothed dc signal via a logarithmic amplifier so that the resultant output represented the logarithm of the amplitude of the smoothed dc voltage modeling the desired Gaussian response;
- (4) computer plotted the result as the transducer was mechanically rotated from  $-90$  degrees to  $+90$  degrees at a constant angular rate.

The rectified smoothed dc signal fed to the computer is an approximation of the integration of the logarithm of the dc envelope signal taken over three FM sweeps (150 ms).

The computer was used to select, from a batch of transducers tested, the four transducers, which, when operating together in a system, most closely matched the system requirements. The left and right receiver transducers must be identical in performance; that is, determined by their maximum sensitivities being equal, and the width of their fields of view being equal at the  $-6$  dB and  $-12$  dB levels. The large oval transducer sensitivity must be such that it can be adjusted in a sensor amplifier to be 4 dB above the maximum sensitivity of the peripheral fields.



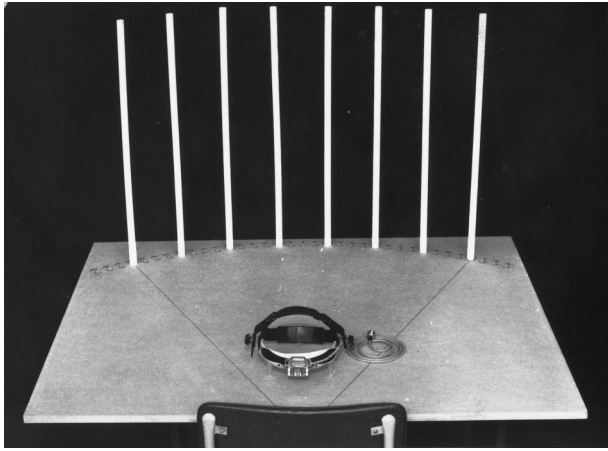


FIG. 5. Arrangement for psychometric testing of the angular resolution of the Tri-sensor by the ability of subjects to correctly count the number of rods. In this arrangement all rods sound the same restricting the variable to angle change only.

The distance accuracy is measurable by an FFT spectrum analyser. For an FFT operating over a frequency band of 5000 Hz producing 514 real lines, the line spacing is 9.72 Hz, requiring a repetitive sweep time of sub-multiples of 102.88 ms. The 50-Hz sweep was modified to 51.4 ms and triggered to synchronize with the FFT. This should produce a stationary major single line when an echo tone is exactly the frequency represented by the FFT line at, say, 1944 Hz, and this corresponds approximately to a distance of 400 mm, depending on the velocity of sound propagation at the time.

The measured echolocation system performance in the form of an audible signal is analyzed instrumentally at the receiver output to determine the accuracy of the direction cue, and relies on the techniques developed for measuring the transducer field responses. The computer was fed with the binaural signals via the field plot test system. This test system additionally calculated the value of the IID and fed this to the computer. A plot of the left and right fields of view and the IID signal was recorded on one screen as the sensor head was rotated. This was intended as a model of the auditory process of a user.

## B. Testing of the bio-acoustic system

The primary purpose of the new design, incorporating the narrow central field of view, was to improve the spatial resolution of the binaural sensors and to remove the ambiguity that occurred when two objects were in the peripheral fields at the same distance. Both objects then produce echo signals of the same frequency and it is then impossible to correctly map the two objects on the basis of amplitude difference only. An experiment was conducted to establish that the new sensor had higher angular spatial resolution.

A simple facility was constructed whereby up to 35 equally spaced 2-cm-diam wooden rods could be placed on a clear table in an arc of 1-m radius from the position of a subject's head (see Fig. 5). This layout especially eliminated any variation between rods in the distal cue. Angular resolution was then entirely dependent on angular discrimination. The rods were equally spaced and varied in number from a

minimum of two up to nine in random sequence. A seated sighted-blindfolded subject was required to count the number of rods using each of three sensors—the Sonicguide, the Canterbury Child's Aid, and the Tri-sensor. All three sensors had the same range code of 5000 Hz representing 2 m. When the Tri-sensor was used, two conditions were selected; one was the use of the central field only, and the other was the use of all three fields. The angular spacing between the rods was a multiple of 3 degrees. Sufficient time was spent teaching the subjects to achieve a plateau of performance with at least one sensor before the testing commenced.

During the test each subject would use one of the three sensors chosen according to a random pattern. A minimum of two rods, and up to a maximum of nine rods, was quietly placed in an arc on the table for each trial before the sensor was switched on. Care was taken to ensure that the rods were perpendicular to the table top. The subject was then required to count the rods. Only "correct" or "incorrect" was recorded. Sufficient 30-min sessions were run over a period of weeks for each subject to complete the counting exercise with each sensor. The number of trials carried out was 214. The exercise with a sensor was considered complete when a subject failed to obtain a correct count on three successive trials.

In a second facility two rods were placed at 1-m radial distance and one rod was placed at 0.9-m radius in such a way that a triangle was formed with the apex facing the subject's position. The base of the triangle was varied in width down to a minimum of 6 degrees. The apex always bisected the base. The subjects were required to determine when only two rods were in place instead of three. This was done by removing any one rod at random. The three rods produced a varying sound pattern as the head was scanned in an arc. With very wide spacing at the base of the triangle, there was no difficulty in discriminating between two rods and three rods when using the three different sensors. The change in the distal cue was dominant in making this possible. Each rod could be heard quite distinctly until the critical case, i.e., when no clear discrimination was possible and errors were made.

## III. RESULTS

### A. Acoustic system performance

#### 1. Sensing array field response

In Fig. 6 are shown the azimuth field plots of the four transducers acting as receivers of the sound source as plotted by the computer. The lack of a smooth plot is attributed to the simple RC-type integration used. It is not noise. Most noticeable are the small sensing field secondaries. This is because the energy in the secondaries is being blurred in angle by the wavelength change during each frequency sweep. The triggered signal envelope for one sweep covering the width of the screen, as seen on a monitor CRT, varies significantly in shape as the transducer rotates. Off-axis, the variation in the width of the field of view with frequency change is very noticeably observed in the envelope shape. When monitoring the audible signal, the character of the repetitive tone changed with angle due to the change in wave-

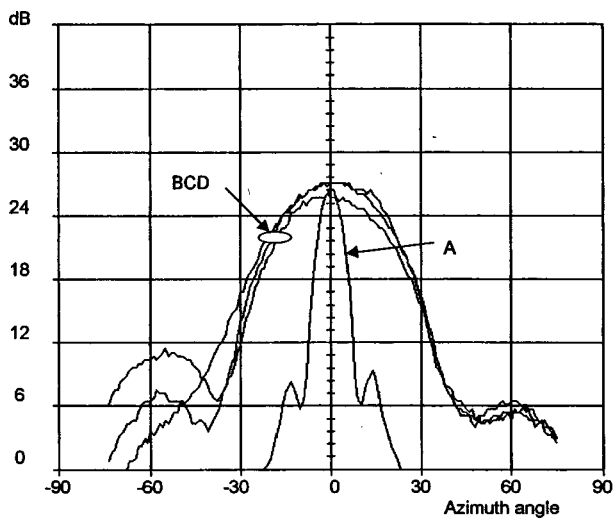


FIG. 6. Plots of the field of view of a computer-selected matched-set of left, right, and central receiving transducers, normalized to zero angular displacement. Plot A is that from the large oval transducer and plots B, C, and D are from the small round transducers overlaid with the plot from the small oval radiating transducer, all acting as receivers. The field plots model the auditory perception of loudness change as the transducer is rotated past an ultrasonic, wide bandwidth, wide-angle radiator at 1 m. The radiation code is similar to that used in the sensor system. The transducers are therefore receiving linear time varying octave wide bandwidth signals as they rotate.

length. The shape of the peripheral field approximates to the desired Gaussian function over the active field of view of 54 degrees.

The plot of the sensor field of view operating as an echolocation system is shown in Fig. 7. The plot shows a linear IID within the narrow central field of  $\pm 6$  degrees with

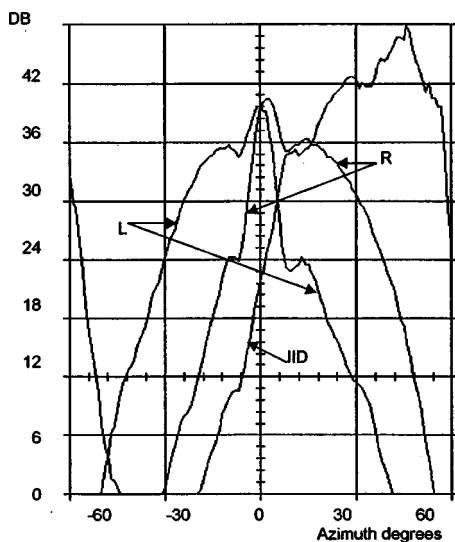


FIG. 7. Plot of the convolution of the radiation field and the two peripheral receiver fields, with the center channel added, resulting from a reflection of the radiated signal from a vertical 40-mm-diam plastic pipe acting as a standard target. These left and the right plots are spaced in angle by 23 degrees. This is so only because the radiation field is similar to the receiver fields. Also plotted is the intensity difference (IID) between the left and the right fields, the logarithm of which is approximately linear over a 60 degree arc. Zero intensity difference occurs at the angle of intersection of the two plots at approximately zero degrees. The central field enables accurate location of an object relative to the direction of the head for fixation on an object.

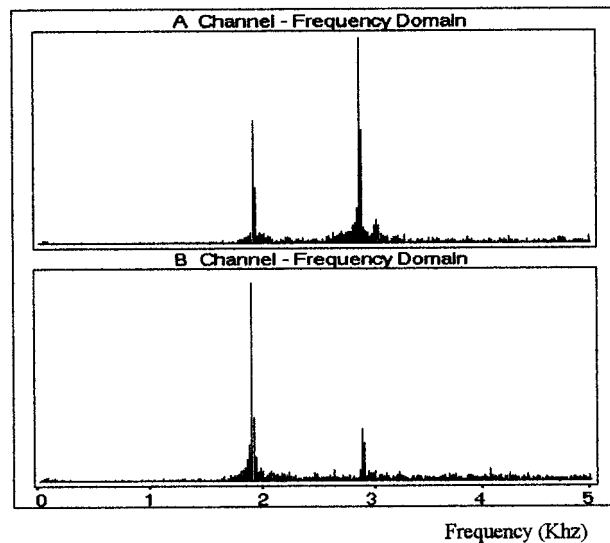


FIG. 8. FFT plots of 20-mm-diam wooden rods in the field of view mounted on a wood board. The signal line spacing is 20 Hz. This corresponds approximately to one mean wavelength. The distal scale is 1 kHz represents 200 cm. The rods are not at a distance which produces a frequency that corresponds exactly to a line frequency. A rod can be moved in distance to produce an echo frequency which exactly corresponds to a line frequency when slight air motion causes the adjacent two lines to move. The system noise prior to the first rod echo is shown to be very low and allows very accurate positioning of a rod to within 100  $\mu$ m relative to a datum when in still air conditions. The low level backscatter from the wood board, after the first echo, is clearly seen.

an IID rate of 2.0 dB/per degree. A discontinuity occurs at approximately  $\pm 6$  degrees when the IID rate reduces to 1.0 dB/degree or less. Beyond  $\pm 15$  degrees when audibly monitoring the echo signal the lateral shift becomes so great that one believes that the signal is from well to the left or right.

## 2. Distance accuracy and direction determination

Echoes in the form of tone pulses of 51.2-ms duration produce FFT frequency spectra with 19.5-Hz spectral lines, such as shown in Fig. 8, where the echo signal lines coincide with alternate FFT lines. This spacing of 19.5 Hz represents 3.9-mm radial distance. Two 20-mm rods are in the field of view at different distances and directions. Radial motion of the rods relative to the sensor head was found to produce a dramatic change in the FFT line amplitudes, so much so that a distance change of 1.0 mm produces a very noticeable change in the relative amplitudes of the major spectral lines. At a distance of 30 cm, a radial distance change of 100  $\mu$ m is just reliably detectable by the change in the line structure against the background noise.

Direction of an object is determined by the lateral shift. Thus the IID plot of Fig. 7 enables a measurement of the direction of a single object reflection within the 54 degree field of view. Discrimination between two objects at the same distance is aided by the central peak in the angular response, the perceptual centralization of the fused echo when fixating on one of the objects, and the axial character of the echo signal.

TABLE I. Angular cue only.

Sensor in use	Field of view	Angular resolution
Tri-sensor (central field only)	9 degrees	12 degrees
Tri-sensor (three fields)	54 degrees	18 degrees
Canterbury Child's Aid	35 degrees	24 degrees
Sonicguide	48 degrees	Greater than 60 degrees

## B. Bio-acoustic system performance

### 1. Angular spatial resolution with and without differential cues

The subjects learned fine control of their neck motion in a slow scanning action. They appeared momentarily to fixate auditorily on a rod in order to add it to the count—just like visual counting of white lines on a black background. A plateau of performance, that of counting eight rods correctly on every trial, was first achieved using the Tri-sensor with only the central channel operating, as shown in Fig. 5. With both central and peripheral fields together a plateau of performance was achieved counting five rods. A plateau of performance was also reached using the Canterbury Child's Aid counting four rods. No plateau of performance within the field of view of 60 degrees was achieved using the Sonicguide. The results are tabulated in Table I.

When the differential distal cue was used and three rods formed a triangle, the frequency change of 250 Hz at 2500 Hz was very noticeable. It dramatically increased the angular resolution as tested for this method situation (see Fig. 9 and Table II).

### 2. Exercise with blind children

The exercise of counting rods was included in a training program of 40 exercises for 7 blind children aged 7 to 11.<sup>17</sup> The children used only the Tri-sensor with the three fields operating. Data were collected. In exercise 12, after 4 training sessions involving 10 trails each, up to 4 rods were correctly counted in 10 test trails. Following further training for 5 sessions involving 10 trials each, 5 to 8 rods were counted correctly 75% of the time in the final session of 10 test trials.

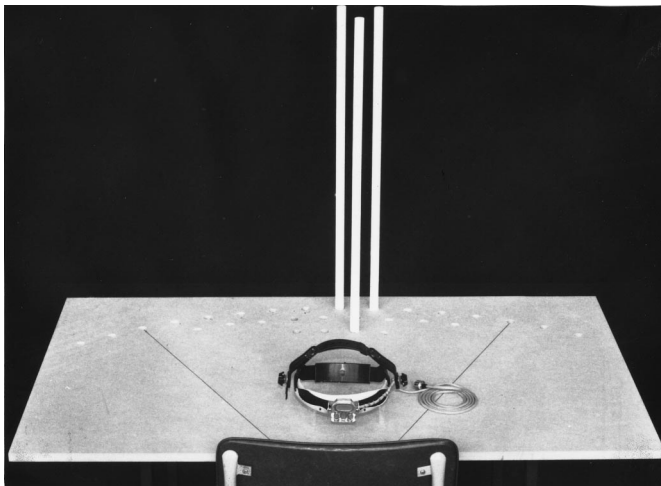


FIG. 9. Sensor resolution using the Tri-sensor when there is a 10% frequency difference between adjacent rods.

TABLE II. Angular and distal cues.

Sensor used	Angular resolution
Tri-sensor (three fields of view)	4.5 degrees
Canterbury Child's Aid	18 degrees
Sonicguide	36 degrees

Eight rods were correctly counted occasionally by some children. These seven children were rated scholastically to have very poor to normal learning ability.

## IV. DISCUSSION AND CONCLUSION

### A. The physical system

A critical step was taken in the measurement of the physical system field of view that is intuitive. By electronically modeling, in the test rig, the auditory system integration that evidently takes place in the binaural process, the integrated signal amplitude variation as plotted during the rotation of the transducer was found to simulate well the variation in loudness and lateral shift in the fused audible tones. The quite large variations in the instantaneous sound level at the two earphones and seen on a CRT during one frequency sweep duration ( $T$ ) are not binaurally perceived as such. Instead, the quality of the fused-tone changes. It can reasonably be assumed that all users will be auditorally stimulated in much the same way, and that they can be expected to hear the same fused sound. But it cannot be assumed that the perception—or interpretation of the information—will be the same as that of the system designer. The best that can reasonably be expected is that similar behavior will be observed under the same controlled conditions. This is precisely the case with visual perception, except that sighted man has a long experience of vision through natural development and naturally expects others with normal vision to see the same, whereas with this “sonocular” perception the experience is new to everyone. The new term “sonocular” is a mix of “*son*” pertaining to sound and “*ocular*” pertaining to the eye, and describes an entirely new form of spatial perception which needs a name.

The IID plot in Fig. 7 is the logarithm of the smoothed (integrated) amplitude difference between the left and right field responses as the angle of rotation of the sensor changes and is key to an understanding of the sensor as a perceptual aid for use by blind persons. When the two left and right signals are equal at zero degrees, the IID is zero. The fused signal is then centralized on the midline. (The amplitude scale from 0 to 48 dB refers only to the field plots. However, the scale of 6 dB per division applies to the IID plot.) Subjects vary greatly in their IID sensitivity, as found from psychophysical experiments carried out for the development of the Binaural sensors.<sup>13</sup> In the Tri-sensor, emphasis is on accurate location of objects by turning one's head toward an object according to the left/right direction cue and centralizing the sensor axis on the object; in other words, “looking at the object.” The IID plot confirms that the required physical sensory information is provided. Laboratory tests of human locomotion guidance by a simulated binaural sonar showed a greatly exaggerated IID relative to the normally<sup>13</sup> produced

locomotion control that was comparable to that with sight.<sup>18</sup> The IID plot shows that information is provided for the required acoustic flow found to be so important in the guidance process. Also, adaptation to the binaural cue had been tested and it was found that, with an exaggerated IID, adaptation was more rapid.<sup>14</sup> In the Tri-sensor, the peripheral field IID sensitivity is 1 dB/degree (measured without the central field) with a transducer physical splay angle  $\alpha$  of 46 degrees. This equates operationally to a field of view splay angle of 23 degrees, when the convolved radiation field and the peripheral fields are equal. In the Tri-sensor the central field IID rate is 2 dB/degrees. This high rate appears to be very important for centralizing the beam accurately on an object reflection of interest. The discontinuities at  $\pm 6$  degrees appear to indicate the limits for a finely controlled neck scan.

The 4.0-dB central peak is clearly seen in Fig. 7. The small angular displacement of the left and the right peaks is due to the summation of the central and the displaced L/R fields separately. This displacement is not auditorally noticeable. Additionally, the envelope shape of the filtered signal from the pole, on-axis, was approximately rectangular, as depicted in Fig. 4. The characteristic rapid onset of the signal followed by a constant amplitude appears to be important for locating the object on the field of view axis.

The characteristic of the echo signature from the central field is noticeably louder and also sounds different. This enables increased discrimination between two close objects by a user viewing first one object axially through a type of looking action and momentary fixation then looking at the other object. When monitoring the audible signal, there is a clear binaural lateral shift which moves more rapidly within the central field than when within the outer parts of the peripheral fields. This sensation, as observed by an experienced engineer/listener, correlates well with the IID plot. However, it is not possible from this personal experience to know what a novice or an experienced blind user would perceive.

In Fig. 8 is shown an FFT screen for two interrupted tones which normally produce two FFT lines which become a single line only at a very precise position of the rod for every shift in distance of approximately one wavelength. Air disturbance or temperature change causes the line structure to vary noticeably. At the mean frequency of 75 KHz the wavelength of sound in air is approximately 4.5 mm, which corresponds to 1.15 times the signal spectral line spacing. The first echo signal frequency in the B channel is in very close juxtaposition to an FFT line. There are a number of signal lines on either side of the longest line. These represent the energy spread caused by the signal blanking in the transmission. The system noise is seen to be very much smaller than the backscatter from the unpolished wood base upon which the rods were placed.

The good signal-to-noise ratio seen in Fig. 8 is due to (1) the high-energy radiation produced by the continuous transmission of the ultrasonic waves, and (2) the receiver matched filter process when using an FFT processor following the multiplier. This is to carry out the required correlation between the radiation signal and the echo signal. The detection threshold is well above system noise for signals from objects of interest to a blind user. Because of the octave bandwidth,

and with only 10% shaped sound blanking, the maximum physical distal spatial resolution of  $c/(\text{system bandwidth})$  is very nearly being obtained in time  $T$  by the system design.<sup>16</sup> In auditory terms this spatial resolution, approximately represented by the signal line spacing, very loosely corresponds to the just noticeable change in frequency for normal hearing in the frequency range 100–5000 Hz.<sup>19</sup>

A second echo signal could, ideally, be resolved by the FFT spectrum analyzer if it were separated in distance from the first echo signal by 1.5 times the line spacing—provided it was of a similar intensity. This could be produced by a second similar rod spaced a small angle from the first rod at a distance of approximately 1.7 mean wavelengths farther away from the sensor. The effect could also be produced by the sharp edges of two sheets of paper separated in distance of 1.7 mean wavelengths and viewed looking toward the edges perpendicularly from across the paper at a distance of, say, 150 mm. In Fig. 4 is depicted the auditory signature which is in the form of the signal envelope for two such specular reflections spaced approximately 2.5 wavelengths apart in distance and at slightly different angles—either vertical or azimuthal—so that both can be seen. The structure of the sound signatures formed by tone complexes is mainly independent of their mean audio frequency from which the distance is read. This we believe justifies the use of the psychological term “invariant”—or capable of being learned—for echo signatures from objects. The nature of the sound signatures obtained from objects is determined both by the sensor’s varying field of view with radiated frequency and the change in the aspect of objects as viewed by the movement of sensor elements. This, together with any Doppler shift, is due to any radial component of motion of the offset sensor head. Ambiguity exists which is normally eliminated by fixation on and small angle scanning of the object with the central field.

## B. The combined physical and bio-acoustic systems

Experience suggests that most other users are not aware of what is taking place; the feature is merely used by naturally learning its use. The desired system performance as a sensor can only be determined by studying the spatial perception of a blind user. In turn, perception can only be determined through many different behavioral responses during a number of experiments from a number of subjects each in a controlled experiment. These experiments do not indicate the uses to which the sensor may be put, but they at least indicate the kind of experiences that could be usefully learned. It is useful to remember that a driver’s eyesight is tested by reading letters at a distance, from which it is inferred that a driver can see well enough to drive in a busy area. We can only acquire such inference from experienced Tri-sensor users.

The psychometric tests reported here are simply a beginning, using the concepts embodied in the Snellen chart to indicate spatial acuity. In these tests, a rod spacing was found when all five subjects obtained a correct count on each trial using the Tri-sensor in both operating conditions and also when using the Canterbury Child’s Aid. This did not happen with the Sonicguide. The Tri-sensor operating in the

central field mode enabled all five subjects to count the rods correctly for an angular spacing of 12 degrees as seen in Fig. 5. When all three fields were operating, the angular spacing for correct counting increased to 18 degrees for a field of view of 54 degrees—a three times improvement over the peripheral fields of view alone.

The simple psychometric test of counting rods using three alternative sensor designs—two of which had previously been used in everyday activities—showed that a dramatic positive change in spatial acuity can be brought about through a more sophisticated sensing system. The ability of subjects to count rods when several are in the field of view at the same time, all producing a similar tone, the frequency of which is the same as the signal from the rod being viewed, may not be analytic. The high rate of IID, the accentuated amplitude of the central field, and the on-axis echo signature characteristic are thought to be the reason.

When there exists a slight difference in distance between two sources of signal, the pitch difference is used to discriminate between the two sources. This is very evident in the second bio-acoustic psychometric test. An acoustic acuity of 18 degrees in the first test for the Tri-sensor is reduced to 4.5 degrees in the second test. The angular separation between the front rod and the either of the rear two rods of 4.5 degrees allows a correct count of two or three rods.

The poor performance with the Sonicguide in these tests is attributed to the low IID rate of 0.3 dB/degrees and the wide angle response of 80 degrees at  $-20$  dB. The sounds were confusing. The sensor used in the major field trial with blind persons was different. Even so, some blind subjects (perhaps 500 World Wide) learned to use the Sonicguide to their great satisfaction over some 25 years. This could be accounted for by their having a binaural sense that was sensitive to the direction cue that was higher than normal, as found possible by Rowell.<sup>13</sup>

When in general use, the sensors operate with both angular and distal cues working together to determine the spatial layout of multiple objects. It is seen from the second table that the Tri-sensor has a spatial resolution that is at least four times better than the Canterbury Child's Aid and eight times better than the Sonicguide for the particular situation chosen.

A significant variable in performing the counting exercise was the neck control. Sometimes the head drooped, reducing the sensor acuity off the vertical axis and sometimes the angular control was insufficient to be sure that a rod was not missed when turning to the next rod. In real life few objects are at the same distance (within a wavelength or two) from an observer (subject), and when in motion the observer's view quickly changes so as to rapidly remove the ambiguity created by equidistant object or features. In the Tri-sensor, the radial resolution is comparable with the just noticeable difference in the pitch of two signals. This represents a distance of approximately one wavelength at the shorter ranges.

In conclusion, it can be said that the performance of the physical sonar forming the information gathering part of the bio-acoustic system has been tested and shown to have spe-

cial characteristics. These indicate that the machine operates as designed so as to produce a combination of viewing fields that overlap to form a composite field. Within this viewing field, the signal from an object with a single specular reflection varies according to the length of the incident wave, the angle of wave incidence, and the angle of reception at the receiving aperture. The instantaneous value (amplitude) of the signal carries no more information than the presence and radial distance to the reflection. The angle of reception cannot be calculated from the instantaneous difference in amplitude between the two signal channels because the difference in the amplitude between the two channels varies during an FM sweep.

It has been established that adding a narrow central field to two peripheral sensing fields produces greater bio-acoustic spatial resolution than for only the peripheral fields operating for the situations reported here. The auditory system apparently processes the new complex sensory information in a way that needs some detailed explanation. The bio-acoustic system appears to have valuable sensing features for a developing blind child which require further in-depth studies. Recent advances in neuroscience are beginning to explain why the basic spatial sensing features are quickly learned and the ability to use them is memorized.

## ACKNOWLEDGMENTS

This work has involved many persons who have made significant contributions; only a few can be mentioned in the references. The late George Clark, an experimenter and a craftsman, who worked closely with me over a period of years developing the transducers, made the most significant contribution.

<sup>1</sup>L. Kay, "A sonar aid to enhance spatial perception of the blind: engineering design and evaluation," *Radio Electron. Eng.* **44**, 605–627 (1974).

<sup>2</sup>L. Kay, "Air sonars with acoustical display of spatial information," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York, 1980), pp. 769–816.

<sup>3</sup>P. W. Araisan, "Evaluation of the Binaural Sensory Aid," *Research Bulletin No. 26* (American Foundation for the Blind, New York, 1973), pp. 51–71.

<sup>4</sup>C. K. Henkel, "The auditory system," in *Fundamental Neuroscience*, edited by D. E. Haines (Churchill Livingstone, New York, 1997), pp. 285–301.

<sup>5</sup>S. Aitkin and T. G. R. Bower, "Intersensory substitution in the blind," *J. Experimental Child Psychol.* **33**, 309–323 (1982).

<sup>6</sup>E. Sampaio and J. L. Dufier, "Suppléance sensorielle électronique pour les jeunes enfants aveugles," *J. Fr. Ophthalmol.* **11**, 161–167 (1988).

<sup>7</sup>E. R. Strelow, "The use of the Binaural Sensory Aid by preschool children," *J. Visual Impairment Blindness* **77**, 429–438 (1983).

<sup>8</sup>K. A. Ferrell, "Can infants use the Sonicguide? Two years experience of project VIEW," *J. Visual Impairment Blindness* **74**, 209–220 (1980).

<sup>9</sup>E. R. Strelow and R. M. Hodgson, "The development of a spatial sensing system for blind children," *The New Outlook* **70**, 22–24 (1976).

<sup>10</sup>M. A. Do and L. Kay, "Resolution in an artificially generated multiple object auditory space using new auditory sensations," *Acustica* **36**, 9–15 (1976/77).

<sup>11</sup>J. A. Brabyn and E. R. Strelow, "Computer analysed measures of human locomotion and mobility," *Behav. Res. Methods Instrum.* **9**, 456–462 (1977).

<sup>12</sup>E. Thelen and L. B. Smith, *A Dynamic Systems Approach to the Development of Cognition and Action* (The MIT Press, London, 1994).

<sup>13</sup>D. Rowell, "Auditory display of spatial information," Doctoral thesis, University of Canterbury, New Zealand (1970).

- <sup>14</sup>J. Keith, "Adaptation to distortion of auditory space," Masters thesis, University of Auckland, New Zealand (1971).
- <sup>15</sup>L. Kay, Outer/Inner Beam Sonar or Radar. US Patent 4292678 (1981).
- <sup>16</sup>L. Kay and M. A. Do, "An artificially generated multiple object auditory space for use where vision is impaired," *Acustica* **36**, 1–8 (1976/77).
- <sup>17</sup>G. Hornby, L. Kay, M. Satherly, and N. Kay, "Spatial awareness training of blind children using the Trisensor," in *Electronic Spatial Sensing for the Blind*, edited by D. H. Warren and E. R. Strelow (Martinus Nijhoff, Boston, 1985), pp. 257–272.
- <sup>18</sup>J. A. Brabyn and E. R. Strelow, "Computer-analysed measures of characteristics of human locomotion and mobility," *J. Visual Impairment Blindness* **71**, 210–213 (1977).
- <sup>19</sup>E. Zwicker and H. Fastl, *Psychoacoustics—Facts and Models* (Springer-Verlag, London, 1990), pp. 163–167.

# Maximum *a posteriori* deconvolution of ultrasonic signals using multiple transducers

Tomas Olofsson and Tadeusz Stepinski

Signals and Systems Group, Department of Material Science, Uppsala University, Box 528, 751 20 Uppsala, Sweden

(Received 18 March 1999; revised 23 December 1999; accepted 28 January 2000)

A new method for deconvolution of ultrasonic pulse-echo measurements employing multiple-transducer setup is proposed in the paper. An optimal way of estimating the material reflection sequence for a linear signal generation model using maximum *a posteriori* estimation is proposed. The method combines the measurements from a number of transducers covering different frequency bands yielding an optimal estimate of the reflection sequence. The main idea of this approach is to complement the information unavailable from one transducer in some frequency bands with the information from the other transducers. The method is based on the assumption that the measurements are performed using transducers with identical apertures and apodization, which are located exactly at the same position relative to the test object during the measurement. An error analysis presented in the paper proves that when the above assumptions are fulfilled, the proposed method, by utilizing more data for estimation, consistently yields more accurate reflection sequence estimates than the classical Wiener filter. Experimental evidence is presented using both simulated and real ultrasonic data as a verification of the correctness of the multiple-transducer model and the estimation scheme. An illustration of the advantages of the method is also given using real ultrasonic data. © 2000 Acoustical Society of America. [S0001-4966(00)02105-6]

PACS numbers: 43.60.Pt, 81.70.Cv [JCB]

## INTRODUCTION

An ultrasonic signal measured in pulse echo-mode (an A scan) can be modeled as a noise corrupted convolution between the transducer's lumped electro-mechanical impulse response and the so-called reflection sequence (RS). Hence, the estimation of the RS is an example of a deconvolution problem. Traditionally, deconvolution of ultrasonic pulse-echo signals found in nondestructive testing, NDT, have been performed for single transducers only.<sup>1-4</sup> Since the RS usually has a much broader bandwidth than the impulse response of the ultrasonic transducer, the deconvolution is an ill-posed problem. Even though the solution space is commonly restricted by introducing some *a priori* assumptions of the RSs, the accuracy of the estimates of the RSs will usually be poor. For commonly used transducers characterized by relatively narrow bands, this problem seems to be inevitable. Unfortunately, we always face a trade-off between two of the factors that determine the quality of the estimates: signal-to-noise ratio (SNR), and bandwidth (high energy excitation versus heavy damping). For broadband transducers the main problem may instead be poor sensitivity and, thus, poor SNR.

The main idea developed in this paper consists of increasing the signal bandwidth by using a multiple-transducer setup. By using several transducers the total amount of energy of the interrogating pulses could be significantly increased, for instance, by choosing a number of relatively narrow-band transducers covering successive frequency bands resulting in an increased SNR. Also, the information lost in one transducer in some frequency bands would be replaced with the information from the other transducers.

The main problem encountered when using a multiple-

transducer setup is finding the way of merging output signals of the individual transducers yielding the best information about the RS. This issue is formulated and solved in the paper in terms of finding an optimal estimate of the RS given a linear signal generation model.

A feasible method for estimation of the RS from a multiple-transducer measurement should have the following properties:

- (1) The estimation error should decrease as the number of measurements increases, and it should approach zero as the number of measurements goes to infinity.
- (2) It should be robust against deviations of transducer parameters (such as impulse responses, apertures, etc.) from the assumed values.
- (3) Parameters required by the method should be easy to tune.
- (4) The computation time should be short enough to facilitate on-line processing.

In this paper we propose a method of estimating the RS based on several A scans acquired at the same point but using different transducers. First, we will introduce a linear signal generation model defining the measured signal as a convolution of the transducer's lumped impulse response and the RS. Then, we formulate the deconvolution problem as maximum *a posteriori* (MAP) estimation of the RS. A method for deconvolving images of the same object obtained using several sensors has been reported by Yaroslavsky,<sup>5</sup> and a similar approach, combined with multi-scale (wavelet) techniques, has been reported.<sup>6</sup> However, to the authors' knowledge, no work has been reported on RS estimation using multiple ultrasonic transducers.

We will show that the proposed method possesses all the desired properties listed above. We assume that a convolution model for our transducer is valid following the scalar diffraction theory as reviewed by Fink *et al.*<sup>7</sup> To achieve the ideal performance of the proposed method, the transducers should have identical apertures and apodization, in other words, they differ only in surface velocity waveforms. Also means for exact positioning the transducers are needed. We will show that under such conditions the first property is fulfilled.

Concerning the second property, we will show that the deviations from assumed conditions can be treated by introducing a model error term with a variance that reflects the amount of deviation from the ideal conditions.

The proposed deconvolution method assumes knowledge of impulse responses from each transducer (can be measured), variance of the measurement noise (can be estimated), and variance of model error. The latter variance is difficult to find empirically but can be estimated, at least roughly, by means of simulations.

Finally, some comment should be made concerning the calculation time. The RS estimate is a linear function of the measured data so the calculations can be performed fast. Taking into account that the speed of scanning in ultrasonic systems is limited by the pulse repetition frequency (PRF), we can show that the calculations involved in the estimation can be performed in about the same time that it is needed for performing the measurements and, thus the fourth property can also be fulfilled.

The paper is organized as follows: In the first section we present the signal generation model together with statistical models of the measurement noise, model error noise, and the RS. In the subsequent section a MAP estimate is derived and the covariance matrix of the estimation error is presented. To obtain better understanding of the estimation accuracy properties of the method, the estimation error is studied in more detail for a simple special case assuming ideal conditions, and some computational aspects are also presented. The paper continues with a section where results obtained with simulated ultrasonic data are presented, illustrating the properties of the method. In the subsequent section we show results obtained using two different sets of real ultrasonic data for testing the method. The first data set is used for verifying the model correctness and the second is used for illustrating the benefits of the method. Some practical considerations are also discussed. The paper ends with conclusions and some guiding lines for designing a measurement setup to make better use of the advantages of the method.

## I. SIGNAL MODEL

In this section we first present a physically based signal generation model and then the statistical models of the involved signals are presented.

### A. Signal generation model

In the following we consider planar transducers vibrating in thickness mode only. We assume a linear propagation acoustic model and linear transducers which implies a linear and time invariant relation between the received pulse-echo

signal,  $y(t)$ , and the transducer excitation pulse,  $e_{xc}(t)$ . The signal generation model in pulse echo testing can then be written as

$$y(t) = h_{em}^b(t) * s(t) * h_{em}^f(t) * e_{xc}(t), \quad (1)$$

where  $h_{em}^f(t)$  and  $h_{em}^b(t)$  are the forward and backward mode electro-mechanical impulse responses, respectively. These two impulse responses map the electrical input to outgoing surface normal velocity, and incoming velocity potential (and implicitly pressure) to electrical output, respectively.  $s(t)$  is the RS of the test object. It is defined as the response that would be measured if the lumped scalar transducer impulse response,  $h(t)$ , equals Dirac's delta function, i.e., if

$$h(t) = e_{xc}(t) * h_{em}^f(t) * h_{em}^b(t) = \delta(t). \quad (2)$$

It is important to realize that  $s(t)$  only depends on the relative position, its aperture, and apodization of the transducer, but not on  $e_{xc}(t)$  and  $h_{em}^f(t)$ , as well as  $h_{em}^b(t)$ . All information concerning frequency dependencies of responses from different field points is captured in  $s(t)$ . As a simple example of this statement we consider the received signal,  $y(t)$ , from a reflection from a small size scatterer located at position  $\mathbf{r}$ . We can write  $y(t)$  as<sup>7</sup>

$$y(t) = h_{em}^b(t) * h_s^b(t, \mathbf{r}) * [R(\mathbf{r})u(t)] * h_s^f(t, \mathbf{r}) * h_{em}^f(t) * e_{xc}(t), \quad (3)$$

where  $R(\mathbf{r})$  is the scalar reflectivity at position  $\mathbf{r}$ ,  $u(t)$  is the inverse Fourier transform of the frequency response of the scatterer, and  $h_s^b(t, \mathbf{r})$  and  $h_s^f(t, \mathbf{r})$  are the backward and forward spatial impulse responses which model diffraction. Clearly, in this example the reflection sequence becomes  $s(t) = h_s^b(t, \mathbf{r}) * [R(\mathbf{r})u(t)] * h_s^f(t, \mathbf{r})$ .

Since convolution is a commutative operator, the order of the convolutions in Eq. (1) can be altered and by using the lumped impulse response  $h(t) = e_{xc}(t) * h_{em}^f(t) * h_{em}^b(t)$  we can rewrite the signal generation model as

$$y(t) = h(t) * s(t). \quad (4)$$

An approximation of the impulse response  $h(t)$  can be found by means of direct normal incidence pulse-echo measurements from a planar target in the far field.<sup>8</sup> However, this measurement procedure introduces an error which can be modeled as the convolution  $h_m(t) = h(t) * H_v(d, a, t)$ . Here  $h_m(t)$  is the measured impulse response,  $H_v(d, a, t)$  is the radiation coupling function,  $d$  is the distance between the transducer and the plane, and  $a$  is the transducer radius. The model error can be corrected using standard deconvolution methods.<sup>8</sup> The radiation coupling function corresponding to the measurements of the impulse responses used in this work was close to Dirac's function and no error correction was made.

Assume now that we have  $J$  transducers,  $T_1, \dots, T_J$ , of piston type that all have the same aperture and apodization but otherwise different electro-mechanical properties, crystal thicknesses and backing, and therefore different lumped impulse responses. Also assume that A scans from each of  $J$  transducers can be obtained for each position relative to the



test object. (Note that the measurements do not necessarily have to be performed at the same time, one possible transducer configuration is presented in Fig. 6.)

The time-discrete signals measured by the transducers, as those described above, can be described approximately using the convolutional model

$$y_j(k) = h_j(k) * s_j(k) + e_j(k), \quad (5)$$

where  $k$  is the time index,  $y_j(k)$  is the signal measured by  $T_j$ ,  $e_j(k)$  is the measurement noise,  $h_j(k)$  is the sampled lumped impulse response of  $T_j$ , and  $s_j(k)$  is the RS associated with  $y_j(k)$ .

Possible model errors are accounted for by splitting the RS into two parts:

$$s_j(k) = s(k) + v_j(k), \quad (6)$$

where  $s(k)$  is a part of the RS that is common to all the transducers used and where  $v_j(k)$  is a noise term that includes possible model errors, such as positioning errors or errors due to differences in apertures of the transducers. We should note that the role of  $s(k)$  in Eq. (6) is, at least partly, defined by the user. For example, by setting  $v_l(k)$  identical to zero, we implicitly say that  $T_l$  plays the role of a reference transducer and the model errors contained in  $v_j(k)$ ,  $j \neq l$  are taken relative to the RS corresponding to measurements obtained using  $T_l$ .

To illustrate the idea of the model error noise term we simulated some A scans from a 1-mm-long 1D crack positioned 30 mm below the transducers. The A scans were calculated using the simulation program FIELD II, developed by Jensen.<sup>9,10</sup> The simulation package is based on the spatial impulse responses method<sup>11,7</sup> and performs calculations in time domain. The transducers were defined to be of circular piston type and their lumped impulse responses were Dirac functions so that the A scans and the RSs are equivalent. The sound speed was set to 6320 m/s [corresponding to longitudinal waves in Eq. (A1)]. Three different situations were treated:

- Ex. A: Two transducers, both with radius 4 mm and 4.1 mm, respectively.
- Ex. B: Two transducers, both with radius 4 mm but at slightly different position.
- Ex. C: Two transducers with the same radius. One of the transducers is parallel to the crack, the other is positioned by a  $1^\circ$  angle with respect to the former.

The geometry of the measurements is given in Fig. 1.

The resulting A scans are shown in Fig. 2. The difference between the pairs of A scans are also shown. This difference signal illustrates the earlier mentioned idea of a having a reference transducer, so in these examples we have that  $v_1(k) \equiv 0$ ,  $s_1(k) = s(k)$  (solid, upper panels),  $s_2(k) = s(k) + v_2(k)$  (dash-dotted, upper panels) and, consequently,  $v_2(k) = s_2(k) - s_1(k)$  (lower panels).

When treating signals of finite length, it is often convenient for analytical purposes to use matrix/vector notation instead of convolutions. This also has the advantage that nonstationary processes, such as transients, can be easily treated. In matrix form Eqs. (5) and (6) become

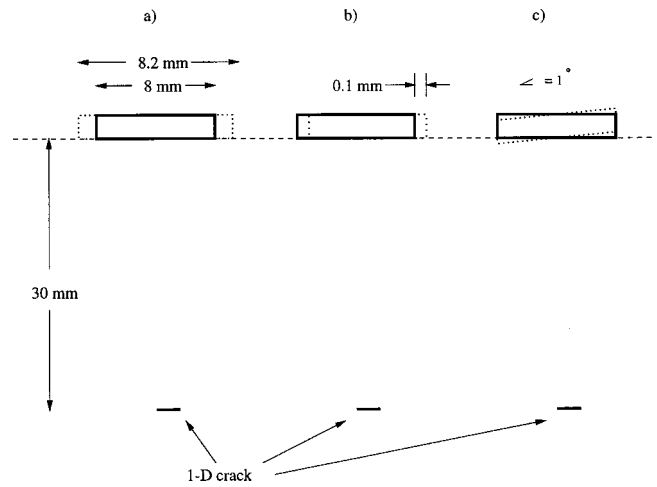


FIG. 1. Geometry of the measurements resulting in the simulated signals shown in Fig. 2. The transducers crystals are represented by rectangles.  $T_1$  is shown as solid and  $T_2$  is shown as dash-dotted. Note that the scale is not preserved in the figure. (a) Two transducers with radius 4 mm (solid) and 4.1 mm (dash-dotted). The reflector is a 1-mm 1D crack at 30 mm. (b) Two transducers with the same radius (4 mm) but at different positions. (c) Two nonparallel transducers. The angle is  $1^\circ$ .

$$\mathbf{y}_j = \mathbf{H}_j \mathbf{s}_j + \mathbf{e}_j \quad (7)$$

and

$$\mathbf{s}_j = \mathbf{s} + \mathbf{v}_j, \quad (8)$$

where bold, lowercase letters represent column vectors, e.g.,  $\mathbf{y}_j = [y_j(1), y_j(2), \dots, y_j(N)]^T$ , and  $\mathbf{H}_j$  is a so-called convolution matrix or finite impulse response (FIR) matrix. These matrices have the corresponding impulse response as their first column and, furthermore, have Toeplitz structure (see Appendix A for a detailed definition).

## B. Statistical models of $\mathbf{e}_j$ , $\mathbf{s}$ and $\mathbf{v}_j$

To be able to perform MAP estimation, statistical models of  $\mathbf{e}_j$ ,  $\mathbf{s}$ , and  $\mathbf{v}_j$  are needed. In this paper we assume all these vectors to be zero mean Gaussian. The Gaussian assumption is made mainly with computational aspects in mind because the corresponding criterion function is then quadratic and, thus the estimate is a linear function of the measured data. Any other assumption will lead to more cumbersome calculations. Note that modifications needed for treating nonzero mean processes are straightforward because the criterion function corresponding to the estimate will still be quadratic in such a case.

We denote the covariance matrices of  $\mathbf{e}_j$ ,  $\mathbf{v}_j$ , and  $\mathbf{s}$  as  $\mathbf{C}_{e_j}$ ,  $\mathbf{D}_j$ , and  $\mathbf{C}_s$ , respectively. Furthermore, we assume that the vectors  $\mathbf{e}_1, \dots, \mathbf{e}_J$ ,  $\mathbf{v}_1, \dots, \mathbf{v}_J$ , and  $\mathbf{s}$  are all mutually independent.

Some comments should be made concerning these Gaussian assumptions. The measurement noise is a combination of electrical noise (thermal) and quantization noise. These noise terms are usually considered to be Gaussian and uniformly distributed respectively. Since both these distributions have short "tails," the distribution of a combination (sum) of samples from such distributions should also have relatively short tails meaning that the combined distribution could be approximated by a Gaussian with good accuracy.

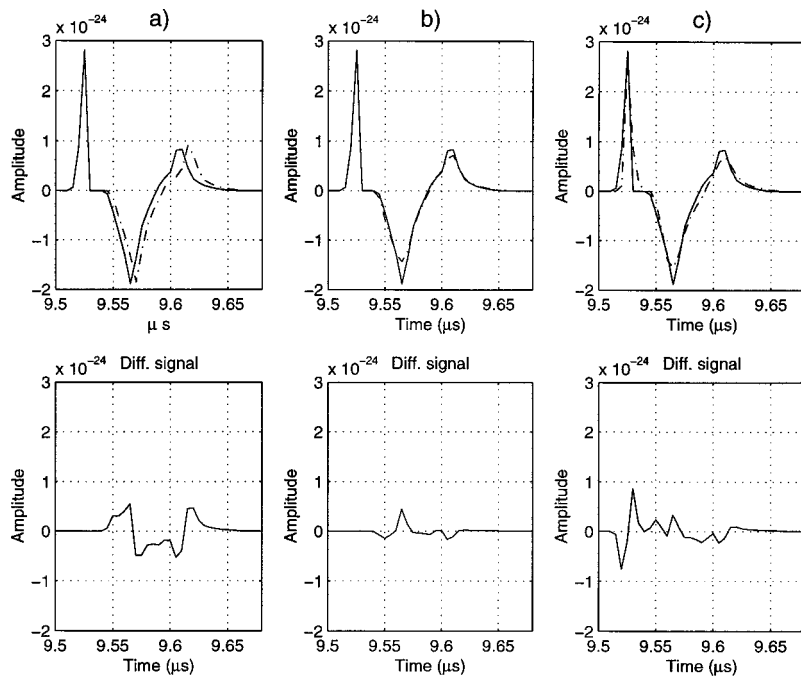


FIG. 2. Simulated A scans having geometry shown in Fig. 1. The A scans from each example are shown together in the upper panels and the difference signals are shown in the lower panels. The transducers were assumed to have a Dirac pulse as impulse response. The reflector is a 1-mm 1D crack at 30 mm. (a) A scans from two transducers with radius 4 mm (solid) and 4.1 mm (dash-dotted). (b) A scans from two transducers at slightly different position. The difference in position is 0.1 mm. (c) A scans from two nonparallel transducers with the same radius (4 mm). The angle between the transducers is 1°.

Furthermore, if the anti-aliasing filtering is performed properly, the components in  $\mathbf{e}_j$  should be almost uncorrelated and a diagonal covariance matrix could be used as a good approximation, i.e.,  $\mathbf{C}_j = \sigma_j^2 \mathbf{I}$ , where  $\sigma_j^2$  is the variance of the amplitude and  $\mathbf{I}$  is the identity matrix. The variance  $\sigma_j^2$  can easily be estimated using measured A scans containing no reflected signals, e.g., water path signals if we consider immersion tests.

The Gaussian assumption on the RS is perhaps more questionable. Due to the central limit theorem and the superposition principle, it may be a relatively good assumption in cases where the material consists of many small reflectors, for instance, copper or cast steel. In materials containing only a few large reflectors, the RS should, on the other hand, consist of sparsely distributed spikes (or more correctly bursts, cf. Fig. 2) and the probability density function (pdf) of the amplitude should have heavier tails than the Gaussian. In such a case the Gaussian density cannot be considered as a good approximation. Work has been done to find good *a priori* distribution for the reflection amplitudes; see Neal and Thompson.<sup>12,13</sup> Results from this work could perhaps be used in this approach. However, one of the aims of the proposed multi-transducer approach is to reduce the need for very realistic *a priori* assumptions about  $\mathbf{s}$  by utilizing more data for the estimation. This is simply because the influence of *a priori* information on the estimate is reduced as the amount of observation data increases which means that the effect of choosing an incorrect pdf will not be as severe as it may be in the single measurement case.

Finally, the motivation of a Gaussian model for  $\mathbf{v}_j$  follows the same line of argument as that concerning  $\mathbf{s}$ . This is because of the linear relation between  $\mathbf{s}_j$  and  $\mathbf{v}_j$ , so whenever the Gaussian assumption is valid for  $\mathbf{s}$ , it will be valid for  $\mathbf{v}_j$ .

## II. ESTIMATION OF THE RS

In this section we first derive an analytical expression of the MAP estimate associated with the model given in the earlier section. Second, we present the covariance matrix of the estimation error and subsequently we treat a simple special case in order to illustrate the accuracy properties of the method. Finally some computational aspects are presented.

### A. Derivation of MAP estimate

The MAP estimate of  $\mathbf{s}$  given the set of measured signals  $\{\mathbf{y}_1, \dots, \mathbf{y}_J\}$  is defined as

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s}} p(\mathbf{s} | \mathbf{y}_1, \dots, \mathbf{y}_J) = \arg \max_{\mathbf{s}} \frac{p(\mathbf{y}_1, \dots, \mathbf{y}_J | \mathbf{s}) p(\mathbf{s})}{p(\mathbf{y}_1, \dots, \mathbf{y}_J)}$$

$$= \arg \max_{\mathbf{s}} p(\mathbf{y}_1, \dots, \mathbf{y}_J | \mathbf{s}) p(\mathbf{s}), \quad (9)$$

where we have used Bayes' theorem together with the fact that the denominator does not depend on  $\mathbf{s}$ . The *a priori* pdf,  $p(\mathbf{s})$ , is given by

$$p(\mathbf{s}) = \frac{1}{(2\pi)^{N/2} (\det \mathbf{C}_s)^{1/2}} e^{-1/2 \mathbf{s}^T \mathbf{C}_s^{-1} \mathbf{s}}, \quad (10)$$

where  $\det \mathbf{C}_s$  is the determinant of  $\mathbf{C}_s$ .

Since we have assumed that all vectors  $\mathbf{e}_1, \dots, \mathbf{e}_J$ ,  $\mathbf{v}_1, \dots, \mathbf{v}_J$ , and  $\mathbf{s}$  are mutually independent, the joint conditional density  $p(\mathbf{y}_1, \dots, \mathbf{y}_J | \mathbf{s})$  can be factorized as

$$p(\mathbf{y}_1, \dots, \mathbf{y}_J | \mathbf{s}) = \prod_{j=1}^J p(\mathbf{y}_j | \mathbf{s}). \quad (11)$$

Furthermore,  $p(\mathbf{y}_j | \mathbf{s})$  is given by the multivariate integral

$$\begin{aligned}
p(\mathbf{y}_j|\mathbf{s}) &= \int p(\mathbf{y}_j|\mathbf{s}, \mathbf{v}_j) p(\mathbf{v}_j) d\mathbf{v}_j \\
&= \int p_j(\mathbf{y}_j - \mathbf{H}_j(\mathbf{s} + \mathbf{v}_j)) p(\mathbf{v}_j) d\mathbf{v}_j, \quad (12)
\end{aligned}$$

where  $p_j(\cdot)$  is the density function of  $\mathbf{e}_j$ . Since we assume both  $\mathbf{v}_j$  and  $\mathbf{e}_j$  to be normal distributed, the integral in Eq. (12) can be easily solved analytically yielding

$$p(\mathbf{y}_j|\mathbf{s}) = \frac{1}{(2\pi)^{N/2} (\det \tilde{\mathbf{C}}_j)^{1/2}} e^{-1/2[(\mathbf{y}_j - \mathbf{H}_j\mathbf{s})^T \tilde{\mathbf{C}}_j^{-1} (\mathbf{y}_j - \mathbf{H}_j\mathbf{s})]}, \quad (13)$$

where

$$\tilde{\mathbf{C}}_j = \mathbf{C}_j + \mathbf{H}_j \mathbf{D}_j \mathbf{H}_j^T. \quad (14)$$

A suitable criterion function corresponding to the MAP estimation problem is given by the negative logarithm of  $p(\mathbf{y}_1, \dots, \mathbf{y}_J|\mathbf{s}) p(\mathbf{s})$ . Utilizing the relations in Eqs. (11) and (13) and excluding all terms that do not depend on  $\mathbf{s}$ , we obtain the criterion function

$$J(\mathbf{s}) = \frac{1}{2} \sum_{j=1}^J [(\mathbf{y}_j - \mathbf{H}_j\mathbf{s})^T \tilde{\mathbf{C}}_j^{-1} (\mathbf{y}_j - \mathbf{H}_j\mathbf{s})] + \frac{1}{2} \mathbf{s}^T \mathbf{C}_s^{-1} \mathbf{s}. \quad (15)$$

Since the criterion function in Eq. (15) is quadratic in  $\mathbf{s}$  the corresponding minimization problem can easily be solved analytically. The solution is given by

$$\hat{\mathbf{s}} = \left( \sum_{j=1}^J \mathbf{H}_j^T \tilde{\mathbf{C}}_j^{-1} \mathbf{H}_j + \mathbf{C}_s^{-1} \right)^{-1} \sum_{j=1}^J \mathbf{H}_j^T \tilde{\mathbf{C}}_j^{-1} \mathbf{y}_j. \quad (16)$$

It is important to note that the estimator in Eq. (16) is a *fixed interval* estimator. As a consequence, the corresponding estimates will show noncausal behavior. For instance, the estimates of a sharp pulse (such as a front wall echo) will usually be preceded by a small ripple.

## B. Estimation accuracy

The estimate given in Eq. (16) is linear in the measurements. It is then straightforward to derive the covariance matrix of the estimation error. Let us introduce the matrix  $\mathbf{U}$  defined as

$$\mathbf{U} = \sum_{j=1}^J \mathbf{H}_j^T \tilde{\mathbf{C}}_j^{-1} \mathbf{H}_j + \mathbf{C}_s^{-1} \quad (17)$$

and the matrices  $\mathbf{F}_j$  defined as

$$\mathbf{F}_j = \mathbf{H}_j \tilde{\mathbf{C}}_j^{-1} \quad j = 1, \dots, J. \quad (18)$$

Using these substitutions, the covariance matrix of the estimation error,  $\tilde{\mathbf{s}}$ , becomes

$$\begin{aligned}
\mathbf{C}_{\tilde{\mathbf{s}}} &= E[\hat{\mathbf{s}} - \mathbf{s}][\hat{\mathbf{s}} - \mathbf{s}]^T \\
&= \left( \sum_{j=1}^J \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j - \mathbf{I} \right) \mathbf{C}_s \left( \sum_{j=1}^J \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j - \mathbf{I} \right)^T \\
&\quad + \mathbf{U}^{-1} \sum_{j=1}^J \mathbf{F}_j \tilde{\mathbf{C}}_j \mathbf{F}_j^T \mathbf{U}^{-1}. \quad (19)
\end{aligned}$$

A derivation of the expression in Eq. (19) is given in Appendix B.

## C. An illustrative special case

Assume ideal conditions where the apertures of all transducers are totally identical. Also assume that the positioning problem has been perfectly solved. Then the term  $\mathbf{v}_j$  is identically zero for all  $j = 1, \dots, J$  and, thus,  $\tilde{\mathbf{C}}_j = \mathbf{C}_j$ . Also, in order to obtain easily interpretable results, we assume  $\mathbf{C}_j = \sigma_j^2 \mathbf{I}$  and  $\mathbf{C}_s = \sigma_s^2 \mathbf{I}$ .

Using the above simplifying assumptions, the estimate becomes

$$\hat{\mathbf{s}} = \left( \sum_{j=1}^J \frac{1}{\sigma_j^2} \mathbf{H}_j^T \mathbf{H}_j + \frac{1}{\sigma_s^2} \mathbf{I} \right)^{-1} \sum_{j=1}^J \frac{1}{\sigma_j^2} \mathbf{H}_j^T \mathbf{y}_j. \quad (20)$$

Since all matrix multiplications involved in finding the estimate  $\hat{\mathbf{s}}$  can be approximated using convolutions, it may be convenient to analyze the estimate in the frequency domain, by means of a Fourier transformation of Eq. (20). We obtain (see Appendix C)

$$\begin{aligned}
\hat{S}(\omega) &\approx \frac{\sum_{j=1}^J \frac{1}{\sigma_j^2} H_j^*(\omega) Y_j(\omega)}{\sum_{j=1}^J \frac{1}{\sigma_j^2} H_j^*(\omega) H_j(\omega) + \frac{1}{\sigma_s^2}} \\
&= \frac{\sum_{j=1}^J \frac{1}{\sigma_j^2} H_j^*(\omega) Y_j(\omega)}{\sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2}}, \quad (21)
\end{aligned}$$

where  $\hat{S}(\omega)$ ,  $Y_j(\omega)$ , and  $H_j(\omega)$  are the Fourier transforms of  $\hat{s}(k)$ ,  $y_j(k)$ , and  $h_j(k)$ , respectively.  $H_j^*(\omega)$  is the complex conjugate of  $H_j(\omega)$  and  $|H_j(\omega)|$  is the absolute value of  $H_j(\omega)$ . If we consider the single transducer case ( $J=1$ ) the expression in Eq. (21) can easily be recognized as the celebrated frequency domain Wiener filter with noise desensitizing factor  $\sigma_1^2/\sigma_s^2$ .

The power spectral density (psd) of the estimation error becomes

$$\begin{aligned}
R_{\tilde{\mathbf{s}}, J}(\omega) &= \left( \frac{\sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2}{\sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2}} - 1 \right)^2 \sigma_s^2 \\
&\quad + \frac{\sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2}{\left( \sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2} \right)^2}. \quad (22)
\end{aligned}$$

It is easy to realize that as the number of terms in the sum  $\sum_{j=1}^J 1/\sigma_j^2 |H_j(\omega)|^2$  increases, the sum itself increases and both terms in Eq. (22) converge to zero and, thus the estimation error goes to zero as the number of transducers grows to infinity. Let  $R_{\tilde{\mathbf{s}}, J+1}$  denote the psd of the estimation

TABLE I. Average computation time for the multiplication between the  $N \times N$  matrix  $\mathbf{G}$  and the  $N \times 1$  vector  $\mathbf{y}$ . Average was taken over 1000 runs.

$N$	Average computation time (ms)
100	0.38
200	1.04
300	2.50
500	9.59
1000	61.95

error associated with a transducer setup  $T_1, \dots, T_{J+1}$ , i.e., transducer  $T_{J+1}$  has been added to the original set. By straightforward but lengthy calculations it can be shown that

$$\int_{\omega} (R_{\bar{s}, J+1}(\omega) - R_{\bar{s}, J}(\omega)) d\omega = - \int_{\omega} \frac{\frac{1}{\sigma_{J+1}^2} |H_{J+1}(\omega)|^2 d\omega}{\left( \sum_{j=1}^{J+1} \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2} \right) \left( \sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2} \right)} \leq 0 \quad (23)$$

with equality if and only if  $H_{J+1}(\omega) \equiv 0$ . This shows that under the ideal conditions described earlier, the estimation error for a given set of transducers is always reduced by including another transducer to the set. Particularly, this holds for  $J=1$  which automatically gives a comparison between the classical Wiener filter and the proposed method. Finally, since the reduction in error for a specific frequency grows as  $H_{J+1}(\omega)$  grows we have a simple means for choosing an optimal combination of transducers.

#### D. Computational considerations

If the estimation is to be performed in real time, we should make modifications so that the calculation can be

performed in as short time as possible. By using the matrices  $\mathbf{U}$  and  $\mathbf{F}_j$  defined in Eqs. (17) and (18) we can rewrite the estimate in Eq. (16) as

$$\hat{\mathbf{s}} = \sum_{j=1}^J \mathbf{U}^{-1} \mathbf{F}_j \mathbf{y}_j = \sum_{j=1}^J \mathbf{G}_j \mathbf{y}_j, \quad (24)$$

where the  $N \times N$  matrix  $\mathbf{G}_j = \mathbf{U}^{-1} \mathbf{F}_j$ . This matrix can be calculated in advance so, once these pre-calculations have been performed, the computation time for the estimation of  $\mathbf{s}$  will at most be proportional to  $N \times N \times J$ . We have estimated the average time to compute the product  $\mathbf{G}_j \mathbf{y}_j$  for different  $N$  on a 330-MHz Pentium II computer running under Linux and the calculations performed in Matlab<sup>®</sup>. The average computation time is given in Table I.

Some comments should also be made regarding the hardware. The calculations involved in the estimation are well suited for one or several signal processors. (The reader should note, however, that all calculations involved in this paper were performed off-line.) By distributing the calculations over one signal processor for each transducer, on-line processing is within the reach even for relatively long signals (e.g.,  $N=1000$ ). The single factor that limits the usefulness of the approach would then be the testing time that is proportional to the number of transducers.

### III. DECONVOLUTION OF SYNTHETIC DATA

In this section we illustrate some properties of the proposed method using simulated ultrasonic data. The first example illustrates the fact that for reflected echoes with a large portion of high frequency components, the proposed method will *not* in general yield results that are much better than those results obtained by using a single broadband transducer. Such echoes typically occur for planar reflectors that are parallel to the transducer surface, located in media with negligible frequency dependent attenuation. In practice, however, reflectors are rarely parallel to the transducer sur-

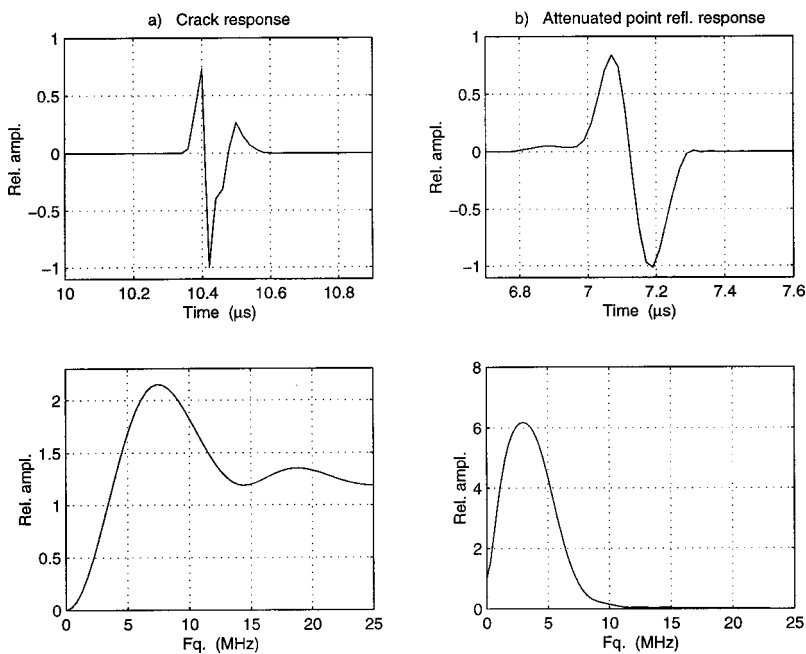


FIG. 3. Simulated A scans from 1-mm 1D crack (left) and from attenuated point reflector (right). The corresponding amplitude spectra are shown in the lower panels.

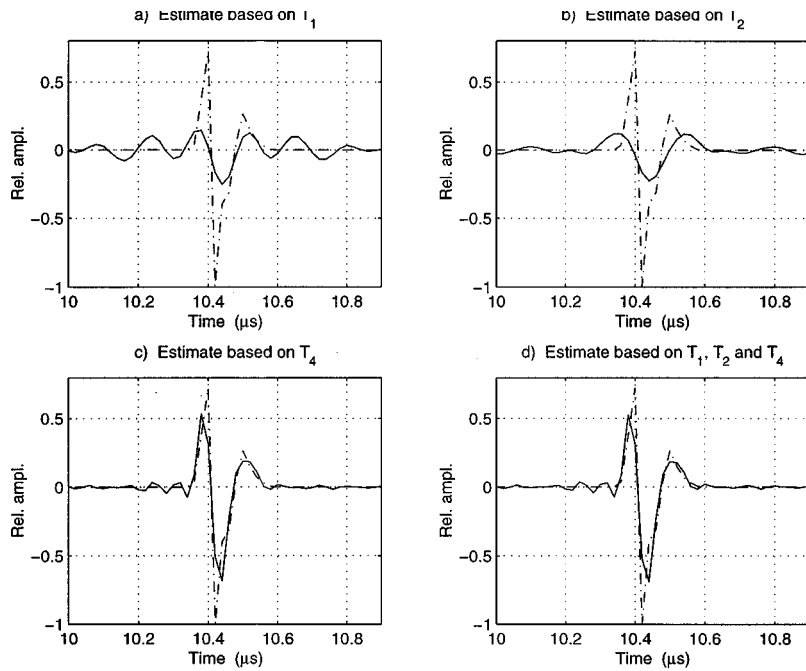


FIG. 4. Estimated RSs (solid) together with the simulated crack response (dash-dotted). The mean estimates are presented. (a) Estimate based on measurements from the single transducer  $T_1$ . (b) Estimate based on  $T_2$ . (c) Estimate based on  $T_4$ . (d) Estimate based on all three transducers.

face and attenuation is usually present, so the reflected signals are often limited to a lower frequency band. For instance, an analysis on how crack responses depend on the crack angle has been presented by Lh emery.<sup>14</sup> The second example in this section illustrates a situation where the broadband transducer is no longer the obviously best choice if single transducer deconvolution is considered. Note that the examples are generated to illustrate the properties of the proposed method, not to illustrate sound modeling results. The most important point is that in both examples, the multi-transducer method yields equally good or better estimate than the best estimates based on only single transducer.

In general we cannot expect to have detailed knowledge of the frequency contents of the RS to be estimated, since

this depends on unknown parameters, such as, crack angle, attenuation, etc. If we had such detailed knowledge, a detection procedure would perhaps be more relevant than an estimation procedure.

The RSs were again calculated using the simulation program FIELD II. The two examples mentioned above are:

- (1) Planar 2-mm-long 1D crack parallel to the surface of the transducers and located 30 mm below the surface (no attenuation).
- (2) Point reflector at a distance 20 mm below the transducers. The elastic waves were subject to a large linear frequency dependent attenuation 0.000 08 dB/[Hz m] (illustrating a highly attenuating plastic material).

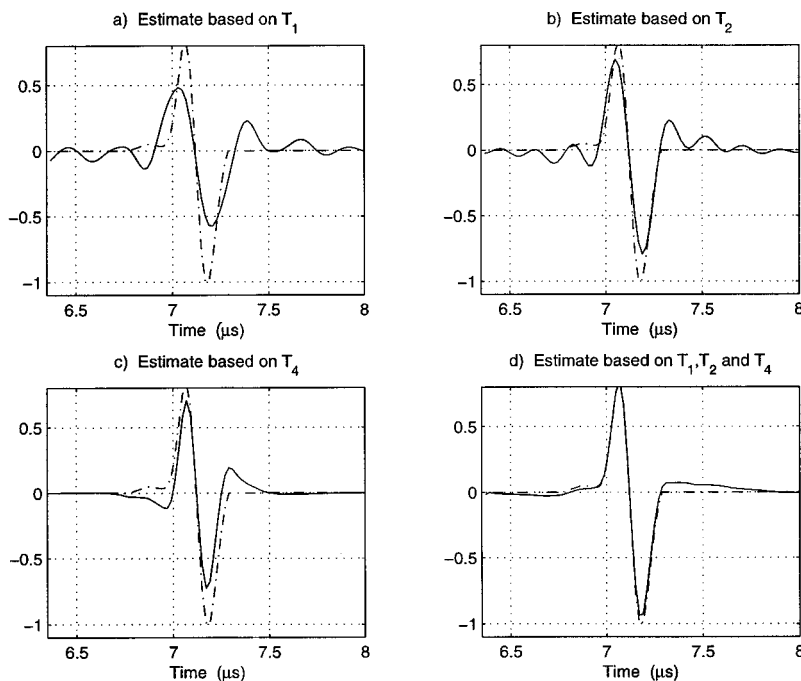


FIG. 5. Estimated RSs (solid) together with the simulated response from a point reflector in an attenuative media (dash-dotted). The mean estimates are presented. (a) Estimate based on measurements from transducer  $T_1$ . (b) Estimate based on  $T_2$ . (c) Estimate based on  $T_4$ . (d) Estimate based on all three transducers.

In Fig. 3 the two simulated RSs are shown (upper panels) together with their corresponding amplitude spectra (lower panels).

The RSs were convolved with the measured impulse responses of three transducers,  $T_1$ ,  $T_2$ , and  $T_4$  [Since the characteristics of  $T_3$  (used in the real data section) is somewhat similar to  $T_2$  and therefore also the corresponding single transducer results, we do not present results corresponding to  $T_3$ .] yielding three simulated A scans. (The time and frequency characteristics of the transducers are shown in Fig. 7 in the subsequent section.) Deconvolution was performed using the single transducers  $T_1$ ,  $T_2$ , and  $T_4$  and the combination  $\{T_1, T_2, T_4\}$ .

No measurement noise was added in the simulation but a  $-20$  dB measurement noise level was *assumed* for all transducers when deconvolving the data. In this way we directly obtained mean values of the deconvolved signals, i.e., the mean with respect to all possible noise realizations. To obtain results that were not influenced too much by *a priori* assumptions, we assumed  $\mathbf{s}$  to be a white stationary sequence, yielding  $\mathbf{C}_s = \sigma_s^2 \mathbf{I}$ . The variance,  $\sigma_s^2$ , was set to the mean squared value of the simulated RSs shown in Fig. 3. The model error vectors,  $\mathbf{v}_j$ , were set to zero.

In Fig. 4 we show the results of deconvolving the A scan from the simulated 1D crack. By comparing the results obtained using  $T_4$  with those obtained with all three transducers, we see that no significant improvement has been made. A significant improvement is, however, not to be expected since the broadband transducer covers the relevant frequencies well and  $T_1$  and  $T_2$  do not contribute much to the overall result. We should note, though, that the multi-transducer result is not worse than the single ( $T_4$ ) result.

In Fig. 5 we show the results from deconvolving the A scan from the point reflector in the attenuative media. Here the reflected signal is dominated by lower frequencies and then  $T_1$  and  $T_2$  contribute more to improve the accuracy compared to the crack example. We can see that the single transducer result obtained with  $T_2$  is of about the same qual-

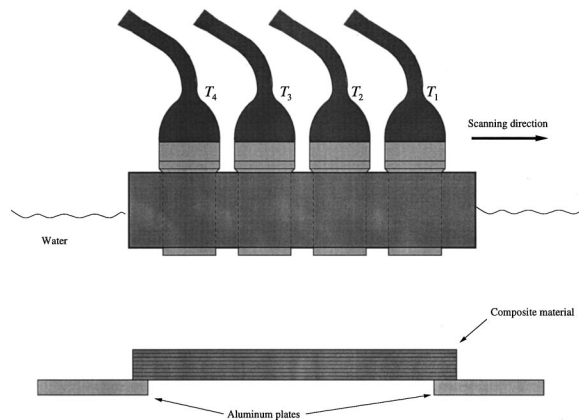


FIG. 6. Schematic measurement setup.

ity, or even better, than results obtained with  $T_4$ . The best result is clearly obtained by the multi-transducer method.

#### IV. DECONVOLUTION OF REAL ULTRASONIC DATA

The purpose of the experiments using real ultrasonic data was twofold. The main purpose was to examine if Eqs. (7) and (8) give a good model of the measurements. Since all calculations are based on these equations, we can get a good indication of the model correctness by comparing the results obtained from measurements obtained using two independent sets of transducers. The only information that is common for these two measurements is the underlying (unknown) RS so if the two corresponding results are similar we have a clear indication that the method is relying on an accurate model. We should note that for real data we do normally not have knowledge about the true  $\mathbf{s}$ . The second purpose was to demonstrate the gain of the multi-transducer approach compared to a single transducer approach in a real measurement situation.

The transducers used in the experiments had not in any way been manufactured with this application in mind; these

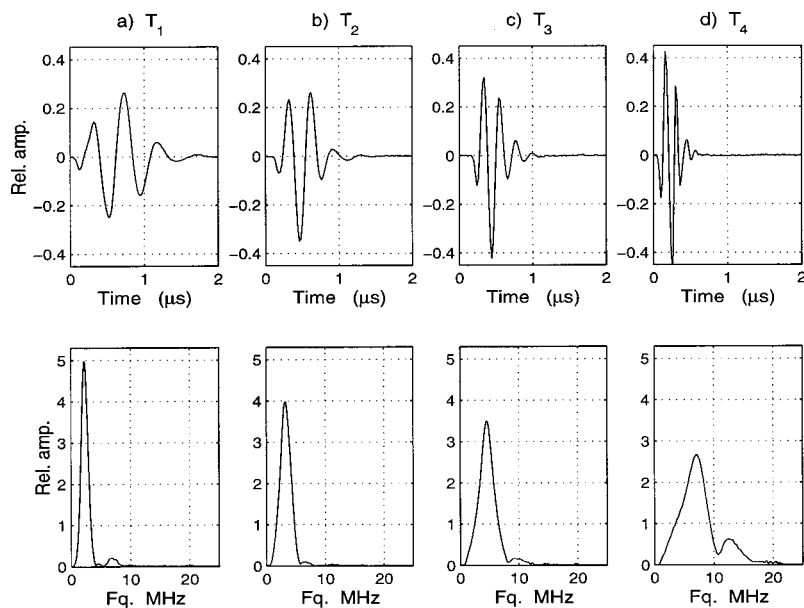


FIG. 7. The impulse responses (upper panels) and corresponding frequency responses (lower panels) of transducers (a)  $T_1$ , (b)  $T_2$ , (c)  $T_3$ , and (d)  $T_4$ . The responses were found by averaging the (synchronized) reflected signals from the front surface of an aluminum plate at distance 55 mm. The responses have been normalized with respect to energy before presentation.

TABLE II. Center frequencies and bandwidths of the transducers used in the experiment.

Transducer	Center frequency (MHz)	3-dB bandwidth (MHz)
$T_1$	2.25	1.35
$T_2$	3.2	2.1
$T_3$	4.5	2.7
$T_4$	7.1	4.6

were standard Panametric nonfocused immersion test transducers with a diameter of 10 mm. Furthermore, we had no means for checking or, even more important, compensating for possible differences in apodization. The transducers were aligned in the scanning direction as indicated in Fig. 6. The distance between the transducers in the setup was chosen as a multiple of the scanning step length, so shifting the B scans this number of A scans relative each other yielded B scans that were covering the same area of the test object.

The time and frequency domain characteristics of the transducers used in the measurements are shown in Fig. 7 and the frequency data are summarized in Table II.

### A. Verification of the model

The first test object was a carbon fiber reinforced epoxy composite material with eight fiber layers. The acquired data from this object are presented as B scans in Fig. 8.

One important issue which ensures obtaining good deconvolution results using the proposed method is synchronization of the measurements in time. In our setup it was possible to measure the impulse responses at the same scanning event as the actual testing. The impulse responses were found by averaging the reflections from the aluminum plate at 1–20 mm in the scanning direction (see Fig. 8). By using this approach we obtained the same internal time delays in both the impulse responses as in the rest of the measurements, so the time synchronization problems could be handled in the best possible way. To obtain high quality

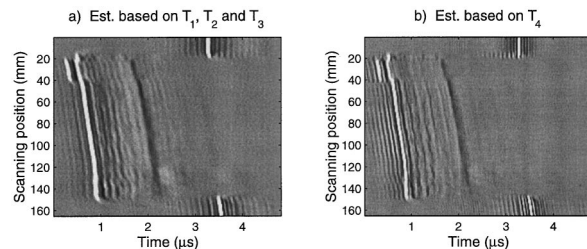


FIG. 9. Estimated RSs presented as B scans. (a) Estimate based on measurements from  $T_1$ ,  $T_2$ , and  $T_3$ . (b) Estimate based on  $T_4$ .

impulse responses, the reflector must be perfectly planar and orthogonal to the incident sound waves. In cases when the impulse responses, due to practical reasons, cannot be measured at the same time as the rest of the data, some other technique, for instance, a cross-correlation technique, must be applied to compensate for synchronization errors.

Deconvolution was in this experiment performed using the proposed method with two different sets of transducers. These were:

- Three transducers,  $T_1$ ,  $T_2$ , and  $T_3$
- Single transducer  $T_4$ .

We assumed the covariance matrices  $\mathbf{C}_j = \sigma_e^2 \mathbf{I}$ ,  $j = 1, \dots, 4$ ,  $\mathbf{D}_j = \sigma_v^2 \mathbf{I}$ ,  $j = 1 \dots 4$ , and  $\mathbf{C}_s = \sigma_s^2 \mathbf{I}$ , where  $\sigma_e^2 = 1$ ,  $\sigma_v^2 = 0.0003$ , and  $\sigma_s^2 = 0.001$

The results are presented in Fig. 9 as B scans and in Fig. 10 estimated reflection sequences corresponding to scanning position 60 mm are shown.

The front echo should after deconvolution appear as a strong and relatively sharp pulse and the eight internal layers in the composite should yield periodic reflections. These periodic reflections should become more and more damped as the sound has traveled a longer way in the attenuative composite for each of these layers. As we mentioned in the earlier section, since the reflections should contain a large por-

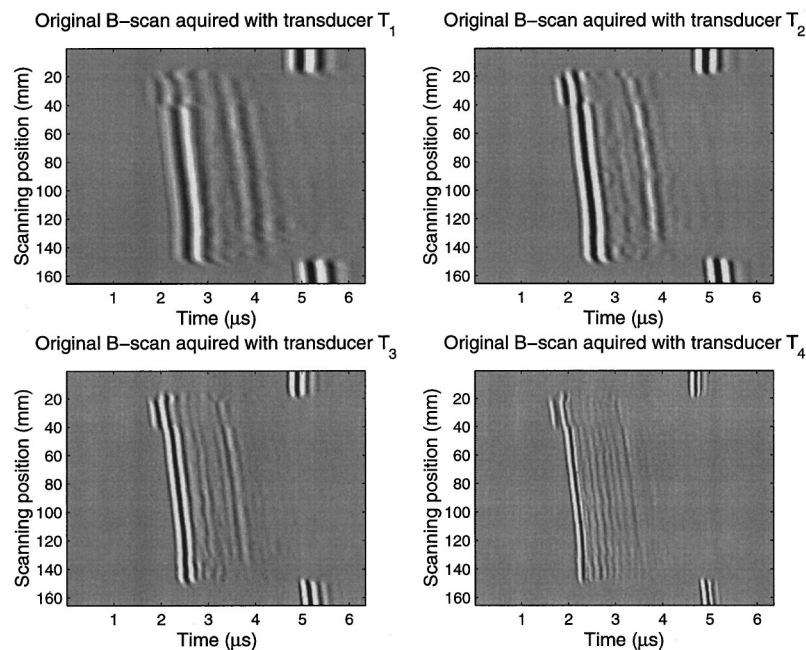


FIG. 8. B scans acquired using measurement setup shown in Fig. 6. At the beginning and end of the scans we can see reflections from the aluminum plates at which the composite material was placed. At scanning positions 25–35 mm we see a reflection from a small metallic marker placed on the composite object. It was used to check that the transducers were aligned correctly.

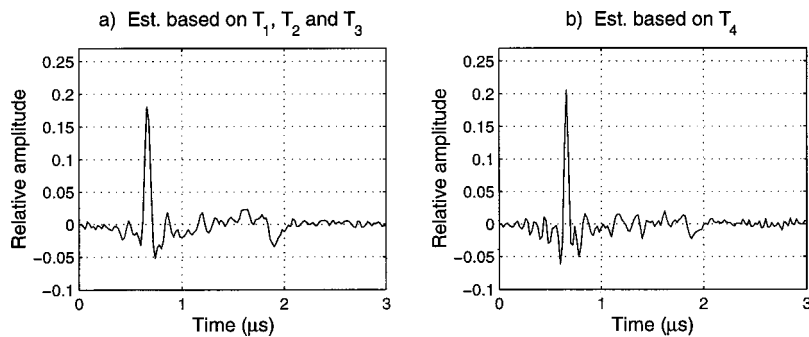


FIG. 10. Example of estimated RSs. The example is from scanning position 60 mm. (a) Estimate based on  $T_1$ ,  $T_2$ , and  $T_3$ . (b) Estimate based on  $T_4$ .

tion of high frequency components (at least the front wall echo), we should not expect the estimation results from using the combination  $T_1$ ,  $T_2$ , and  $T_3$  to be better than the results obtained by using the most broadband transducer, namely  $T_4$ . However, we would expect the SNR to be slightly improved when using several transducers. We could also expect that the reflections that have been subject to heavy frequency dependent attenuation are relatively better estimated using low frequency transducers.

By examining the B scans in Fig. 9 and the estimated RSs in Fig. 10 we see that the estimates obtained using  $T_1$ ,  $T_2$ , and  $T_3$  carry essentially the same information as the estimates obtained using the broadband transducer,  $T_4$ . The resolution is somewhat better for  $T_4$  estimates but this is to be expected since  $T_4$  alone has a larger bandwidth than  $T_1$ ,  $T_2$ , and  $T_3$  together. The estimate of the back wall echo is somewhat larger in the first example. The lower center frequencies of the transducers  $T_1$ ,  $T_2$ , and  $T_3$  should yield better estimates of this reflection and this could explain the larger amplitudes in the estimates corresponding to  $T_1$ ,  $T_2$ , and  $T_3$ .

Since the estimates based on  $T_1$ ,  $T_2$ , and  $T_3$  do not deviate much from the estimates based on the single broadband transducer,  $T_4$ , we conclude that the signal generation

model in Eqs. (7) and (8) is adequate for a multi-transducer deconvolution method to be based upon.

In all estimates shown in Figs. 9 and 10 we can see examples of the noncausal effects mentioned in the end of Sec. II.

### B. Illustration of advantages with the multi transducer approach

Data for the the final experiment were obtained from a side drilled hole in aluminum. The hole radius was 2 mm and the measurements were again performed in immersion with the test object positioned in the far field of all transducers similar to the setup shown in Fig. 6. The impulse responses were this time measured as the normal incidence front surface echoes from the test object. This ensured good synchronization of the impulse responses relative the measurements of the responses from the side drilled hole.

The responses from points on the transducer central axis are more impulsive than the responses from off-axis points due to differences in spatial impulse responses. This means that the reconstruction of responses from on-axis points requires a large total bandwidth compared to the reconstruction of the off-axis responses. In Fig. 11 we show the deconvol-

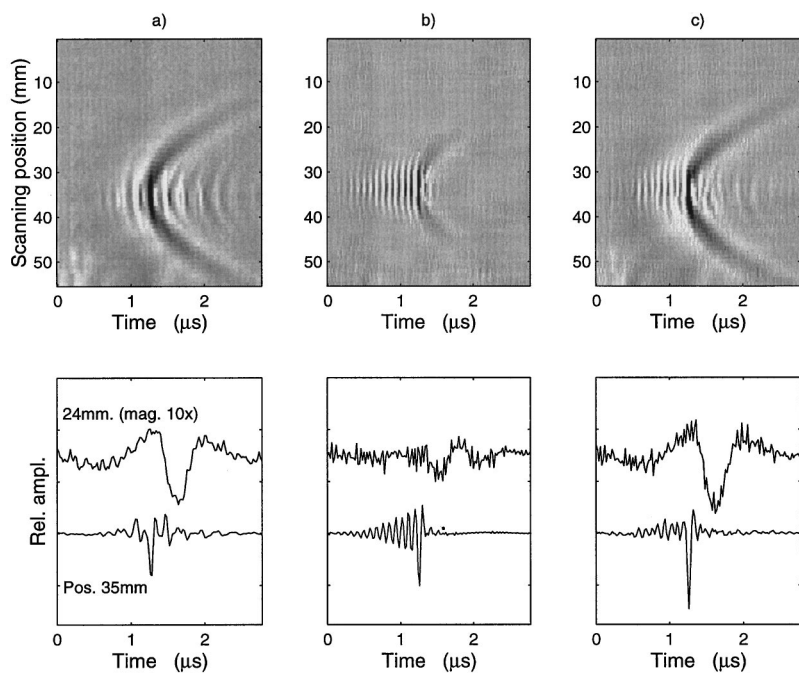


FIG. 11. (a) (upper) Estimated RSs obtained using  $T_1$  presented as B scan and (lower) examples of A scans from lateral positions 24 mm and 35 mm. B scan. (b) Estimated RSs obtained using  $T_4$ . (c) Estimated RSs obtained using  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$ . To separate the A scans in the lower plots, the ones from position 24 mm are offset slightly and furthermore magnified by a factor of 10.



lution results from the side drilled hole for single transducers  $T_1$  and  $T_4$  as well as for the multi-transducer deconvolution result obtained using all four transducers. We can see that  $T_1$ , in comparison with  $T_4$ , gives a larger response to off-axis reflections (lateral pos. 24 mm) whereas the response from the on-axis reflections (pos. 35 mm) is slightly smaller. Some ringing is present in both  $T_1$  and  $T_4$  results. The frequency content of the pulse at 24 mm is concentrated in a low frequency region which  $T_4$ , despite its wide band response, cannot cover, yielding an indistinct response. The multi-transducer estimates are distinct both for on-axis and off-axis responses. Note also that the ringing is reduced by utilizing more transducers. Ringing can be thought of as an effect of the ill-posedness of the deconvolution. Utilizing more data will, if it is done properly, reduce this effect.

The overall behavior is best seen in the B scans in Fig. 11. The  $T_4$  estimates (middle) clearly lack low frequency components to recover the off-axis responses. Using all transducers, however, yields good estimation accuracy for both on-axis and off-axis responses.

## V. CONCLUSIONS AND POSSIBLE IMPROVEMENTS

We have presented a new multi-transducer deconvolution method. The estimation error associated with the method has been studied showing that, by utilizing additional measurements, the method consistently yields estimates with better accuracy than the Wiener filter. The method's limitations are that for optimal performance the transducers must have identical apertures, be perfectly aligned, and be positioned at exactly the same position relative to the test object. If those conditions are not fulfilled, the RS to be estimated is not well defined. We have shown how to treat this ambiguity by incorporation of additional model error noise terms; although such an approach does not utilize measurement data in an optimal manner, it merely robustifies the estimation against model errors. We have shown that the calculations can be performed fast enough to facilitate on-line processing, if they are organized in a proper way and, if possible, implemented using signal processors.

The method has been tested on simulated as well as real ultrasonic data. By using simulated data we have illustrated the (obvious) fact that different testing applications call for different transducers. We have shown, however, that the multi-transducer approach is capable of combining the measurements so that the best features from each measurements are preserved in the final estimation result.

In the first experiment using real ultrasonic data, estimates based on two independent sets of measurements have shown strong resemblance, demonstrating that the model that was used for deriving the multi-transducer approach is accurate enough for this purpose. The results from the second real data experiment have demonstrated that the multi-transducer method yields RS estimates that combine the best features from the different measurements; the temporal resolution properties of the estimate of the on-axis response was determined mainly by the higher frequency broadband transducer

but to obtain an accurate estimate of the off-axis response, measurements from low frequency transducers were needed. However, it should be pointed out that the evaluation of the results obtained with real data is very difficult. There are no simple means for obtaining the *true* RS so we have to rely on some rather simple physical reasoning to interpret the results.

To obtain good estimation results, synchronization in both time and position are important issues. The use of the proposed method would be simplified if these problems could be solved already at transducer manufacturing, for instance, by applying array techniques. Manufacturing the transducers at the same time, using the same equipment, etc., should make it possible to obtain transducers with properties that do not need the use of the robustifying model error noise term described in the paper and further improvements of the estimation results should therefore be possible.

## ACKNOWLEDGMENTS

The authors wish to thank Eider Martinez for performing the measurements and Ping Wu and Mats Gustafsson for valuable comments.

## APPENDIX A: DEFINITION OF FIR MATRICES

An impulse response matrix,  $\mathbf{H}$ , corresponding to a causal impulse response  $\mathbf{h}=[h(0),h(1),\dots,h(K)]$  having  $K+1$  nonzero samples is given by

$$\mathbf{H}=\begin{pmatrix} h(0) & & & & \\ h(1) & h(0) & & & \mathbf{0} \\ \vdots & \ddots & \ddots & & \\ h(K) & \ddots & \ddots & \ddots & \\ 0 & \ddots & \ddots & \ddots & h(0) \end{pmatrix}, \quad (\text{A1})$$

where the elements in the matrix are given by  $\mathbf{H}_{p,q}=h(p-q)$  as seen in Eq. (A1).

## APPENDIX B: DERIVATION OF THE COVARIANCE MATRIX OF THE ESTIMATION ERROR

The matrices  $\mathbf{U}$  and  $\mathbf{F}_j$  used below are defined in Eqs. (17) and (18), respectively. Assuming that the covariance matrices involved in the calculations are correct, and using the substitution  $\tilde{\mathbf{e}}_j=\mathbf{H}_j\mathbf{v}_j+\mathbf{e}_j$ , the covariance matrix of the estimation error can be written

$$\begin{aligned}
\mathbf{C}_s &= E[\hat{\mathbf{s}} - \mathbf{s}][\hat{\mathbf{s}} - \mathbf{s}]^T \\
&= E \left[ \sum_j \mathbf{U}^{-1} \mathbf{F}_j (\mathbf{H}_j \mathbf{s} + \tilde{\mathbf{e}}_j) - \mathbf{s} \right] \\
&\quad \times \left[ \sum_j \mathbf{U}^{-1} \mathbf{F}_j (\mathbf{H}_j \mathbf{s} + \tilde{\mathbf{e}}_j) - \mathbf{s} \right]^T \\
&= E \left\{ \sum_j \sum_l [\mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j \mathbf{s} \mathbf{s}^T \mathbf{H}_l^T \mathbf{F}_l^T \mathbf{U}^{-1} \right. \\
&\quad + \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j \mathbf{s} \tilde{\mathbf{e}}_l^T \mathbf{F}_l^T \mathbf{U}^{-1} + \mathbf{U}^{-1} \mathbf{F}_j \tilde{\mathbf{e}}_j \mathbf{s}^T \mathbf{H}_k^T \mathbf{F}_l^T \mathbf{U}^{-1} \\
&\quad + \mathbf{U}^{-1} \mathbf{F}_j \tilde{\mathbf{e}}_j \tilde{\mathbf{e}}_l^T \mathbf{F}_l^T \mathbf{U}^{-1}] - \sum_j [\mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j \mathbf{s} \mathbf{s}^T + \mathbf{U}^{-1} \mathbf{F}_j \tilde{\mathbf{e}}_j \mathbf{s}^T \\
&\quad \left. + \mathbf{s} \mathbf{s}^T \mathbf{H}_j^T \mathbf{F}_j^T \mathbf{U}^{-1} + \mathbf{s} \tilde{\mathbf{e}}_j^T \mathbf{F}_j^T \mathbf{U}^{-1}] + \mathbf{s} \mathbf{s}^T \right\} \\
&= \left( \sum_j \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j \right) \mathbf{C}_s \left( \sum_j \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j \right)^T \\
&\quad + \left( \sum_j \mathbf{U}^{-1} \mathbf{F}_j \right) \delta_{j,l} \tilde{\mathbf{C}}_j \left( \sum_j \mathbf{U}^{-1} \mathbf{F}_j \right)^T - \sum_j \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j \mathbf{C}_s \\
&\quad - \sum_j \mathbf{C}_s \mathbf{H}_j^T \mathbf{F}_j^T \mathbf{U}^{-1} + \mathbf{C}_s \\
&= \left( \sum_j \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j - \mathbf{I} \right) \mathbf{C}_s \left( \sum_j \mathbf{U}^{-1} \mathbf{F}_j \mathbf{H}_j - \mathbf{I} \right)^T \\
&\quad + \mathbf{U}^{-1} \sum_{j=1} \mathbf{F}_j \tilde{\mathbf{C}}_j \mathbf{F}_j^T \mathbf{U}^{-1}, \tag{B1}
\end{aligned}$$

where  $\delta_{j,l}$  is Kronecker's delta. We have used that the inverse of a symmetric matrix is also symmetric.

### APPENDIX C: FREQUENCY DOMAIN REPRESENTATION OF ESTIMATES AND ESTIMATION ERROR

Here we derive the frequency domain representation of the estimate given in Eq. (20). We are considering all involved processes to be stationary. In the matrix/vector "language" this is equivalent with saying that the corresponding covariance matrices are Toeplitz. We use the convention given in Sec. I, namely that a signal  $x(k)$  corresponds to a column vector  $\mathbf{x} = [\dots, x(k-1), x(k), x(k+1), \dots]^T$ . It is important to note that the result below are only *approximate*. The correspondence between convolutions and matrix/vector products hold when the impulse responses are short compared to the size of the matrices. So, for instance, when we seemingly state below that the inverse of a Toeplitz matrix is also Toeplitz, we merely say that the inverse to a (nearly) Toeplitz matrix is (nearly) Toeplitz provided that this latter matrix is diagonal dominant.

The arbitrary convolution  $z(k) = h(k) * y(k)$  can be represented as product between a Toeplitz matrix and a vector, i.e.,  $\mathbf{z} = \mathbf{H}\mathbf{y}$  (and vice versa). The elements in  $\mathbf{H}$  are given by  $\mathbf{H}_{p,q} = h(p-q)$ .

Since  $\mathbf{H}^T$  is also Toeplitz, the product  $\mathbf{z} = \mathbf{H}^T \mathbf{y}$  can be represented by the (noncasual) convolution  $z(k) = h(-k) * y(k)$  where  $h(-k)$  is a time-reversed impulse response.

A general observation is that matrix multiplications can be seen simply as a sequence of matrix/vector multiplications. Examining the product  $\mathbf{S} = \mathbf{H}^T \mathbf{H}$ , column by column, we see that the elements can be written as the convolutions

$$s_l(k) = h(-k) * h(k+l), \tag{C1}$$

where  $l$  is the column index. The columns in  $\mathbf{S}$  are simply (time) shifted versions of the same column vector and  $\mathbf{S}$  is, thus also a Toeplitz matrix.

Now consider the matrix  $\mathbf{U}$  which is used in the estimate in Eq. (20). First we note that  $\mathbf{U}$ , by construction, is positive definite and, therefore, invertible. This matrix must also be a Toeplitz matrix since it is a sum of Toeplitz matrices. We will show that the product  $\mathbf{a} = \mathbf{U}^{-1} \mathbf{x}$  where  $\mathbf{x}$  is an arbitrary vector, can also be represented by a convolution. This can be done by first considering the product  $\mathbf{x} = \mathbf{U}\mathbf{a}$  which can be written as the convolution

$$x(k) = u(k) * a(k) \tag{C2}$$

and which can be represented in the frequency domain as  $X(\omega) = U(\omega)A(\omega)$ , where  $A(\omega)$ ,  $U(\omega)$ , and  $X(\omega)$  are the Fourier transform of  $a(k)$ ,  $u(k)$ , and  $x(k)$ , respectively. This, of course, means that  $A(\omega) = X(\omega)/U(\omega)$ . Since  $\mathbf{U}$  is invertible we know that  $\mathbf{a}$  exists and is limited.  $A(\omega)$  must then also be limited and, thus we can define a sequence  $u^{-1}(k)$  with corresponding Fourier transform  $U^{-1}(\omega) = 1/U(\omega)$  such that

$$a(k) = u^{-1}(k) * x(k). \tag{C3}$$

This, furthermore, means that the sum of products

$$\hat{\mathbf{s}} = \sum_j \mathbf{U}^{-1} \frac{1}{\sigma_j^2} \mathbf{H}^T \mathbf{y}_j \tag{C4}$$

can be represented by the convolution

$$\hat{s}(k) = \sum_j \frac{1}{\sigma_j^2} u^{-1}(k) * h_j(-k) * y_j(k) \tag{C5}$$

yielding the frequency domain representation

$$\hat{S}(\omega) = \sum_j \frac{1}{U(\omega) \sigma_j^2} H_j^*(\omega) Y_j(\omega). \tag{C6}$$

To find  $U(\omega)$  we just consider one column in  $\mathbf{U}$ . It can be written (again as a convolution)

$$u(k) = \sum_j \frac{1}{\sigma_j^2} h_j(-k) * h_j(k) + \frac{1}{\sigma_s^2} \delta(k), \tag{C7}$$

where  $\delta(k)$  is the Dirac delta function. In the frequency domain this becomes

$$U(\omega) = \sum_j \frac{1}{\sigma_j^2} H_j^*(\omega) H_j(\omega) + \frac{1}{\sigma_s^2}. \quad (\text{C8})$$

By taking the Fourier transform of the expression in Eq. (C5) we finally obtain

$$\hat{S}(\omega) = \frac{\sum_j \frac{1}{\sigma_j^2} H_j^*(\omega) Y_j(\omega)}{\sum_j \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2}}. \quad (\text{C9})$$

The power spectral density of the estimation error,  $R_{\tilde{s}}(\omega)$ , is the Fourier transform of the autocorrelation function of  $\tilde{s}(k)$  which is found as the center column in the matrix given in Eq. (B1). Using the simplifications  $\tilde{\mathbf{C}}_j = \mathbf{C}_j = \sigma_j^2 \mathbf{I}$ , and  $\mathbf{C}_s = \sigma_s^2 \mathbf{I}$  and using that  $\mathbf{F}_j = \mathbf{C}_j^{-1} \mathbf{H}_j = (1/\sigma_j^2) \mathbf{H}_j$ , the matrix can be written as

$$\begin{aligned} E(\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T) &= \sigma_s^2 \left( \sum_j \mathbf{U}^{-1} \frac{1}{\sigma_j^2} \mathbf{H}_j \mathbf{H}_j - \mathbf{I} \right) \\ &\quad \times \left( \sum_j \mathbf{U}^{-1} \frac{1}{\sigma_j^2} \mathbf{H}_j \mathbf{H}_j - \mathbf{I} \right)^T \\ &\quad + \sum_j \mathbf{U}^{-1} \frac{1}{\sigma_j^2} \mathbf{H}_j \mathbf{H}_j^T \mathbf{U}^{-1}. \end{aligned} \quad (\text{C10})$$

By writing out the center column explicitly and using the commutation property of convolutions and then following much of the same line of arguments as above (Multiplication between a Toeplitz matrix and a vector corresponds to convolution, transposing the Toeplitz matrix corresponds to time reversing the corresponding signal, etc. These rules should, however, be used with some care), it can be shown that

$$\begin{aligned} R_{\tilde{s}}(\omega) &= \left( \frac{\sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2}{\sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2}} - 1 \right)^2 \sigma_s^2 \\ &\quad + \frac{\sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2}{\left( \sum_{j=1}^J \frac{1}{\sigma_j^2} |H_j(\omega)|^2 + \frac{1}{\sigma_s^2} \right)^2}. \end{aligned} \quad (\text{C11})$$

<sup>1</sup>G. Clark, D. Tilly, and W. Cook, "Ultrasonic signal/image restoration for quantitative NDE," *NDT International* **19**, 169–176 (1986).

<sup>2</sup>C. Chen and S. Sin, "High-resolution deconvolution techniques and their applications in ultrasonic NDE," *Int. J. Imaging Syst. Technol.* **1**, 223–242 (1989).

<sup>3</sup>S. Sin and C. Chen, "A comparison of deconvolution techniques for the ultrasonic nondestructive evaluation of materials," *IEEE Trans. Image Process.* **1**, 3–10 (1992).

<sup>4</sup>S. Neal, P. Speckman, and M. Enright, "Flaw signature estimation in ultrasonic nondestructive evaluation using the Wiener filter with limited prior information," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **40**, 347–353 (1993).

<sup>5</sup>L. Yaroslavsky and H. Caulfield, "Deconvolution of multiple images of the same object," *Appl. Opt.* **33**, 2157–2162 (1993).

<sup>6</sup>E. Miller and A. Willisky, "Wavelet transforms and multiscale estimation techniques for the solution of multisensor inverse problems," in *Proceedings of the SPIE, The International Society for Optical Engineering* (SPIE, Orlando, 1994), Vol. 2242, pp. 28–39.

<sup>7</sup>M. Fink and J.-F. Cardoso, "Diffraction effects in pulse-echo measurement," *IEEE Trans. Sonics Ultrason.* **SU31**, 313–329 (1984).

<sup>8</sup>D. Cassereau, D. Guyomar, and M. Fink, "Time deconvolution of diffraction effects—Application to calibration and prediction of transducer waveforms," *J. Acoust. Soc. Am.* **84**, 1073–1085 (1988).

<sup>9</sup>J. Jensen, "Field: A program for simulating ultrasound systems," *Med. Biol. Eng. Comput.* **34**, 351–353 (1996).

<sup>10</sup>J. Jensen and B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**, 262–267 (1992).

<sup>11</sup>P. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," *J. Acoust. Soc. Am.* **49**, 1627–1638 (1971).

<sup>12</sup>S. Neal and O. Thompson, "Utilization of prior information in ultrasonic NDE: An analysis of flaw scattering as a random variable," *J. Nondestruct. Eval.* **11**, 57–67 (1992).

<sup>13</sup>S. Neal and D. Thompson, "The measurement and analysis of acoustic noise as a random variable," in *Review of Progress in Quantitative NDE* (Plenum, New York, 1990), Vol. 9, pp. 625–632.

<sup>14</sup>A. Lh emery and R. Raillon, "Theoretical predictions and experimental measurements of echo responses from angled targets," in *Review of Progress in Quantitative NDE* (Plenum, New York, 1992), Vol. 11, pp. 81–88.

# Design of an optimal wave-vector filter for enhancing the resolution of reconstructed source field by near-field acoustical holography (NAH)

Bong-Ki Kim and Jeong-Guon Ih

*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 May 1999; revised 30 July 1999; accepted for publication 8 March 2000)

In near-field acoustical holography using the boundary element method, the reconstructed field often diverges due to the presence of small measurement errors. In order to handle this instability in the inverse problem, the reconstruction process should include some form of regularization for enhancing the resolution of source images. The usual method of regularization has been the truncation of wave vectors associated with small singular values, although the determination of an optimal truncation order is difficult. In this article, an iterative inverse solution technique is suggested in which the mean-square error prediction is used. A statistical estimation of the minimum mean-square error between measured pressures and the model solution is required for yielding the optimal number of iterations. The continuous curve of an optimal wave-vector filter is designed, for suppressing the high-order modes that can produce large reconstruction errors. Experimental results from a baffled radiator reveal that the reconstruction errors can be reduced by this form of regularization, by at least 48% compared to those without any regularization. In comparison to results using the optimal truncation method of regularization, the new scheme is shown to give further reductions of truncation error of between 7% and 39%, for the example in this article. © 2000 Acoustical Society of America. [S0001-4966(00)04006-6]

PACS numbers: 43.60.Sx, 43.50.Yw, 43.40.Rj [JCB]

## INTRODUCTION

As a method for source identification in vibroacoustic problems, near-field acoustical holography (NAH)<sup>1-3</sup> is widely used in many areas nowadays. In this technique, the measured field pressure data on the hologram plane is decomposed into the space and wave number domain by the spatial Fourier transform. The pressure on a target plane is reconstructed by the inverse spatial Fourier transform. The shape of the source surface or target plane should be regular. With concepts similar to the conventional NAH based on the spatial Fourier transform, boundary element method (BEM)-based acoustical holography has been suggested.<sup>4,5</sup> This technique has the following advantages compared with conventional NAH: the pressure need not be measured in separable coordinates, thus a reduced number of measurements with uneven spacing is possible; reflections from all directions can be considered; concave regions of the source can be reconstructed; and wrap-around error due to the finite aperture size is not involved. Due to these merits, the actual reconstruction of source fields using this technique has been performed by many researchers.<sup>3,5-9</sup>

The precision of reconstruction will depend on how closely to the source surface the measurement is made. However, in order to utilize the beneficial feature of uneven spacing of sensor positions, the measurement points should also be located independently from each other. For this purpose, the measuring points can be allocated so as to produce a minimum singularity in the transfer matrix by means of the effective independence (EfI) method.<sup>10</sup>

Additionally, in order to overcome the divergence phe-

nomenon due to nonpropagating wave components, proper filtering is needed during the inverse operation.<sup>5</sup> Because the higher-order modes in acoustical holography can be considered as nonpropagating wave components, low-pass filtering such as the Wiener filtering<sup>11-13</sup> can be taken in the optimal sense, and spatial regularization by using the singular value decomposition (SVD) is adopted<sup>5,14,15</sup> for taking account of the measurement noise. Another possible approach is regularization by using an iterative inverse solution.<sup>16,17</sup> This method in the image restoration area has been known not to require an inverse matrix and the precision of solutions can be enhanced by incorporating bounding conditions during iteration steps. In this method, the choice of iteration number is most important due to the trade-off between the variance error arising from measurement error and the bias error resulting from regularization. Unfortunately, no method for determining the optimal iteration number has been suggested.

In this paper, iterative inverse solutions are utilized for regularization in order to reduce the reconstruction error caused from measurement errors. In using the iterative inverse solution technique, the use of mean-square-error prediction is suggested. The optimal number of iterations is determined by estimating the minimum mean-square error (MSE) statistically between the measured pressure and the true model. An optimal wave-vector filter is designed for suppressing the higher-order modes producing large reconstruction errors. The filter coefficients take the values between zero and one, and a generalized function of the optimal wave-vector filter can be realized by curve fitting the obtained filter coefficients using an exponential function. As a demonstration example, a rectangular vibrating plate with a

rigid baffle is considered and experimental investigations are carried out to show the effectiveness of the proposed method.

## I. THEORETICAL BACKGROUND

### A. Basics of backward reconstruction using BEM

The matrix/vector equations can be obtained by the interpolation and the numerical integration of the well-known Kirchhoff–Helmholtz boundary integral equation as follows:<sup>5</sup>

$$\mathbf{D}_s \mathbf{p}_s = \mathbf{M}_s \mathbf{v}_s, \quad (1)$$

$$\mathbf{p}_f = \mathbf{D}_f \mathbf{p}_s + \mathbf{M}_f \mathbf{v}_s. \quad (2)$$

Here,  $\mathbf{p}_s$  and  $\mathbf{v}_s$  denote pressure and velocity on the source surface,  $\mathbf{D}_s$  and  $\mathbf{M}_s$  are dipole and monopole matrices on the surface,  $\mathbf{D}_f$  and  $\mathbf{M}_f$  are similar matrices corresponding to field pressures, and  $\mathbf{p}_f$  represents the field pressure vector. If Eq. (1) is substituted into Eq. (2), the field pressure can be related directly to the surface velocity as

$$\mathbf{p}_f = (\mathbf{D}_f \mathbf{D}_s^{-1} \mathbf{M}_s + \mathbf{M}_f) \mathbf{v}_s \equiv \mathbf{G} \mathbf{v}_s, \quad (3)$$

where  $\mathbf{G}$  is the vibro-acoustic transfer matrix between source and field. Equation (3) holds provided  $\mathbf{D}_s^{-1}$  and  $\mathbf{M}_s^{-1}$  exist. The surface velocity can then be calculated from the overdetermined least-squared solutions and singular value decomposition (SVD) of  $\mathbf{G}$  as

$$\hat{\mathbf{v}}_s = (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \mathbf{p}_f = \mathbf{W} \text{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_n)^{-1} \mathbf{U}^H \mathbf{p}_f. \quad (4)$$

Here,  $\hat{\mathbf{v}}_s$  is the estimated surface velocity vector, the superscript ‘‘ $H$ ’’ signifies the Hermitian operator, and the subscript ‘‘ $n$ ’’ is the rank of  $\mathbf{G}$ . The SVD for  $G$  is defined as

$$\mathbf{G} = \mathbf{U} \text{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_n) \mathbf{W}^H, \quad (5)$$

where

$$\Lambda_1 \geq \Lambda_2 \geq \dots \geq \Lambda_n \geq 0, \quad \mathbf{u}_i^H \mathbf{u}_j = \delta_{ij}, \quad \text{and} \quad \mathbf{w}_i^H \mathbf{w}_j = \delta_{ij}. \quad (6)$$

Here,  $\Lambda_i$  is the singular value,  $\mathbf{u}_i$  and  $\mathbf{w}_i$  are orthonormal vectors, and  $\delta_{ij}$  is the Kronecker delta.

### B. Reconstruction error and singularity factor

When the field pressures are contaminated with measurement noise, the measured field pressures can be rewritten as

$$\tilde{\mathbf{p}}_f = \mathbf{G} \mathbf{v}_s + \mathbf{n}. \quad (7)$$

Referring to Eq. (4), one can estimate the reconstructed velocity as

$$\hat{\mathbf{v}}_s = (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \tilde{\mathbf{p}}_f. \quad (8)$$

By substituting Eq. (7) into (8), the reconstruction error can be expressed as

$$\hat{\mathbf{v}}_s - \mathbf{v}_s = \mathbf{L} \mathbf{n}, \quad (9a)$$

where

$$\mathbf{L} \equiv (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H. \quad (9b)$$

If the measurement noises are uncorrelated Gaussian random with variance  $\sigma^2$  and zero mean, one can write the square of the reconstruction error as<sup>5</sup>

$$E[(\hat{\mathbf{v}}_s - \mathbf{v}_s)^H (\hat{\mathbf{v}}_s - \mathbf{v}_s)] = \sigma^2 S_F, \quad (10a)$$

where

$$S_F \equiv \text{tr}[\text{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_n)^{-2}]. \quad (10b)$$

Here  $S_F$  is called the singularity factor, which indicates the degree of singularity of a vibroacoustic transfer matrix.

### C. Selection of measurement points

The singularity of the transfer matrix is caused by the dependency between columns or rows of the transfer matrix. The contribution of measurement locations to the independence of the transfer matrix can be represented by<sup>5</sup>

$$\mathbf{E}_f = \text{diag}(\mathbf{U} \mathbf{U}^H). \quad (11)$$

The  $i$ th diagonal element of  $\mathbf{E}_f$  represents the contribution of the corresponding sensor location to the linear independence of the transfer matrix. The optimal measurement locations over the frequency range of interest can be determined by the averaged EfI value as

$$\mathbf{E}_f^{\text{avg}} = \frac{1}{(f_2 - f_1)} \text{diag} \left[ \int_{f_1}^{f_2} \mathbf{U} \mathbf{U}^H df \right]. \quad (12)$$

The optimal locations can be selected from a large set of candidate points by deleting the measurement points having the smallest EfI values. It should be noted that the selected positions are not globally optimal, but they are optimal for a given number of sensor positions.

### D. Regularization by iterative inverse solution and design of optimal wave-vector filter

The surface velocity in Eq. (3) can be estimated from the measured field pressure data by using the inverse iteration as<sup>12</sup>

$$\hat{\mathbf{v}}_s^{l+1} = \hat{\mathbf{v}}_s^l + \beta \mathbf{G}^H (\tilde{\mathbf{p}}_f - \mathbf{G} \hat{\mathbf{v}}_s^l) \quad (13a)$$

$$= \beta \mathbf{G}^H \tilde{\mathbf{p}}_f + (\mathbf{I}_n - \beta \mathbf{G}^H \mathbf{G}) \hat{\mathbf{v}}_s^l, \quad (13b)$$

where  $\hat{\mathbf{v}}_s^l$  denotes the estimated source velocity at the  $l$ th step,  $\mathbf{I}_n$  is the identity matrix with rank  $n$ , and  $\beta$  is a convergence parameter. Because Eq. (13b) can be considered as a geometric series, the necessary and sufficient condition for convergence as the iteration number  $l$  is increased is given by

$$|1 - \beta \Lambda_i^2| < 1, \quad \text{for } i = 1, 2, \dots, n. \quad (14)$$

If the inverse of  $\mathbf{G}$  is obtainable, the limit value of this geometric series for  $l \rightarrow \infty$  is equivalent to the pseudo-inverse solution as

$$\hat{\mathbf{v}}_s^\infty = (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \tilde{\mathbf{p}}_f. \quad (15)$$

From this, it can be stated that the direct pseudo-inverse solution can be obtained by increasing the number of iterations sufficiently. However, if measurement errors are involved, the reconstruction error does not converge to a mini-

imum value even if the number of iterations is infinitely increased, because the measurement errors can be amplified by nonradiating wave components during the backward reconstruction. Consequently, the iteration process should be terminated after some finite number of iterations chosen to achieve the minimum reconstruction error. It is particularly noteworthy that the high-order modes with small singular values possess a slower convergence rate than the lower modes, because the convergence rate depends on the geometric ratio of Eq. (13b). In other words, termination of the iteration process provides the low-pass filtering of the reconstruction field.

The direct implementation of the iterative filtering method can be written as

$$\hat{\mathbf{v}}_s^l = \beta \sum_{i=0}^l (\mathbf{I} - \beta \mathbf{G}^H \mathbf{G})^i \mathbf{G}^H \tilde{\mathbf{p}}_f \quad (16a)$$

$$= \beta \sum_{i=0}^l \mathbf{W} \Lambda (1 - \beta \Lambda^2)^i \mathbf{U}^H \tilde{\mathbf{p}}_f \quad (16b)$$

$$= \mathbf{W} \text{diag}[(1 - (1 - \beta \Lambda_1^2)^{l+1})/\Lambda_1, \dots, (1 - (1 - \beta \Lambda_n^2)^{l+1})/\Lambda_n] \mathbf{U}^H \tilde{\mathbf{p}}_f. \quad (16c)$$

By comparing Eq. (16c) with Eq. (4), the wave-vector filter matrix  $\mathbf{F}_l$  can be defined as

$$\hat{\mathbf{v}}_s^l = \mathbf{W} \mathbf{F}_l \text{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_n)^{-1} \mathbf{U}^H \tilde{\mathbf{p}}_f, \quad (17)$$

where

$$\mathbf{F}_l = \text{diag}(1 - (1 - \beta \Lambda_1^2)^{l+1}, \dots, 1 - (1 - \beta \Lambda_n^2)^{l+1}). \quad (18)$$

All the components of the filter matrix should be less than or equal to one. Divergence in the reconstruction field due to very small singular values can be suppressed by the filter.

The next objective is to find the optimal number of iterations,  $l$ , in Eq. (18). In the same fashion as in Ref. 1, the reproduced field pressure from the regularized solution can be expressed as

$$\hat{\mathbf{p}}_f = \mathbf{P}_l \tilde{\mathbf{p}}_f, \quad (19)$$

where

$$\mathbf{P}_l = \mathbf{U} \text{diag}(\Psi_1, \Psi_2, \dots, \Psi_n) \mathbf{U}^H$$

and

$$\Psi_i = 1 - (1 - \beta \Lambda_i^2)^{l+1}.$$

The mean-square error (MSE) of the regularized solutions can be derived as (see Appendix A)

$$\xi_l^2 = E[(\hat{\mathbf{p}}_f - \mathbf{G} \mathbf{v}_s)^H (\hat{\mathbf{p}}_f - \mathbf{G} \mathbf{v}_s)] \quad (21a)$$

$$= \sigma^2 \sum_{i=1}^n \Psi_i^2 + \mathbf{b}_l^H \mathbf{b}_l, \quad (21b)$$

where

$$\mathbf{b}_l = (\mathbf{P}_n - \mathbf{P}_l)(\mathbf{G} \mathbf{v}_s) \quad \text{and} \quad \mathbf{P}_n = \mathbf{U} \mathbf{U}^H. \quad (22)$$

The first term of Eq. (21b) represents the variance error and the latter is the bias error. If the number of iterations is increased, the variance error increases while the bias error

decreases. Consequently, there is a trade-off between the variance and bias errors, which can produce the minimum MSE. However, it is impossible to calculate the MSE directly because  $\mathbf{v}_s$  is unknown. Therefore, in order to obtain the MSE in terms of the measurement field, the bias error of the regularized model can be estimated as

$$\hat{\mathbf{b}}_l = (\mathbf{P}_n - \mathbf{P}_l) \tilde{\mathbf{p}}_f. \quad (23)$$

From this, the estimated MSE variance after the  $l$ th iteration can be derived as (see Appendix B)

$$\hat{\xi}_l^2 = \sigma^2 \left( 2 \sum_{i=1}^n \Psi_i - n \right) + \hat{\mathbf{b}}_l^H \hat{\mathbf{b}}_l, \quad (24)$$

where the condition of  $\xi_l^2 = E[\hat{\xi}_l^2]$  should be satisfied. From Eq. (24), one can determine the optimal number of iterations that yields the minimum MSE. The wave-vector filter in Eq. (18) can then be optimally designed with this number.

## E. Estimation of measurement error variance

In most realistic source identification problems, the measurement error variance  $\sigma^2$  is unknown *a priori*. However, the estimation scheme of measurement error variance is developed before the experiments, because it is essential to estimate the optimal regularization order and to design the optimal wave-vector filter. In order to estimate the measurement error variance, Eq. (7) can be modified as

$$\mathbf{n} = \tilde{\mathbf{p}}_f - \mathbf{G} \mathbf{v}_s = (\tilde{\mathbf{p}}_f - \mathbf{G} \hat{\mathbf{v}}_s) + \mathbf{G}(\hat{\mathbf{v}}_s - \mathbf{v}_s). \quad (25)$$

Substituting Eq. (9a) into Eq. (25) and using the SVD of transfer matrix  $\mathbf{G}$ , the following relationship can be derived:

$$(\mathbf{I}_m - \mathbf{U}_n \mathbf{U}_n^H) \mathbf{n} = \tilde{\mathbf{p}}_f - \mathbf{G} \hat{\mathbf{v}}_s. \quad (26)$$

Here,  $\mathbf{U}_n$  is composed of the first  $n$  columns of the matrix  $\mathbf{U}$ . The relationship of

$$\mathbf{U}_n \mathbf{U}_n^H + \mathbf{U}_{m-n} \mathbf{U}_{m-n}^H = \mathbf{I}_m \quad (27)$$

is satisfied because  $\mathbf{U}_n$  is the unitary matrix.  $\mathbf{U}_{m-n}$  are made up of the last  $(m-n)$  columns of the left singular matrix  $\mathbf{U}$ . Therefore, Eq. (26) can be rewritten as

$$(\mathbf{U}_{m-n} \mathbf{U}_{m-n}^H) \mathbf{n} = \tilde{\mathbf{p}}_f - \mathbf{G} \hat{\mathbf{v}}_s. \quad (28)$$

Taking the Frobenious norm and expectation of each side of Eq. (28), the following relationship can be obtained:

$$\text{tr}[(\mathbf{U}_{m-n} \mathbf{U}_{m-n}^H) E(\mathbf{n} \mathbf{n}^H)] = E[\|\tilde{\mathbf{p}}_f - \mathbf{G} \hat{\mathbf{v}}_s\|_F^2]. \quad (29)$$

If every measurement position has the same variance  $\sigma^2$ , the estimation of measurement error variance can be given by

$$\hat{\sigma}^2 = \frac{1}{m-n} \|\tilde{\mathbf{p}}_f - \mathbf{G} \hat{\mathbf{v}}_s\|_F^2, \quad (30)$$

where

$$\text{tr}(\mathbf{U}_{m-n} \mathbf{U}_{m-n}^H) = m-n, \quad \sigma^2 = E[\hat{\sigma}^2]. \quad (31)$$

Now, one can estimate the measurement error variance from the measured pressures and estimated surface velocity without regularization.

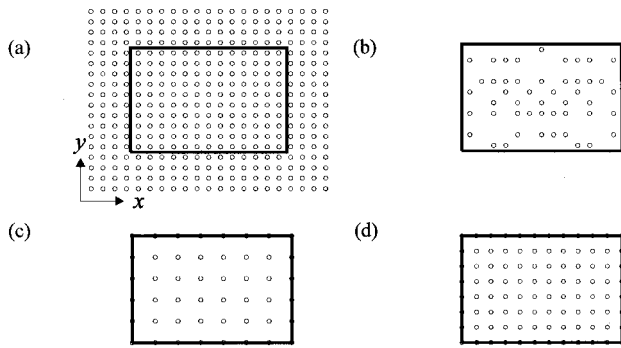


FIG. 1. Distribution of measurement points at the hologram plane: (a) initial 378 evenly distributed candidate points, (b) 48 measurement points selected by the EfI method, (c) evenly spaced 48 points, and (d) evenly spaced 96 points.

## II. APPLICATION EXAMPLE

### A. Test model and selection of measurement points

A vibrating rectangular flat plate of  $0.6 \times 0.4 \text{ m}^2$  with a rigid flat baffle is taken as an example. The boundary element model of the plate has 63 nodes and 24 quadratic isoparametric elements. The maximum characteristic length of an element is 0.2 m which limits the applicable high frequency to 572 Hz under the  $\lambda/3$  criterion (maximum modeling error  $\approx 1\%$ ). The number of initial candidate measurement points is chosen as 378 and they are uniformly distributed on a plane parallel to the baffle, located at 0.035 m from the plate as depicted in Fig. 1(a). As can be seen in Fig. 1(b), the measurement points are reduced to 48 points by applying the averaged EfI method within the frequency range of 50–500 Hz in which 30 uniformly distributed frequencies are selected for averaging. Additionally, two measurement-point sets composed of 48 and 96 points as shown in Fig. 1(c) and (d) are taken for investigating the validity of measuring points chosen by the EfI method. Measuring points in these sets are uniformly distributed on a measurement plane, parallel to the baffle, and located at 0.035 m from the plate. The purpose of choosing the measurement set in Fig. 1(d) is to show how many measurements points can be eliminated by the EfI process while maintaining the same singularity factor with the original evenly spaced measurement set. Calculated singularity factors are compared between three measurement sets as illustrated in Fig. 2. The singularity factor of the measurement set determined by the EfI method is much smaller than that of evenly distributed measurement points in the given frequency range, and this means that a small reconstruction error is expected. The high-frequency excitation yields a smaller singularity than excitations in low frequencies. In Fig. 2, one can see that the uneven measuring points chosen by the EfI method produce a far smaller singularity factor than the evenly spaced field point set with the same number of points. It is noted that the same level of singularity factor is obtainable when evenly spaced measurement points are chosen with a twofold increase in the number of points selected as compared to the EfI method.

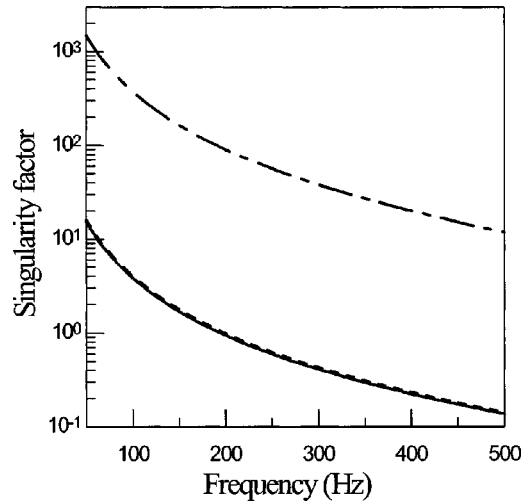


FIG. 2. Calculated singularity factor of different measurement sets in Fig. 1: —, unevenly located 48 points selected by EfI method from 378 candidate points; ---, evenly distributed 48 points; - · -, evenly distributed 96 points.

### B. Experimental setup and measurements

A rectangular plate having dimensions of  $0.6 \times 0.4 \text{ m}^2$ , with a rigid baffle, was made of a 1-mm-thick steel plate, and the boundaries were fixed to the baffle made of a 10-mm-thick square-holed aluminum plate. Outer regions of the aluminum plate were baffled by a 25-mm-thick plastic plate with dimensions of  $3.3 \times 2.36 \text{ m}^2$ . It was feared that the sound radiated from the bottom side of the panel might affect the upper side. For this purpose, a thick steel panel was placed at the test section and a loudspeaker transfer function test was performed before conducting the main experiment. The thin plate was excited by a shaker (B&K 4809) positioned at  $(x, y) = (0.15 \text{ m}, 0.267 \text{ m})$  and a band-limited random signal centered at 250 Hz was supplied to the shaker for the frequency range of 0–500 Hz. This signal was fed by a signal generator module in a 16-channel spectrum analyzer (HP 3566A) through a power amplifier (B&K 2706). The measurement plane was located at 0.035 m from the vibrating source plane. Eight microphones (B&K 4130) of  $\frac{1}{2}$ -in. diameter were used for simultaneous data acquisition and scanned over 48 measurement positions. Because the whole sound field could be considered as being perfectly coherent with the excitation signal, field pressures were referenced to only one fixed reference microphone. The reference microphone was located above a corner of the vibrating plate and 0.04 m from the plate.

Because the precise calibration of microphones is very important in the measurement, especially in this backward reconstruction, eight microphones were simultaneously calibrated by using the in-duct method. By generating the reference signal from the opposite end of the duct, for the frequency range of interest, the frequency response functions between the reference microphone and others were measured and investigated. Phase mismatches were well within 1 degree for the frequency range of interest and the calibration data had been used consistently to correct the magnitude and phase mismatch during data acquisition. Insulation of airborne sound was carefully made to cut off the sound origi-

TABLE I. Reconstruction error of the surface velocity recovered from the measured pressure data at 48 near-field points. Here, set I is composed of the points selected by the EfI method from 378 candidate points and set II contains the evenly distributed points. Regularization was conducted by using the set I data. RTM means the regularization using the wave-vector truncation and RIM denotes that using the inverse iteration method.

Frequency (Hz)	Noise variance		S/N ratio		$e(\text{w/o regularization})$ (%)		$e(\text{w/regularization})$ (%)	
	Set I	Set II	Set I	Set II	Set I	Set II	RTM	RIM
111.5	0.8133	0.1352	51.4	59.2	72.9	1094.7	32.0	24.8
270.3	1.0516	0.4418	37.8	46.8	121.7	4453.1	78.2	39.2

nating from the vibration exciter and via direct and flanking paths. For this purpose, a steel box packed with sound-absorbing material was employed in which the shaker was supported by a 25-mm-thick rubber pad lying on a 75-mm-thick sand layer at the bottom. Tests had been carried out in a full anechoic chamber.

In order to compare the reconstructed vibration field with the actual one, frequency response functions (FRFs) at 35 points on the vibrating plate were independently measured, referenced to the force signal at the excitation point. Miniature accelerometers (PCB 353B16) were mounted on the plate. A very subtle change of dynamic characteristics of the plate was observed due to loading, but the change was not a considerable one. The coherence between vibration and acoustical signals was very high and all the peak frequencies having large vibration amplitudes also existed in the acoustic field. From many peaks, two resonant peaks at 111.5 and 270.3 Hz were selected for demonstrating the reconstruction of the vibroacoustic field of the vibrating plate. Both the vibration mode shape and the signal-to-noise ratio were different for each resonance peak.

Two sets of 48 field pressure data at the hologram plane were measured by using the 16-channel spectrum analyzer (HP 3566A). Measuring positions of set I were determined from the EfI results as shown in Fig. 1(b) and those of set II were evenly spaced, as can be seen in Fig. 1(c). By using Eq. (30), the error variance of measured pressures was estimated as shown in Table I.

### C. Reconstruction of source velocity

The vibration fields of the plate are reconstructed first without any regularization processing. The measured vibrating shapes of the plate are as shown in Fig. 3, for the two selected resonance frequencies. One can clearly observe a (2,2) bending mode shape of the plate at 111.5 Hz, whereas the measured vibration pattern at 270.3 Hz shows one of high-order bending mode shapes.

Figures 4 and 5 show the reconstructed surface velocities using the measurement sets I and II, respectively, without conducting any regularization. As expected, the restoration process based on the measurements using set I, selected by the EfI method, produces more precise vibration fields than set II.

Although the result given in Fig. 5 contains basic features of the actual vibrating motion, the precision is not satisfactory yet and enhancement of the resolution is required for the reconstruction. The MSE can be estimated from Eq. (24) by using the estimated error variance and measured pressures, and the reconstruction error can be calculated by comparing the restored data with the actual measured velocity field of the vibrating surface. The reconstruction error in percentage terms is given by

$$e = \frac{\|\mathbf{v}_s - \hat{\mathbf{v}}_s\|}{\|\mathbf{v}_s\|} \times 100(\%). \quad (32)$$

Figure 6 shows the calculated results with an increasing

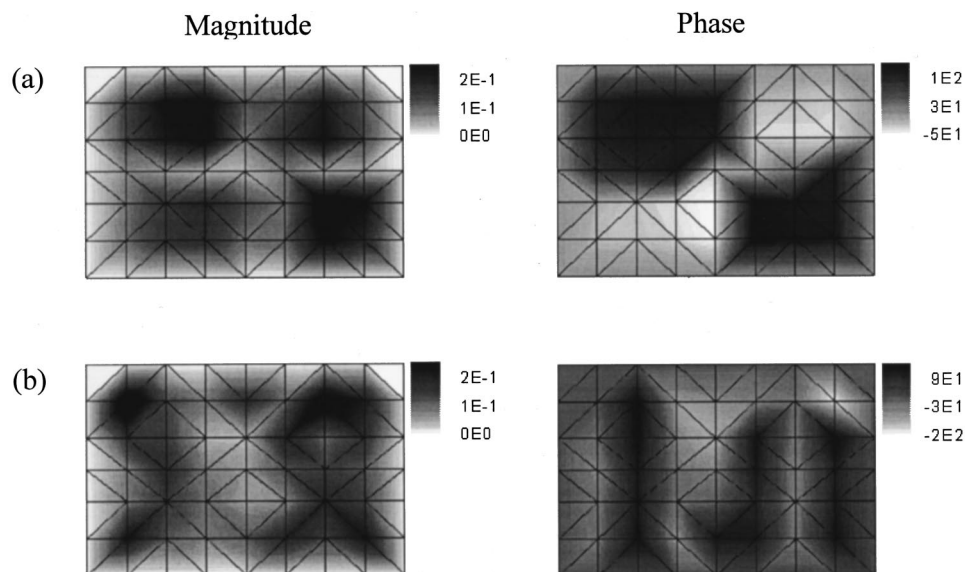


FIG. 3. Distribution of measured actual surface normal velocity: (a) 111.5 Hz and (b) 270.3 Hz.



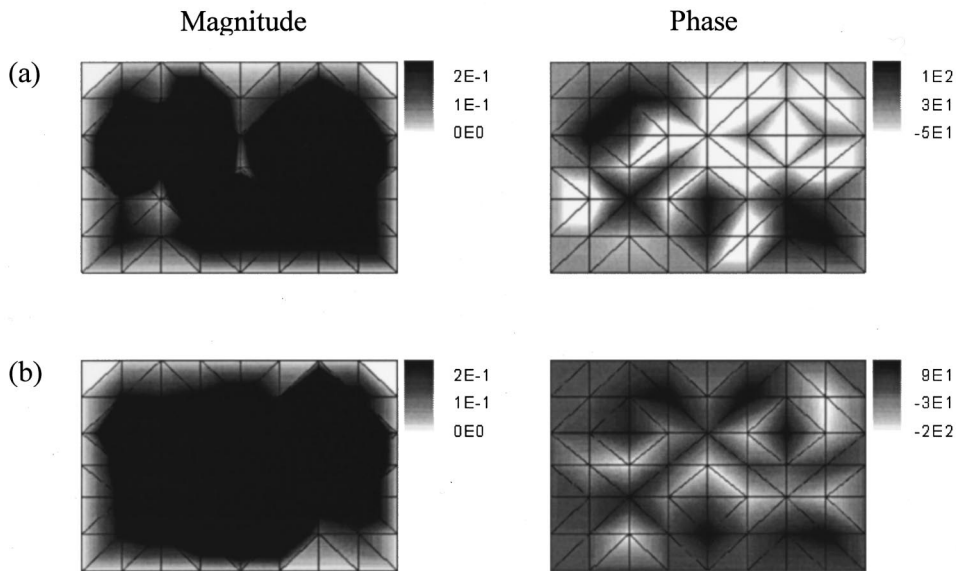


FIG. 4. Surface velocity reconstructed from pressure measurements at evenly spaced 48 field points without regularization: (a) 111.5 Hz ( $e=1094.7\%$ ) and (b) 270.3 Hz ( $e=4453.1\%$ ).

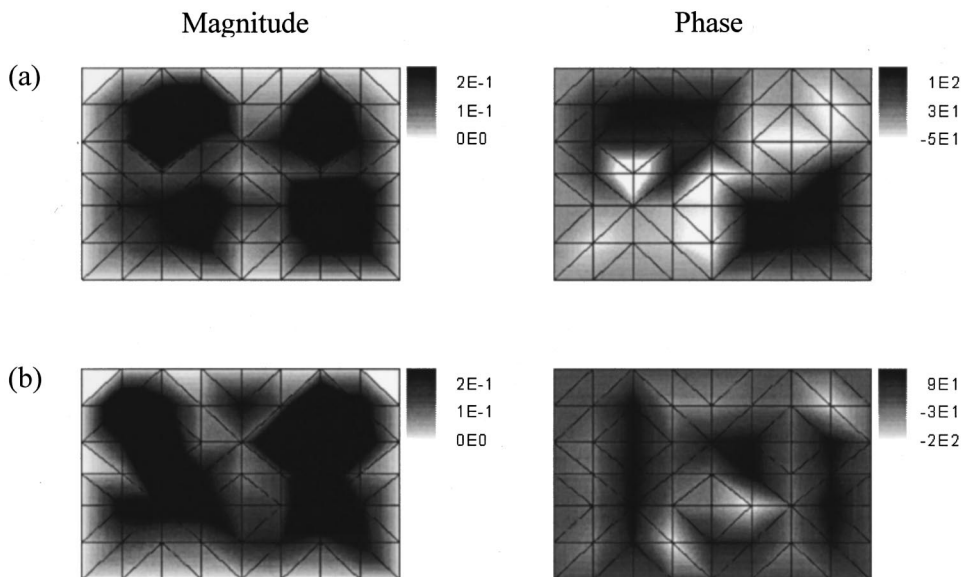


FIG. 5. Surface velocity reconstructed from pressure measurements at 48 field points selected by Efi method without regularization: (a) 111.5 Hz ( $e=72.9\%$ ) and (b) 270.3 Hz ( $e=121.7\%$ ).

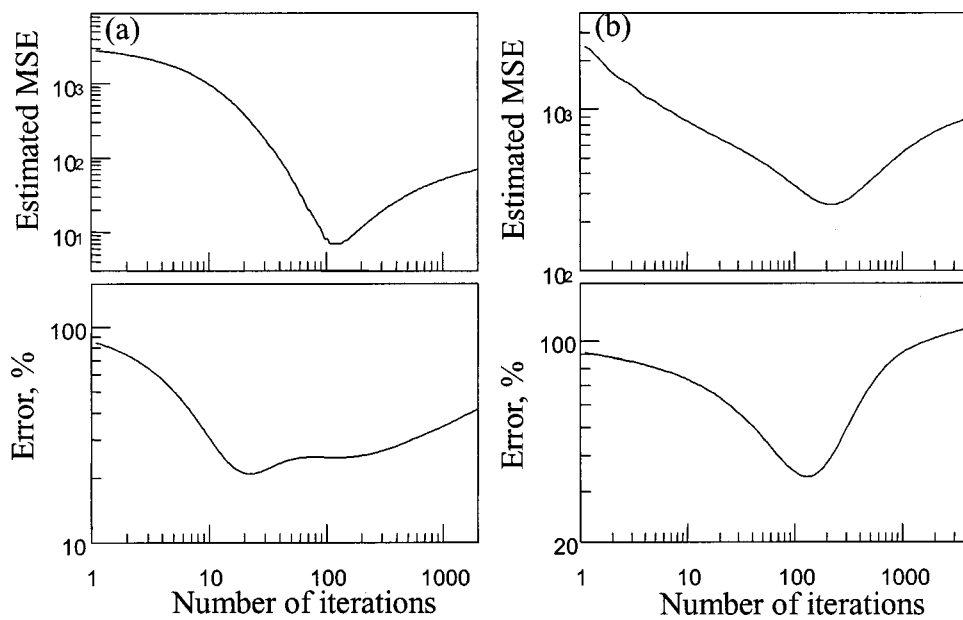


FIG. 6. Estimated mean-squared and reconstruction errors with increasing the number of iterations: (a) 111.5 Hz and (b) 270.3 Hz.

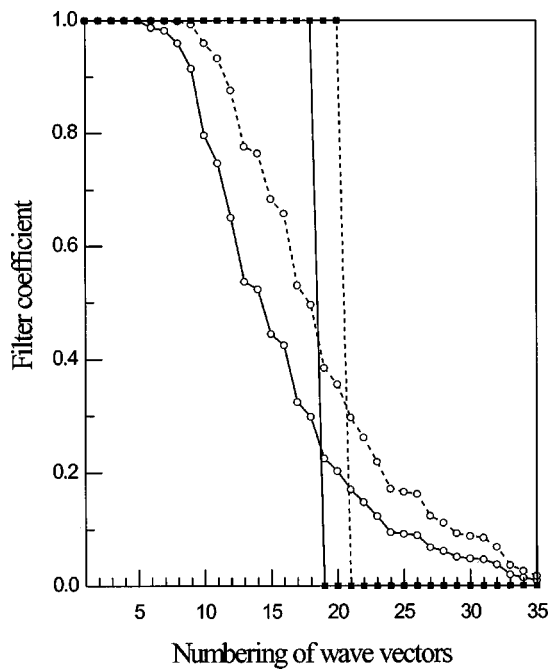


FIG. 7. Shapes of optimal wave-vector filters for the regularization: boxed line with box, truncation method at 111.5 Hz; dotted line with box, truncation method at 270.3 Hz; line with circle, iteration method at 111.5 Hz; dotted line with circle, iteration method at 270.3 Hz.

number of iterations. By comparing the two curves of estimated MSE and reconstruction error, one can see that the number of iterations for attaining the minimum MSE differs slightly from that for the minimum reconstruction error. However, the reconstruction errors are greatly reduced near the optimal iteration number anyway. The reconstruction errors are described in detail in Table I.

With the chosen optimal number of iterations, one can design the optimal wave-vector filter for the regularization. The filter coefficients for each resonance frequency are plotted in Fig. 7 as a function of wave-vector number. In this figure, the filter coefficients for regularization by the trunca-

tion method<sup>5,18</sup> are compared with those for the iterative method. One can be assured from Fig. 7 that the high-frequency excitation includes more high-order mode information than the low-frequency one. This is because the high-frequency excitation generates more high-order radiation modes.<sup>19</sup> In other words, more information on high-order modes is needed for the high-frequency excitations during the backward reconstruction. Figures 8 and 9 show the regularized results of reconstructed velocity distributions. In comparison with Fig. 5, one can observe that the resolution of reconstructed velocities is dramatically enhanced by the optimal regularization process. It is noted that smaller reconstruction error can be achieved by using the iterative regularization technique than the truncation method. Although there is another wave-vector filtering method, by using the estimated regularization parameter,<sup>5</sup> it is omitted here to avoid unnecessary duplication. It is observed that the present iterative method is slightly better than the regularization parameter method in the example treated here. However, it is thought that many further various application examples are required in order to determine which method is better.

#### D. Continuous wave-vector filter

In order to obtain the generalized function of the wave-vector filter, the filter coefficients from previously obtained optimal wave-vectors are curve fitted using the following exponential function:

$$F(n_k) = \exp[-\gamma(n_k - 1)^\tau]. \quad (33)$$

Here,  $n_k$  is the order of the wave vector. The values of  $\gamma$  and  $\tau$  should be appropriately determined to minimize the squared-error between the filter coefficient and the fitting function. By simulating several examples, it has been found that the exponent  $\tau$  of Eq. (33) is proportional to the cube of the wave-vector order in the iterative regularization method. The curve fitting result is shown in Fig. 10, where  $\gamma$  is selected as  $2.5881e-3$  when  $f=111.5$  Hz and  $1.3635e-3$  when  $f=270.3$  Hz. From the results,  $\gamma$  is seen to decrease with

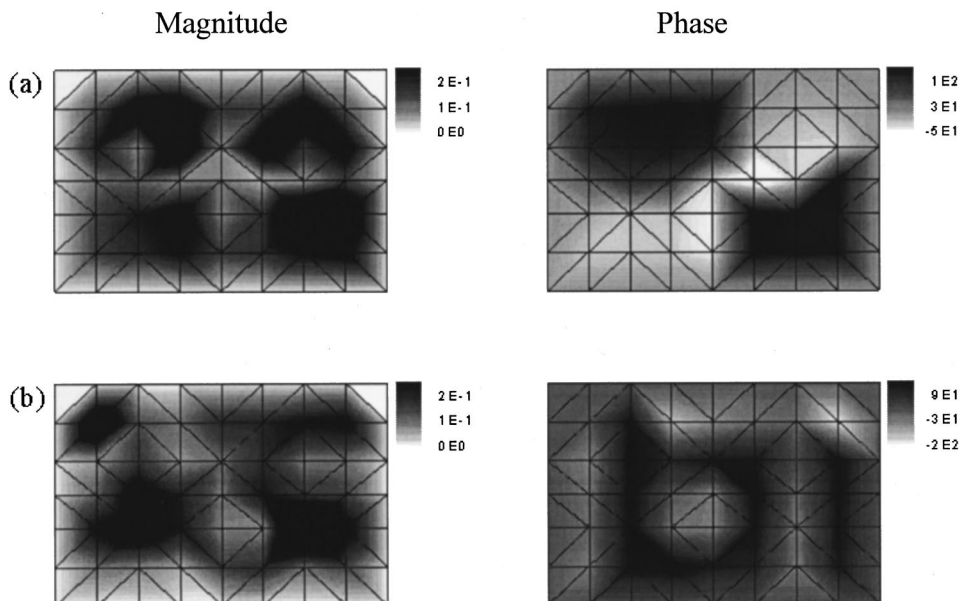


FIG. 8. Resolution enhanced surface velocity by the regularization using the wave-vector truncation at (a) 111.5 Hz ( $e=32.0\%$ ) and (b) 270.3 Hz ( $e=78.2\%$ ).

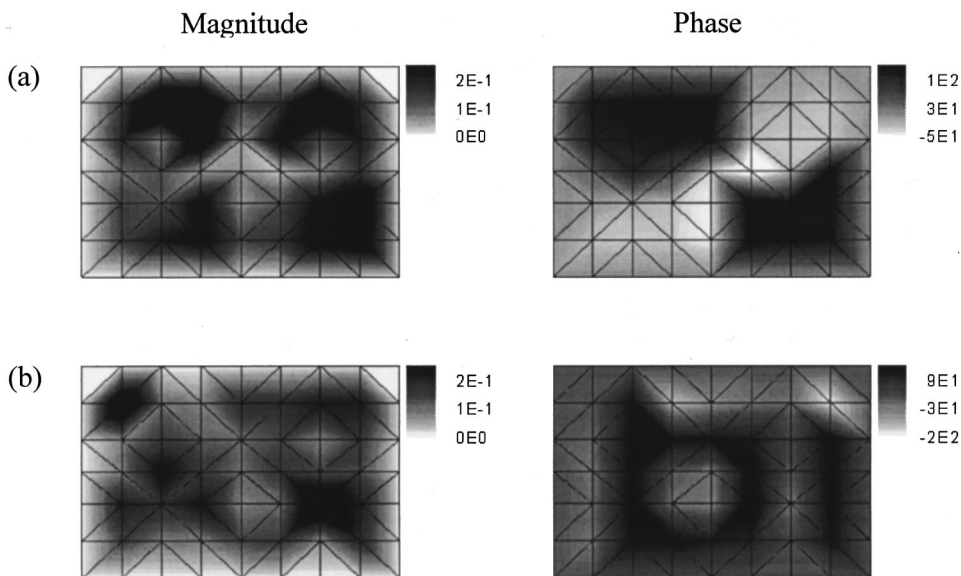


FIG. 9. Resolution enhanced surface velocity by the iterative regularization at (a) 111.5 Hz ( $e=24.8\%$ ) and (b) 270.3 Hz ( $e=39.2\%$ ).

increase of the frequency because the high-frequency excitation should contain richer information on higher-order modes than the low-frequency one, as explained before.

### III. CONCLUSIONS

In this paper, the resolution of reconstruction fields in NAH was enhanced by the establishment of a sensor-positioning method and a design method for optimal wave-vector filtering. The sensor positioning was done by using the effective independence (Efi) method, to ensure independence between sensor positions. Simulations with baffled rectangular radiator model revealed that the uneven measuring points chosen by the Efi method produced a far smaller singularity factor than the evenly spaced field point set with the same number of points. In the test example, the same level of singularity factor was obtained when evenly spaced measurement points were chosen with a twofold increase in the number of points selected as compared to the Efi method. It was found that termination of the iterative inverse solution at a selected stage yielded the minimum mean-square error between the model and the measured pressures. Through an example application of a vibrating rectangular plate, it was found that the estimated optimal number of iterations agreed

well with that producing the minimum reconstruction error. The overall shape of the resultant optimal wave-vector filter had an exponentially decreasing trend with increasing order of the filter, and the high-frequency excitations included more high-order mode information than the low-frequency one. In addition, optimal filter coefficients from the obtained wave vectors by iterations were curve fitted using an exponential function and, then, a generalized function of optimal wave-vector filter was realized. The reconstruction results, after regularization using the optimal wave-vector filter, showed very fine precision when collated with the measured velocity distribution of the source surface. Experimental results from a baffled radiator reveal that the reconstruction errors can be greatly reduced by this form of regularization, by at least 48% compared to those without any regularization. In comparison to results using the optimal truncation method of regularization, the new scheme is shown to give further reductions of truncation error of between 7% and 39%, for the example in this paper.

In this paper, the concept of inverse iteration technique was utilized without employing any bounding conditions during iterations for reducing the computation time. Therefore, there remains good opportunity to improve the technique by incorporating the bounding conditions such as the source information. Although only a simple test case was treated in this paper, the proposed sensor positioning and regularization methodology for reconstructing the source field can be applied to the source identification and visualization problems of many practical noise sources. In particular, the suggested optimal wave-vector filter will be a useful tool for enhancing the resolution of the reconstruction field of any kind of ill-posed acoustical problem.

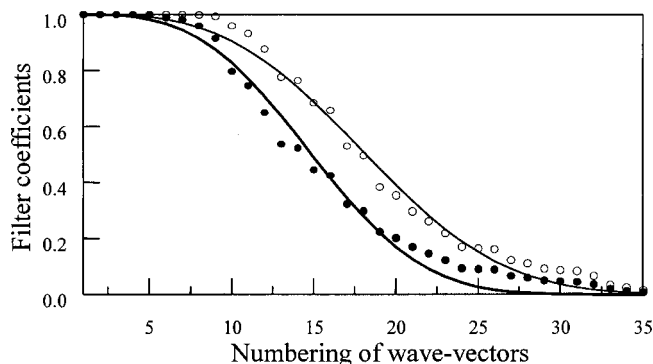


FIG. 10. Exponential curve fitting of wave-vector filter coefficients (marked by symbols) in Fig. 7 obtained by using the iterative regularization method: —, 111.5 Hz; - - -, 270.3 Hz.

### ACKNOWLEDGMENTS

This article was completed while the second author was visiting the Department of Mathematical Sciences, Loughborough University in U.K. He is grateful to the faculty and staff of Mathematical Sciences for their hospitality, and ex-

tends special thanks to Dr. Keith Peat, and to the Leverhulme Trust and the LG Younam Foundation for providing financial support.

## APPENDIX A: DERIVATION OF THE MEAN-SQUARE ERROR

The MSE between the true model and the reproduced field pressure after regularization can be written as

$$\begin{aligned}\xi_l^2 &= E[(\hat{\mathbf{p}}_f - \mathbf{G}\mathbf{v}_s)^H(\hat{\mathbf{p}}_f - \mathbf{G}\mathbf{v}_s)] \\ &= E[(\mathbf{P}_l\tilde{\mathbf{p}}_f - E[\mathbf{P}_l\tilde{\mathbf{p}}_f])^H(\mathbf{P}_l\tilde{\mathbf{p}}_f - E[\mathbf{P}_l\tilde{\mathbf{p}}_f])] \\ &\quad + (E[\mathbf{P}_l\tilde{\mathbf{p}}_f] - \mathbf{G}\mathbf{v}_s)^H(E[\mathbf{P}_l\tilde{\mathbf{p}}_f] - \mathbf{G}\mathbf{v}_s).\end{aligned}\quad (\text{A1})$$

Using the property of  $E[\mathbf{P}_l\tilde{\mathbf{p}}_f] = \mathbf{P}_l\mathbf{G}\mathbf{v}_s$ , the MSE can be derived as

$$\begin{aligned}\xi_l^2 &= E[(\mathbf{P}_l(\tilde{\mathbf{p}}_f - \mathbf{G}\mathbf{v}_s))^H(\mathbf{P}_l(\tilde{\mathbf{p}}_f - \mathbf{G}\mathbf{v}_s))] \\ &\quad + (\mathbf{P}_l\mathbf{G}\mathbf{v}_s - \mathbf{G}\mathbf{v}_s)^H(\mathbf{P}_l\mathbf{G}\mathbf{v}_s - \mathbf{G}\mathbf{v}_s)\end{aligned}\quad (\text{A2})$$

$$\begin{aligned}&= \sigma^2 \text{tr}[\mathbf{U} \text{diag}(\Psi_1, \Psi_2, \dots, \Psi_N)^2 \mathbf{U}^H] \\ &\quad + (\mathbf{G}\mathbf{v}_s)^H(\mathbf{P}_n - \mathbf{P}_l)^2(\mathbf{G}\mathbf{v}_s)\end{aligned}\quad (\text{A3})$$

$$= \sigma^2 \sum_{i=1}^n \Psi_i^2 + (\mathbf{G}\mathbf{v}_s)^H(\mathbf{P}_n - \mathbf{P}_l)^2(\mathbf{G}\mathbf{v}_s), \quad (\text{A4})$$

in which  $\mathbf{P}_n\mathbf{G}\mathbf{v}_s = \mathbf{G}\mathbf{v}_s$ .

## APPENDIX B: DERIVATION OF THE ESTIMATED MEAN-SQUARE ERROR

The estimated bias error of the regularized model can be given by

$$\hat{\mathbf{b}}_l = (\mathbf{P}_n - \mathbf{P}_l)\tilde{\mathbf{p}}_f. \quad (\text{B1})$$

From this, one can obtain the MSE of the bias error as follows:

$$\begin{aligned}E[(\mathbf{b}_l - \hat{\mathbf{b}}_l)^H(\mathbf{b}_l - \hat{\mathbf{b}}_l)] &= E[\text{tr}[(\mathbf{b}_l - \hat{\mathbf{b}}_l)(\mathbf{b}_l - \hat{\mathbf{b}}_l)^H]] \\ &= \text{tr}[\sigma^2(\mathbf{P}_n - \mathbf{P}_l)^2] \\ &= \sigma^2 \left( \sum_{i=1}^n \Psi_i^2 - 2 \sum_{i=1}^n \Psi_i + n \right).\end{aligned}\quad (\text{B3})$$

Because

$$E[(\mathbf{b}_l - \hat{\mathbf{b}}_l)^H(\mathbf{b}_l - \hat{\mathbf{b}}_l)] = E[\hat{\mathbf{b}}_l^H \hat{\mathbf{b}}_l] - \mathbf{b}_l^H \mathbf{b}_l, \quad (\text{B4})$$

by substituting Eq. (B3) into (B4), the MSE in Eq. (21b) can be rewritten as

$$\xi_l^2 = \sigma^2 \left( 2 \sum_{i=1}^n \Psi_i - n \right) + E[\hat{\mathbf{b}}_l^H \hat{\mathbf{b}}_l]. \quad (\text{B5})$$

Consequently, the estimated MSE under the condition of  $\xi_l^2 = E[\hat{\xi}_l^2]$  is also the sum of variance and bias errors due to the regularization and is given by

$$\hat{\xi}_l^2 = \sigma^2 \left( 2 \sum_{i=1}^n \Psi_i - n \right) + \hat{\mathbf{b}}_l^H \hat{\mathbf{b}}_l. \quad (\text{B6})$$

- <sup>1</sup>E. G. Williams and J. D. Maynard, "Holographic imaging without the wavelength resolution limit," *Phys. Rev. Lett.* **45**, 554–557 (1980).
- <sup>2</sup>J. D. Maynard, E. G. Williams, and Y. Lee, "Near-field acoustic holography: I. Theory of generalized holography and the development of NAH," *J. Acoust. Soc. Am.* **78**, 1395–1413 (1985).
- <sup>3</sup>E. G. Williams, B. H. Houston, and J. A. Bucaro, "Broadband nearfield acoustical holography for vibrating cylinders," *J. Acoust. Soc. Am.* **86**, 674–679 (1989).
- <sup>4</sup>M. R. Bai, "Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries," *J. Acoust. Soc. Am.* **92**, 533–548 (1992).
- <sup>5</sup>B.-K. Kim and J.-G. Ih, "On the reconstruction of vibro-acoustic field over the surface enclosing an interior space using the boundary element method," *J. Acoust. Soc. Am.* **100**, 3003–3016 (1996).
- <sup>6</sup>W. A. Veronesi and J. D. Maynard, "Digital holographic reconstruction of source with arbitrarily shaped surfaces," *J. Acoust. Soc. Am.* **85**, 588–598 (1989).
- <sup>7</sup>K. Gardner and R. J. Bernhard, "A noise source identification technique using an inverse Helmholtz integral equation method," *ASME J. Vib., Acoust., Stress, Reliab. Des.* **110**, 84–90 (1988).
- <sup>8</sup>J.-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," in *Proc. 1998 Int. Compressor Eng. Conference, Purdue, IN* (1998), pp. 525–529.
- <sup>9</sup>E. G. Williams, "An overview of the inverse problem for sound reconstruction in interior spaces," *J. Acoust. Soc. Am.* **105**, 968(A) (1999).
- <sup>10</sup>D. C. Kammer, "Effect of modal error on sensor placement for on-orbit modal identification of large space structures," *J. Guid. Control Dyn.* **15**, 334–341 (1992).
- <sup>11</sup>M. R. Bai, "Acoustical source characterization by using recursive Wiener filtering," *J. Acoust. Soc. Am.* **97**, 2657–2663 (1995).
- <sup>12</sup>H. Fleischer and V. Axelrad, "Restoring an acoustic source from pressure data using Wiener filtering," *Acustica* **60**, 172–175 (1986).
- <sup>13</sup>H. C. Andrews and B. R. Hunt, *Digital Image Restoration* (Prentice-Hall, Englewood Cliffs, NJ, 1977).
- <sup>14</sup>E. G. Williams, B. H. Houston, and P. C. Herdic, "Reconstruction of the surface velocity and interior acoustic intensity from an aircraft fuselage using nearfield acoustical holography," *Proc. Noise-Con 96, Seattle, WA* (1996), pp. 193–198.
- <sup>15</sup>H. Lee and D. P. Sullian, "Fundamental limitation of resolution enhancement by wave-field extrapolation," *J. Acoust. Soc. Am.* **84**, 611–617 (1988).
- <sup>16</sup>G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-37**, 2024–2036 (1989).
- <sup>17</sup>J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," *Proc. IEEE* **78**, 856–883 (1990).
- <sup>18</sup>W. A. Veronesi and J. D. Maynard, "Near-field acoustic holography (NAH). II. Holographic reconstruction algorithms and computer implementation," *J. Acoust. Soc. Am.* **81**, 1307–1322 (1987).
- <sup>19</sup>J.-G. Ih, J.-Y. Kim, and B.-K. Kim, "Acoustic field prediction of complex radiators by using near-field measurements and BEM," in *Proc. 11th IMAC, Orlando, FL* (1993), pp. 1396–1401.

# Group delays of distortion product otoacoustic emissions: Relating delays measured with $f_1$ - and $f_2$ -sweep paradigms

Vera F. Prijs, Sandra Schneider, and Ruurd Schoonhoven

Leiden University Medical Center, Department of ENT/Audiology, P.O. Box 9600, 2300 RC Leiden, The Netherlands

(Received 29 November 1999; revised 25 February 2000; accepted 8 March 2000)

A theoretical analysis is presented of group delays of distortion product otoacoustic emissions (DPOAEs) measured with the phase-gradient method. The aim of the analysis is to clarify the differences in group delays  $D_1$  and  $D_2$ , obtained using the  $f_1$ - and the  $f_2$ -sweep paradigms, respectively, and the dependence of group delays on the order of the DPOAE. Two models are considered, the place-fixed and the wave-fixed models. While in the former model the generation place is assumed to be invariant with both  $f_1$ - and  $f_2$ -sweeps, in the latter model the shift of generation place is fully accounted for. By making a simple local approximation of the cochlear scale invariance, a mathematical conversion from phase-place to phase-frequency gradients is incorporated in the wave-fixed model. Under the assumption that the DPOAE (as recorded at the best  $f_2/f_1$  ratio) is dominated by the contribution from the generation site and not by, e.g., reflection components, the analysis leads to simple expressions for the ratio and difference between  $D_1$  and  $D_2$ . Validation of the models against experimental data indicates that lower sideband DPOAEs ( $2f_1 - f_2$ ,  $3f_1 - 2f_2$ ,  $4f_1 - 3f_2$ ) are most consistent with the wave-fixed model. Upper sideband components ( $2f_2 - f_1$ ), in contrast, are not properly described by either the place-fixed or the wave-fixed model, independent whether DPOAE generation is assumed to originate at the  $f_2$  or at the more basally located  $f_{dp}$  characteristic place. © 2000 Acoustical Society of America.

[S0001-4966(00)04206-5]

PACS numbers: 43.64.Jb, 43.64.Kc [BLM]

## LIST OF SYMBOLS

$f_i$	frequency of component $i$ ; $i = 1, 2, dp$
$f_1, f_2$	primary frequencies; $f_2 > f_1$
$f_{dp}$	distortion product frequency, $f_{dp} = (n+1)f_1 - nf_2$ ; $n$ integer, where $n \geq 1$ for the lower sideband DPOAEs: $f_{dp} < f_1, f_2$ and $n \leq -2$ for the upper sideband DPOAEs: $f_{dp} > f_1, f_2$
$x$	place along the basilar membrane
$X_i$	characteristic place of component $i$

$x_g$	place of maximal DPOAE generation
$L_i$	level of component $i$ in the ear canal
$\varphi(x; \omega_i)$	phase at place $x$ of component $i$ with $\omega_i = 2\pi f_i$
$\tilde{\varphi}(x; \omega)$	forward phase change
$\bar{\varphi}(x; \omega)$	retrograde phase change
$D_i$	DPOAE group delay for $f_i$ -sweep
$\vec{D}$	forward delay
$\bar{D}$	backward delay

## INTRODUCTION

Distortion product otoacoustic emissions (DPOAEs) are generated in the cochlea when the ear is stimulated with two primaries having slightly different frequencies  $f_1$  and  $f_2$  ( $f_1 < f_2$ ). The group delay of a DPOAE can be determined using the phase-gradient method, i.e., from the slope of the DPOAE phase-frequency function. Thereto, variation of DPOAE frequency is accomplished by keeping one of the primaries fixed while changing the other (Kimberley *et al.*, 1993). In the  $f_1$ -sweep paradigm  $f_1$  is varied at fixed values of  $f_2$ , while in the  $f_2$ -sweep paradigm  $f_2$  is varied at fixed  $f_1$ .

For the lower sideband (LSB) DPOAEs (with  $f_{dp}$  smaller than the primaries  $f_1$  and  $f_2$ , i.e.,  $2f_1 - f_2$ ,  $3f_1 - 2f_2$ , and  $4f_1 - 3f_2$ ), these two methods give different group delay values. The group delays determined with  $f_2$ -sweep appear to be systematically larger (O'Mahoney and Kemp, 1995; Moulin and Kemp, 1996a, b; Whitehead *et al.*,

1996; Bowman *et al.*, 1997; Schneider *et al.*, 1999). However, for the upper sideband (USB) DPOAE  $2f_2 - f_1$  the delays measured with the two methods are basically equal (Moulin and Kemp, 1996b; Schneider *et al.*, 1999). Additionally,  $f_1$ -sweep group delays do not depend on the order of the DPOAE, while the  $f_2$ -sweep group delays show a small but significant dependence on order for the lower sidebands. Furthermore, in contrast with  $f_1$ -sweep delays,  $f_2$ -sweep delays demonstrate a large difference between lower and upper sidebands (Wable *et al.*, 1996; Moulin and Kemp, 1996b; Whitehead *et al.*, 1996; Schneider *et al.*, 1999).

The origin of the differences between upper and lower sideband DPOAEs is sought in different generation sites. Suppression experiments, onset latencies and amplitude correlations have shown that the main source of the lower sideband DPOAEs is close to the  $f_2$  characteristic place  $X_2$ , while the upper sideband DPOAE  $2f_2 - f_1$  originates from a

more basal region, close to  $X_{dp}$  (Brown and Kemp, 1984; Kummer *et al.*, 1995; Martin *et al.* 1987, 1998).

With regard to LSB DPOAEs the notion is now accepted that two sources contribute to the acoustic ear canal response. The region of nonlinear interaction near  $X_2$  generates distortion products  $f_{dp}$ . Propagation takes place in two directions, basally towards the ear canal, and apically towards the DPOAE characteristic place. There, the apically traveling component is partially reflected by the mechanism of linear coherent reflection, resulting in a backward traveling wave. In the ear canal, the DPOAE is a combination of a component generated at  $X_2$ , and a component reflected at  $X_{dp}$  (Shera and Guinan, 1999; Talmadge *et al.*, 1998). Interference of both components is held responsible for the fine structure in DPOAE amplitude and the corresponding irregularities in the phase-frequency profiles (Brown *et al.*, 1996; Heitmann *et al.*, 1998; Talmadge *et al.*, 1999). The relative strength of the two contributions depends, among others, on the primary levels and the primary frequency regions (Fahey and Allen, 1997; Knight and Kemp, 2000). Near the  $f_2/f_1$  ratio producing maximal DPOAE amplitudes, where the aforementioned DPOAE group delay studies were performed, the DPOAE is thought to be dominated by the contribution from the generation place  $X_2$  (Knight and Kemp, 2000). The present study concentrates on the same condition and therefore focuses on an analysis of the consequences of sweeping either  $f_1$  or  $f_2$  for the phase variation of the DPOAE generated at  $X_2$ .

The difference between  $f_1$ - and  $f_2$ -sweep group delays for the LSB components is thought to originate from a difference in shift of the generation place with swept  $f_1$  or  $f_2$ . Two models have been proposed in this context: the place- and the wave-fixed models (Kemp, 1986; O'Mahoney and Kemp, 1995; Moulin and Kemp, 1996b). In the place-fixed model, the generation place is thought to be unchanged during both the  $f_1$ - and the  $f_2$ -sweep paradigms. The generation place is fixed to the cochlear partition, close to  $X_2$ . It is tentatively associated with an anatomical irregularity in the cochlea. The wave-fixed model assumes a generation place which is fixed to the maximum in the traveling wave envelope of  $f_2$ , i.e., to  $X_2$ . As a consequence, in this model the generation place shifts along with a swept  $f_2$ . Based on general assumptions, Moulin and Kemp (1996b) were able to give several qualitative and some quantitative predictions regarding the relations between  $f_1$ -sweep and  $f_2$ -sweep group delays and their dependence on the order of the DPOAE. A full quantitative description of the wave-fixed model was not feasible due to the lack of an appropriate description of the spatial phase changes of the  $f_1$ -wave read out by a moving  $X_2$  in the  $f_2$ -sweep paradigm. The intermediate model, introduced by Moulin and Kemp (1996b), is in fact place-fixed for small changes in frequency, but for larger changes it accounts for "jumps" of the generation site to another "place-fixed" site. A different approach was taken by Bowman *et al.* (1997, 1998), who assign the difference in  $f_1$ - and  $f_2$ -sweep group delays to the build-up time of the cochlear filter, which they consider to make up a larger part of the  $f_2$ -sweep group delay than of the  $f_1$ -sweep group delay.

They do not explicitly take possible spatial shifts of the generation place into consideration.

The aim of the present study is to elaborate on a mathematical analysis for the place- and the wave-fixed hypotheses, yielding relations between the group delays determined with  $f_1$ - and with  $f_2$ -sweeps ( $D_1$  and  $D_2$ , respectively). The problem in giving a theoretical description of the wave-fixed model that allows verification against experimental data is the shifting of the presumed generation place  $X_2$  when sweeping  $f_2$ . This generates phase-place gradients which therefore need to be translated into phase-frequency gradients. This conversion is achieved by using the scale invariance of the cochlea (Zweig and Shera, 1995; Talmadge *et al.*, 1998, 1999). Additionally, the place- and wave-fixed models will be elaborated for the upper sideband DPOAEs under the assumption that these components originate from a place basal to  $X_2$ , near the DP characteristic place  $X_{dp}$ . The analytical results of the theory will be compared to guinea pig data published before (Schneider *et al.*, 1999, 2000). A tentative extension of the theory to LSB DPOAE reflection components from the  $X_{dp}$  characteristic place is discussed briefly.

## I. THEORY

Given any combination of primary frequencies ( $f_1, f_2$ ) (with  $f_2 > f_1$ ) that produces a measurable DPOAE, group delays can be obtained by applying small frequency changes of either of the two primaries around ( $f_1, f_2$ ) and determining the DPOAE's phase-frequency gradient. The DPOAE group delays derived with  $f_1$ - and  $f_2$ -sweep measurement paradigms will be denoted as  $D_1$  and  $D_2$ , respectively. The goal of the present analysis is to derive descriptions of  $D_1$  and  $D_2$  in both the place-fixed and the wave-fixed models that allow for interpretation of properties, mutual relations, and dependence on the order of the DPOAE.

In this section the assumptions are given, followed by the main steps of reasoning and the resulting expressions for  $D_1$ ,  $D_2$  and their relations. For further mathematical details the reader is referred to the Appendix. The first part of this section is dedicated to the basic conventions and assumptions. Then the situation is elaborated where the main DPOAE contribution originates from the region where the cochlear waves associated with  $f_1$  and  $f_2$  show maximal overlap, i.e., where the generation site  $x_g$  is assumed to be close to  $X_2$ , the characteristic place of  $f_2$ . Finally, the condition for the upper sideband components where the dominant generation site is likely to be near  $X_{dp}$ , i.e., basal to  $X_1$  and  $X_2$ , is treated separately.

The following general assumptions are made:

I. DPOAE generation is thought to be concentrated at a single generation site  $x_g$ . In the place-fixed model changes in  $x_g$  during a primary sweep are neglected:  $dx_g = 0$ . In the wave-fixed model changes in  $x_g$  are accounted for.

II. Contributions to the DPOAE in the ear canal from sites other than the generation site  $x_g$  are neglected.

The DPOAE group delay is defined as

$$D \equiv - \frac{d\varphi_{dp}}{d\omega_{dp}}, \quad (1)$$

where  $\varphi_{dp}$  is the phase of the DPOAE in the ear canal and  $\omega_{dp}$  is  $2\pi f_{dp}$ . The frequencies of the DPs are related to the primary frequencies as

$$\omega_{dp} = (n+1)\omega_1 - n\omega_2, \quad (2)$$

with  $n \geq 1$  for the lower sideband DPOAEs and  $n \leq -2$  for the upper sideband DPOAEs. The phase of the DPOAE ( $\varphi_{dp}$ ) in the ear canal is thought to consist of a forward component  $[\tilde{\varphi}(x_g; \omega_{dp})]$ , related to the phases of the primaries at  $x_g$  according to the relation of the  $\omega_i$ 's as given in Eq. (2), and a retrograde component  $[\tilde{\varphi}(x_g; \omega_{dp})]$ . The same applies to the group delay  $D$  which can thus be written as the sum of  $\vec{D}$  and  $\vec{D}$ , respectively. Since  $\varphi_{dp}$  depends on both  $x_g$  and  $\omega_{dp}$ , partial derivatives have to be used. The forward delay can be written as

$$\vec{D} = - \left( \frac{\partial \tilde{\varphi}(x_g; \omega_{dp})}{\partial \omega_{dp}} + \frac{\partial \tilde{\varphi}(x_g; \omega_{dp})}{\partial x_g} \frac{dx_g}{d\omega_{dp}} \right). \quad (3)$$

Combination of Eqs. (2) and (3) results in

$$\begin{aligned} \vec{D} = & -(n+1) \left( \frac{\partial \varphi(x_g; \omega_1)}{\partial \omega_1} \frac{d\omega_1}{d\omega_{dp}} + \frac{\partial \varphi(x_g; \omega_1)}{\partial x_g} \frac{dx_g}{d\omega_{dp}} \right) \\ & + n \left( \frac{\partial \varphi(x_g; \omega_2)}{\partial \omega_2} \frac{d\omega_2}{d\omega_{dp}} + \frac{\partial \varphi(x_g; \omega_2)}{\partial x_g} \frac{dx_g}{d\omega_{dp}} \right), \end{aligned} \quad (4a)$$

where  $\partial \varphi(x_g; \omega_i) / \partial \omega_i$  is the phase-frequency gradient for primary  $i$  at  $x_g$ , and  $\partial \varphi(x_g; \omega_i) / \partial x_g$  is the phase-place gradient for primary  $i$  at  $x_g$ . Accordingly, the backward delay can be written as

$$\vec{D} = - \frac{\partial \tilde{\varphi}(x_g; \omega_{dp})}{\partial \omega_{dp}} - \frac{\partial \tilde{\varphi}(x_g; \omega_{dp})}{\partial x_g} \frac{dx_g}{d\omega_{dp}}. \quad (4b)$$

According to the relation of the  $\omega_i$ 's, the gradients  $d\omega_i / d\omega_{dp}$  depend on which primary is swept:

$$d\omega_{dp} = (n+1)d\omega_1 \quad \text{and} \quad d\omega_2 = 0 \quad \text{for } f_1\text{-sweeps} \quad (5)$$

and

$$d\omega_{dp} = -nd\omega_2 \quad \text{and} \quad d\omega_1 = 0 \quad \text{for } f_2\text{-sweeps}. \quad (6)$$

The various phases in Eq. (4) are a function of the  $\omega_i$ 's and of the DPOAE generation place  $x_g$ . Since  $x_g$  is assumed constant in the place-fixed hypothesis, only the frequency dependence will remain. The same consideration applies for the wave-fixed hypothesis during an  $f_1$ -sweep when  $x_g$  equals the constant  $X_2$ . However, in the wave-fixed hypothesis, where, e.g.,  $x_g = X_2$  changes during an  $f_2$ -sweep, the DPOAE group delay is not only determined by phase-frequency gradients at one place but also by phase-place gradients at one frequency. To that end we will translate each phase-place gradient into a phase-frequency gradient so that all phase changes can be described as phase-frequency gradients. In addition, the original wave-fixed model has to be extended with a description of the characteristic frequency-place relation. We therefore make the following assumptions:

III. In a restricted region near  $x_g$  the phase distribution of a primary along the basilar membrane is frequency invariant, apart from a shift along the membrane that corresponds to the shift in frequency.

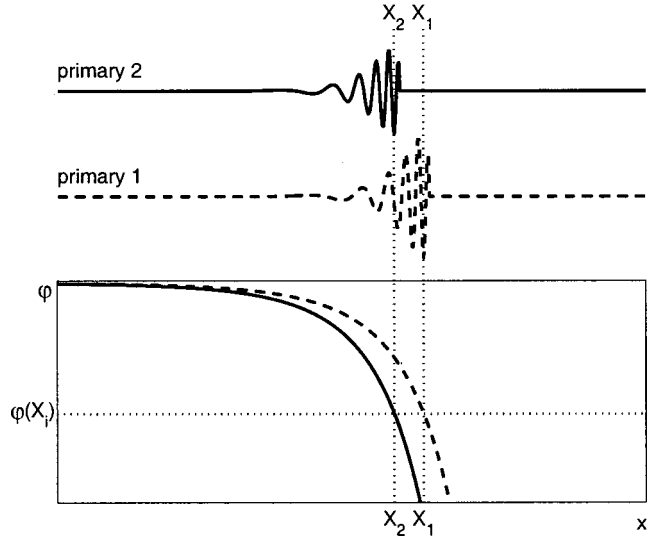


FIG. 1. Schematic representation of the mechanical response of the basilar membrane in response to primaries with frequencies  $f_1$  and  $f_2 (> f_1)$ . The figure demonstrates the principal consequence of cochlear scale invariance, i.e., that the phase of the response at the characteristic place is independent of frequency.

IV. The characteristic frequencies are logarithmically distributed along the basilar membrane.

Assumption III, the ‘‘frequency-shift invariance,’’ is a local approximation of the scale invariance of the cochlea (Zweig and Shera, 1995; Talmadge *et al.*, 1999, 2000) that was also used in the present context by, e.g., Moulin and Kemp (1996a, b). A visualization of this concept is given in Fig. 1, which also illustrates that, as a consequence, the phase at the characteristic place is independent of frequency. Assumption IV implies that

$$\frac{d\omega_i}{dX_i} = - \frac{\omega_i}{C}, \quad (7)$$

where  $C$  is a positive constant. Combination of assumptions III and IV results in the relation of the phase-place and phase-frequency gradients that takes an essential part in the analysis (see the Appendix):

$$\frac{\partial \varphi(x_g; \omega_i)}{\partial x_g} = \frac{\partial \varphi(x_g; \omega_i)}{\partial \omega_i} \frac{\omega_i}{C}. \quad (8)$$

Equations (1)–(8) form the base of the further analysis. In the original place- and wave-fixed hypotheses  $x_g$  is identified with the  $X_2$  of the primary pair ( $f_1, f_2$ ). This concept is elaborated in the next section. Since several studies suggest that for the USB DPOAEs  $x_g$  is basally from  $X_2$ , close to  $X_{dp}$ , a second section is dedicated to the assumption that for these DPOAEs  $x_g = X_{dp}$ .

## A. Analysis with $x_g$ related to $X_2$

### 1. The place-fixed model

In this model  $dx_g = 0$ , during both  $f_1$ - and  $f_2$ -sweeps. Therefore, only phase-frequency gradients are left in Eq. (4). With implementation of the  $d\omega_i$  relations [Eqs. (5) and (6)] the following descriptions of the  $f_1$ - and  $f_2$ -sweep group delays are found:

$$D_1 = -\frac{\partial\varphi(X_2; \omega_1)}{\partial\omega_1} - \frac{\partial\tilde{\varphi}(X_2; \omega_{dp})}{\partial\omega_{dp}}, \quad (9a)$$

$$D_2 = -\frac{\partial\varphi(X_2; \omega_2)}{\partial\omega_2} - \frac{\partial\tilde{\varphi}(X_2; \omega_{dp})}{\partial\omega_{dp}}. \quad (9b)$$

Both delays are the sum of the forward delay of the sweeping primary and the retrograde delay of the DP, which is equal in both expressions. Relations between  $D_1$  and  $D_2$  are

$$D_2 - D_1 = -\frac{\partial\varphi(X_2; \omega_2)}{\partial\omega_2} + \frac{\partial\varphi(X_2; \omega_1)}{\partial\omega_1} \quad (10a)$$

and

$$\frac{D_2}{D_1} = \frac{-\partial\varphi(X_2; \omega_2)/\partial\omega_2 - \partial\tilde{\varphi}(X_2; \omega_{dp})/\partial\omega_{dp}}{-\partial\varphi(X_2; \omega_1)/\partial\omega_1 - \partial\tilde{\varphi}(X_2; \omega_{dp})/\partial\omega_{dp}}. \quad (10b)$$

Equation (10a) means that, in the place-fixed model, the difference of the DPOAE group delay as measured with  $f_1$ - and  $f_2$ -sweeps is equal to the difference in forward group delay of the respective primaries at the place of maximal DP generation, here  $X_2$ . This difference is independent of the order of the DPOAE. Since the forward delay of  $f_2$  to  $X_2$  is larger than that of the smaller frequency  $f_1$  (see Fig. 1),  $D_2$  is larger than  $D_1$ .

## 2. The wave-fixed model

In this model  $x_g = X_2$  is constant during  $f_1$ -sweep and changes with frequency during  $f_2$ -sweep. Hence, during the  $f_1$ -sweep the hypotheses of place- and wave-fixed models are equal and  $D_1$  in the wave-fixed model is identical to the expression given in Eq. (9a) for the place-fixed model. During an  $f_2$ -sweep the first primary is constant, and as a result of the frequency-shift invariance the phase of the second primary at the moving generation site  $X_2$  does not change either (cf. Fig. 1). Therefore,  $\vec{D}_2$  is affected only by  $X_2$  moving through the  $f_1$ -wave, as induced by the  $f_2$ -sweep. The retrograde delay is affected by both the change in generation site and in DP frequency. Here we use the translation of phase-place gradients into phase-frequency gradients (see the Appendix). Combining the forward and retrograde delays we can describe  $D_2$  as

$$D_2 = \frac{(n+1)}{n} \frac{\omega_1}{\omega_2} \left( -\frac{\partial\varphi(X_2; \omega_1)}{\partial\omega_1} - \frac{\partial\tilde{\varphi}(X_2; \omega_{dp})}{\partial\omega_{dp}} \right). \quad (11)$$

Clearly,  $D_2$  depends on the order of the DPOAE. The relations of  $D_1$  and  $D_2$  are

$$D_2 - D_1 = \frac{(n+1)\omega_1 - n\omega_2}{n\omega_2} \times \left( -\frac{\partial\varphi(X_2; \omega_1)}{\partial\omega_1} - \frac{\partial\tilde{\varphi}(X_2; \omega_{dp})}{\partial\omega_{dp}} \right) \quad (12a)$$

and

$$\frac{D_2}{D_1} = \frac{(n+1)}{n} \frac{\omega_1}{\omega_2}. \quad (12b)$$

Thus, in the wave-fixed model, the ratio of the delays  $D_2$  and  $D_1$  depends only on the order of the DPOAE and on the

frequency ratio of the primaries at which the phase gradients are taken.

## B. Analysis for the USB DPOAEs with $x_g$ related to $X_{dp}$

In order to investigate the implication of a more basal generation site for the upper sideband DPOAEs in both the place- and wave-fixed models we calculated  $D_1$  and  $D_2$  using the assumption that this site is at the characteristic place of the DP:  $x_g = X_{dp}$ .

### 1. The place-fixed model

When the DPOAE generation site is at  $X_{dp}$  this location changes when either  $f_1$  or  $f_2$  is swept. In the place-fixed model this variation is neglected and consequently  $dx_g = 0$ . The delays thus depend only on the changes in the  $\omega_i$ 's so that Eqs. (4a) and (4b) can be reduced and  $D_1$  and  $D_2$  can be expressed as [cf. Eqs. (9a) and (9b)]:

$$D_1 = -\frac{\partial\varphi(X_{dp}; \omega_1)}{\partial\omega_1} - \frac{\partial\tilde{\varphi}(X_{dp}; \omega_{dp})}{\partial\omega_{dp}}, \quad (13a)$$

$$D_2 = -\frac{\partial\varphi(X_{dp}; \omega_2)}{\partial\omega_2} - \frac{\partial\tilde{\varphi}(X_{dp}; \omega_{dp})}{\partial\omega_{dp}}. \quad (13b)$$

The second term in Eqs. (13a) and (13b) represents the retrograde group delay of the DP from  $X_{dp}$  back to the ear canal, which is identical in  $D_1$  and  $D_2$ . Therefore, the difference of  $D_2$  and  $D_1$  equals the difference in the forward group delays of both primaries at the generation site:

$$D_2 - D_1 = -\frac{\partial\varphi(X_{dp}; \omega_2)}{\partial\omega_2} + \frac{\partial\varphi(X_{dp}; \omega_1)}{\partial\omega_1}. \quad (14a)$$

The ratio of both delays can be expressed as

$$\frac{D_2}{D_1} = \frac{-\partial\varphi(X_{dp}; \omega_2)/\partial\omega_2 - \partial\tilde{\varphi}(X_{dp}; \omega_{dp})/\partial\omega_{dp}}{-\partial\varphi(X_{dp}; \omega_1)/\partial\omega_1 - \partial\tilde{\varphi}(X_{dp}; \omega_{dp})/\partial\omega_{dp}}. \quad (14b)$$

### 2. The wave-fixed model

In the wave-fixed model the generation place  $x_g (= X_{dp})$  changes both when  $f_1$  and when  $f_2$  is swept. The frequency-shift invariance implies that the retrograde delays from  $X_{dp}$  are zero. Using the relations between DP and primary frequencies, the logarithmical frequency map, and the transformation of phase-place into phase-frequency gradients (see the Appendix), we can write  $D_1$  and  $D_2$  as

$$D_1 = -n \frac{\omega_2}{\omega_{dp}} \left( \frac{\partial\varphi(X_{dp}; \omega_2)}{\partial\omega_2} - \frac{\partial\varphi(X_{dp}; \omega_1)}{\partial\omega_1} \right), \quad (15a)$$

$$D_2 = -(n+1) \frac{\omega_1}{\omega_{dp}} \left( \frac{\partial\varphi(X_{dp}; \omega_2)}{\partial\omega_2} - \frac{\partial\varphi(X_{dp}; \omega_1)}{\partial\omega_1} \right). \quad (15b)$$

Note that both  $D_1$  and  $D_2$  are proportional to the difference of the  $-\partial\varphi(X_{dp}; \omega_i)/\partial\omega_i$  for  $i=1, 2$ , i.e., of the group delays of the primaries at  $X_{dp}$ . The difference between  $D_1$  and  $D_2$  reads



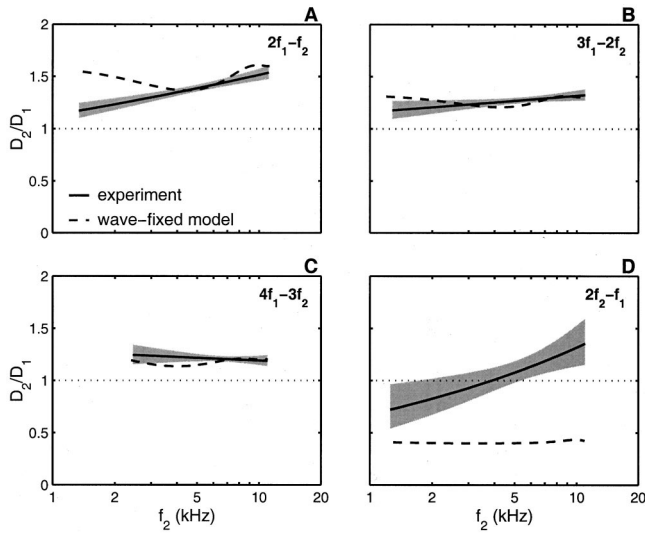


FIG. 2. Validation of the wave-fixed model against experimental data for the guinea pig (Schneider *et al.*, 1999). The ratio of group delays measured with  $f_2$ - and  $f_1$ -sweep paradigms,  $D_2/D_1$ , is plotted as a function of  $f_2$ . Solid line and shaded area represent mean and 95% confidence interval for the experimental data; the dashed lines give the predictions from the wave-fixed model [Eq. (12b)].

$$D_2 - D_1 = -\frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} + \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1}, \quad (16a)$$

which is equal to the expression for the place-fixed model [Eq. (14a)]. The ratio of  $D_2$  and  $D_1$  is

$$\frac{D_2}{D_1} = \frac{(n+1)\omega_1}{n\omega_2}, \quad (16b)$$

which is equal to the expression found for the lower sideband DPOAEs using the wave-fixed hypothesis with  $x_g = X_2$  [Eq. (12b)].

## II. RESULTS

In Fig. 2 the ratio  $D_2/D_1$ , predicted by the wave-fixed hypothesis [Eqs. (12b) and (16b)] is compared with previously published experimental data from the guinea pig (Schneider *et al.*, 1999). The DPOAEs with frequencies  $2f_1 - f_2$ ,  $3f_1 - 2f_2$ ,  $4f_1 - 3f_2$ , and  $2f_2 - f_1$  were measured, with both the  $f_1$ - and the  $f_2$ -sweep paradigm. Group delays  $D_1$  and  $D_2$  were determined at the primary frequency ratio  $f_2/f_1$  that yielded the maximum DPOAE level. The group delay versus  $f_2$  curves were fitted with a double-logarithmic regression line. The  $f_1$ - and  $f_2$ -sweep fits were subtracted which resulted in the fits for  $D_2/D_1$  as a function of  $f_2$  as presented in this figure (solid lines, with 95% confidence interval; for details see Schneider *et al.*, 1999). For the prediction of the wave-fixed hypothesis, the experimental ratios  $f_2/f_1$  that yielded the maximum DPOAE levels were substituted in Eq. (12b). This resulted in the dashed lines in Fig. 2. For the lower sideband DPOAEs  $2f_1 - f_2$ ,  $3f_1 - 2f_2$ , and  $4f_1 - 3f_2$ , there is good agreement between the experimental group delay ratio  $D_2/D_1$  and the ratio predicted by the wave-fixed theory [Fig. 2(a)–(c)], except for the lower frequencies of the  $2f_1 - f_2$  [Fig. 2(a)]. The experi-

mental results of the upper sideband DPOAE  $2f_2 - f_1$  are not consistent with the wave-fixed theory with the assumptions made in our theoretical analysis [Fig. 2(d)].

## III. DISCUSSION

Distortion product otoacoustic emissions provide a simple and powerful technique to study fundamental properties of cochlear filter processes. After many studies had focused on DPOAE amplitudes, the analysis of DPOAE phase to derive measures of group delay was introduced by Kimberley *et al.* (1993). While these and other authors primarily used the  $f_1$ -sweep paradigm, others advocated the  $f_2$ -sweep paradigm (O'Mahoney and Kemp, 1995; Moulin and Kemp, 1996a; Wable *et al.*, 1996). The origin of the difference between the two sweep paradigms and between the delays of upper and lower sideband DPOAEs has not been fully understood. Such understanding is required before  $f_1$ - and  $f_2$ -sweep group delays can be interpreted in terms of other measures of delay in the cochlea (e.g., Anderson *et al.*, 1971; Goldstein *et al.*, 1971; Eggermont, 1979; Versnel *et al.*, 1992; Kimberley *et al.*, 1993; Whitehead *et al.*, 1996; Talmadge *et al.*, 2000). Similarly, a better comprehension of DPOAE group delays may contribute to our knowledge of fundamental properties of the cochlear filter process. The lack of understanding of DPOAE group delays is due to incomplete knowledge of (1) retrograde phase delays, (2) the generation site of USB versus LSB DPOAEs, (3) group delays of possible contributions from the  $f_{dp}$  characteristic place for LSB components, and (4) spatial phase changes of one primary read out when sweeping the other. The principal aim of the present article was to give a comprehensive description of particularly the last issue. Based on several assumptions, mathematical descriptions of  $D_1$  and  $D_2$  could be derived for both the place-fixed and the wave-fixed model, several of which could be tested against experimental data.

### A. Assumptions and theoretical considerations

The analysis is based on four main assumptions. The assumed frequency-shift invariance of the frequency-place relation (assumption III) is a local approximation of the generally accepted concept of cochlear scale invariance (Sondhi, 1978; Viergever, 1980; Zweig and Shera, 1995), incorporated in the analysis of DPOAE group delays by several authors (Moulin and Kemp, 1996b; Talmadge *et al.*, 1998, 1999, 2000). The assumed logarithmic tonotopic map (assumption IV) finds similarly general application.

The neglect of contributions to the DPOAE recorded in the ear canal other than from the generation site (assumption II) is not self-evident. For LSB DPOAEs contributions from particularly the DP characteristic place have been identified (Kummer *et al.*, 1995; Brown *et al.*, 1996; Martin *et al.*, 1987, 1998). In the two-source model, recently described by Shera and Guinan (1999) and Talmadge *et al.* (1998, 1999), the DPOAE in the ear canal is a combination of the component generated by nonlinear interaction of the primaries around  $X_2$ , and a component arising from linear coherent reflection of the apically traveling wave close to  $X_{dp}$ . The relative strength of those two components depends among others on the primary levels and the primary fre-

quency ratio (Fahey and Allen, 1997; Knight and Kemp, 2000). The interference of the reflection component with the generation component is thought to be responsible for the fine structure in the DPOAE amplitude and for irregularities in the phase-frequency profile (Brown *et al.*, 1996; Heitmann *et al.*, 1998; Talmadge *et al.*, 1999). In this study, we have mainly focused at the  $X_2$  generation site component since all group delay studies referred to in the present context were performed at the best  $f_2/f_1$  ratio where that component presumably dominates the response (Knight and Kemp, 2000). Possible effects of residual reflection components on measured phase-frequency profiles, and therefore on DPOAE group delays, are further averaged out when considering group average data as was done, e.g., in Fig. 2. Yet, some extension of the theory to the reflection component is feasible. As a corollary the last section of the Appendix indicates (assuming that both the generation and the reflection process obey wave-fixed mechanisms) that the  $D_2/D_1$  relation derived for the wave-fixed model applies also for the reflection component itself. Therefore the results regarding the  $D_2/D_1$  ratio hold also if the measured DPOAE is dominated by the reflection component. The intermediate situation, where both components are approximately equal, is more difficult to interpret. More fundamentally, while DP generation at  $X_2$  is likely to obey wave-fixed mechanisms, reflection at  $X_{dp}$  is thought to be due to local anatomical irregularities and therefore must have an essentially place-fixed character, a notion supported by recent experimental data (Talmadge *et al.*, 1999). An extension of the analysis for the wave-fixed generation component with a place-fixed reflection component greatly complicates the analysis and prohibits the formulation of transparent expressions for either  $D_2 - D_1$  or  $D_2/D_1$ .

Finally, the assumption was made that DP generation is concentrated on a single place  $x_g$  (assumption I). In the real situation the generation of nonlinear distortion is obviously distributed over a certain range along the cochlear partition where  $f_1$  and  $f_2$  activation patterns overlap. The assumption is thus based on a “center of mass” representation of this distributed source. The model, in fact, does not require  $x_g$  to be equal to, e.g., the place  $X_2$  with characteristic frequency  $f_2$  as defined, e.g., with a neural frequency threshold curve, but merely states that  $x_g$  is fixed to (and shifting with) the envelope of the  $f_2$  response. This also means that, for example,  $x_g$  may depend on the level of the primaries.

After application of these assumptions the analysis leads to expressions for  $D_1$  and  $D_2$  fully in terms of phase-frequency gradients, in both the place-fixed and the wave-fixed models. Several aspects of these expressions are noteworthy. First, in the comparisons of  $D_1$  and  $D_2$  no explicit assumptions were made regarding (in)equality of forward and retrograde group delays. The expressions for  $D_2 - D_1$  in the place-fixed model and for  $D_2/D_1$  in the wave-fixed model [Eqs. (10a) and (12b)] are independent of the absolute values of the retrograde delay. In contrast, the retrograde delay term is explicitly included in the other expressions, Eqs. (10b) and (12a). Second, the  $D_2/D_1$  ratio in the wave-fixed model [Eq. (12b)] depends only on the order of the DPOAE and on the frequency ratio of the primaries. Thus,

$D_2/D_1$  does not depend on parameters such as stimulus level, the absolute primary frequencies, and, e.g., a possible pathological condition of the cochlea, as long as the assumptions are not violated. Several of these predictions can be subjected to experimental verification. Both the place-fixed and the wave-fixed models lead to the conclusion that  $D_2 > D_1$ , as can be seen from Eqs. (10a) and (12b). Finally, since the predicted  $D_2/D_1$  for the USB component in the wave-fixed model is identical whether  $X_2$  or  $X_{dp}$  is the assumed generation place [Eqs. (12b) and (16b)], no distinction between those two alternatives can be made on the basis of the present analysis.

## B. Comparison with earlier theoretical analyses of $f_1$ - and $f_2$ -sweep group delays

Several general properties of  $f_1$ - and  $f_2$ -sweep group delays were already indicated by Moulin and Kemp (1996b). The present analysis reconfirms and extends several of their results, and particularly gives a more thorough theoretical foundation to others. In their analysis of the place-fixed model forward delays were predicted to be larger for the  $f_2$ - than for the  $f_1$ -sweep paradigm; retrograde delays were implicitly assumed to be equal, making  $D_2 > D_1$ ; their mathematical expressions were partly identical to those in the present paper. The same authors elaborated the wave-fixed model on the basis of additional qualitative assumptions, and also predicted forward delays larger with the  $f_2$ - than with the  $f_1$ -sweep. A quantitative expression could not be given, since no explicit expression for the spatial phase changes of the  $f_1$ -wave, read out by the moving  $X_2$ , was available. No explicit indication was given of the difference of retrograde delays in the wave-fixed model. Similarly, Moulin and Kemp (1996b) derived predictions regarding the relative group delays of different order DPOAEs. The predicted equality of forward delays in the place-fixed model, irrespective of the order of the DPOAE or the choice of sweep paradigm, was more formally reproduced in the present theory. For the wave-fixed model, forward delays were predicted to depend on DPOAE order in a qualitatively very similar way as in the present theory. Again, however, since no explicit expression could be given regarding the spatial phase changes of the  $f_1$ -wave, no definite quantitative results could be formulated by Moulin and Kemp (1996b).

In a different approach, Bowman *et al.* (1997, 1998) developed a description of DPOAE delays in which the build-up time of the cochlear filter plays an important role. They posed that, since generation takes place in the  $X_2$ -region where rapid phase changes of  $f_2$  occur, the filter build-up time is included in the DPOAE delay for the  $f_2$ -sweep, but not for the  $f_1$ -sweep. In their view, DPOAE delays are the sum of a forward delay of the primaries to the generation place, a filter response time, and a backward delay of the DP component. A possible link between that approach and the place-fixed and wave-fixed models is not evident at the present stage. Yet, given that the concept of a change of generation place in the  $f_2$ -sweep paradigm was not explicitly accounted for by these authors, their approach is essentially place-fixed. The observation of Bowman *et al.* (1997) that  $D_2 - D_1$  decreases with increasing stimulus lev-

els is at least qualitatively consistent with a stimulus-independent  $D_2/D_1$  ratio predicted by the present theory, given the decrease of both  $D_2$  and  $D_1$  with intensity demonstrated by Bowman *et al.* (1997). Recently, the expression for the  $D_2/D_1$  ratio in the wave-fixed model [Eq. (12b)] was simultaneously presented in preliminary form by ourselves (Schneider *et al.*, 2000) and (for  $n=1$ ) by Talmadge *et al.* (2000).

### C. Validation against experimental data

The formal mathematical character of the expressions for  $D_1$  and  $D_2$  does not allow a direct quantitative validation of the theory against experimental data. The only simple expression allowing such validation is the  $D_2/D_1$  ratio in the wave-fixed model [Eq. (12b)]. From the other expressions the difference  $D_2 - D_1$  in the place-fixed model [Eq. (10a)] allows some qualitative interpretation that can be compared to experimental data.

The results shown in Fig. 2(a)–(c) indicate that the theoretically predicted  $D_2/D_1$  ratio quantitatively matches the experimental data for each LSB DPOAE component, and that it shows the appropriate dependence on the order of the DPOAE. The experimental data used for this comparison are consistent with other published data on  $D_2/D_1$  (Moulin and Kemp, 1996b; Whitehead *et al.*, 1996). Regarding the values of  $D_1$  and  $D_2$  themselves, for the LSB components,  $D_1$  was found to be invariant with the order of the DPOAE, while  $D_2$  decreases with increasing order (Moulin and Kemp, 1996b; Schneider *et al.*, 1999). This dependence is theoretically predicted in the wave-fixed model but not in the place-fixed model, where the delay difference  $D_2 - D_1$  should be order independent [Eqs. (10a) and (12a)]. Therefore we conclude that, for the above-mentioned data, the variation with  $f_1$ - versus  $f_2$ -sweep of LSB DPOAE group delays and their dependence on DPOAE order are consistent with the wave-fixed model, and not with the place-fixed model.

The theoretical prediction of the group delay ratio  $D_2/D_1$  for the  $2f_2 - f_1$  upper sideband component shows a large discrepancy with the experimental data [Fig. 2(d)]. Apparently, the wave-fixed model combined with the further assumptions made is not appropriate for this component, independent whether  $X_2$  or  $X_{dp}$  is assumed to be the dominant generation site; this may be due to inappropriate description of the shift of the generation site with changing primaries, and therefore does not necessarily conflict with the notion that for the USB components  $x_g$  is probably close to  $X_{dp}$ . Some further quantitative validation in terms of a comparison of USB versus LSB components is possible if both would obey the place-fixed model. From the theory the difference  $D_2 - D_1$  is larger for the LSB generated at  $X_2$  than for the USB generated at  $X_{dp}$  [Eq. (10a) versus (14a)], because particularly the first term in Eq. (10a) is dominant (cf. Fig. 1). Our experimental data are consistent with this prediction when a comparison is made, e.g., between  $2f_1 - f_2$  and  $2f_2 - f_1$  components (Schneider *et al.*, 1999). This comparison, however, is precarious since the place-fixed model had already appeared inappropriate for the LSB components. As far as the present analysis reaches, the place-fixed model

cannot be rejected for the USB components if generated basically from the primary region.

## IV. CONCLUSIONS

A theoretical analysis was presented of the phase behavior of distortion product otoacoustic emissions, which aims at elucidating the differences between delays observed in the  $f_1$ - and in the  $f_2$ -sweep paradigm,  $D_1$  and  $D_2$ , respectively. By making a local approximation of the cochlear scale invariance, a mathematical conversion from phase-place to phase-frequency gradients could be formulated. This allows incorporating the effects of a spatial shift of the  $X_2$  generation place in the  $f_2$ -sweep paradigm in the wave-fixed model. In this way expressions could be derived for  $D_1$ ,  $D_2$ , and their ratios and differences in either of the two models, under the condition where the DPOAE recorded in the ear canal is dominated by the contribution from the generation site, and not, e.g., by reflection components. Experimental validation of these expressions indicates that group delays of lower sideband DPOAEs ( $2f_1 - f_2$ ,  $3f_1 - 2f_2$ ,  $4f_1 - 3f_2$ ) are most consistent with the wave-fixed model. In particular, the difference between the group delays obtained with  $f_1$ - and  $f_2$ -sweeps could be fully explained by the shift of the  $X_2$  generation site in the  $f_2$ -sweep paradigm. Group delays of the upper sideband DPOAE ( $2f_2 - f_1$ ) are inconsistent with both wave-fixed and place-fixed models, independent of whether DP generation is assumed to take place at the  $f_2$  characteristic place or at the  $f_{dp}$  characteristic place. Apparently, the assumptions made regarding the location of the generation site combined with its shift when sweeping either of the two primaries are inappropriate.

## ACKNOWLEDGMENT

This study was supported by the Netherlands Organization for Scientific Research and by the Heinsius Houbolt Fund.

## APPENDIX

### 1. Frequency-shift invariance

As a consequence of assumption III the phase  $\varphi(X_i; \omega_i)$  of a particular frequency  $\omega_i$  at its characteristic place  $X_i$  is equal for all frequencies (see Fig. 1):

$$\frac{d\varphi(X_i; \omega_i)}{d\omega_i} = \frac{\partial\varphi(X_i; \omega_i)}{\partial\omega_i} + \frac{\partial\varphi(X_i; \omega_i)}{\partial X_i} \frac{dX_i}{d\omega_i} = 0. \quad (A1)$$

The frequency-shift invariance prescribes that when the radial frequency is changed from  $\omega_i$  to  $\omega'_i$  the new phase at  $x_g$  can be found from the original phase distribution for  $\omega_i$  at place  $x'_g$  (Fig. A1):

$$\varphi(x_g; \omega'_i) = \varphi(x'_g; \omega_i), \quad (A2)$$

where  $x'_g = x_g + \Delta x_g$ , and  $(-\Delta x_g)$  corresponds to the change in characteristic place from  $X_i$  to  $X'_i$ :  $\Delta x_g = -\Delta X_i = -(X'_i - X_i)$ . Therefore, we can write

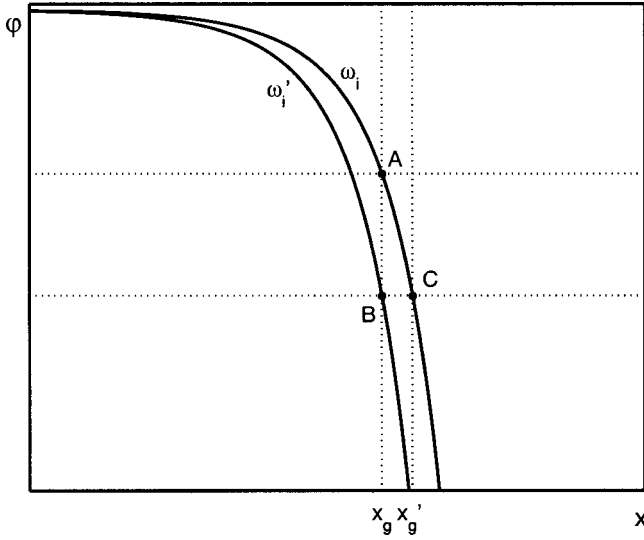


FIG. A1. Illustration of the conversion from phase-place to phase-frequency gradients [Eq. (A6)]. The phase of  $\omega_i$  at place  $x_g$  is given by  $\varphi(x_g; \omega_i)$  (point A). When the frequency is changed from  $\omega_i$  to  $\omega_i'$ , the new phase  $\varphi(x_g; \omega_i')$  of  $\omega_i'$  at place  $x_g$  (point B) can be found from the original phase distribution of  $\omega_i$  at place  $x_g'$  [ $\varphi(x_g'; \omega_i)$ ; point C], with  $x_g' = x_g + \Delta x_g$  and  $\Delta x_g = X_i - X_i'$ , the difference between the characteristic places of the frequencies  $\omega_i$  and  $\omega_i'$ .

$$\begin{aligned} \frac{\varphi(x_g'; \omega_i) - \varphi(x_g; \omega_i)}{\Delta x_g} &= \frac{\varphi(x_g; \omega_i') - \varphi(x_g; \omega_i)}{\Delta x_g} \\ &= - \frac{\varphi(x_g; \omega_i') - \varphi(x_g; \omega_i)}{\Delta X_i}. \end{aligned} \quad (\text{A3})$$

Now we can formulate the desired relation of a phase-place gradient and a phase-frequency gradient:

$$\frac{\partial \varphi(x_g; \omega_i)}{\partial x_g} = - \frac{\partial \varphi(x_g; \omega_i)}{\partial \omega_i} \frac{d\omega_i}{dX_i}, \quad (\text{A4})$$

where  $d\omega_i/dX_i$  can be found from assumption IV expressed as

$$\frac{d\omega_i}{dX_i} = - \frac{\omega_i}{C}, \quad (\text{A5})$$

where  $C$  is a constant. Combination of Eqs. (A4) and (A5) results in

$$\frac{\partial \varphi(x_g; \omega_i)}{\partial x_g} = \frac{\partial \varphi(x_g; \omega_i)}{\partial \omega_i} \frac{\omega_i}{C}. \quad (\text{A6})$$

## 2. The wave-fixed model

### a. $x_g$ related to $X_2$

Since in the wave-fixed model for the  $f_2$ -sweep the DP generation place is thought to shift according to the  $X_2$ , the forward DPOAE group delay [Eq. (4a)] can be expressed in relation to the primary parameters  $X_2$ ,  $\omega_1$ , and  $\omega_2$ :

$$\begin{aligned} \tilde{D}_2 &= -(n+1) \left( \frac{\partial \varphi(X_2; \omega_1)}{\partial \omega_1} \frac{d\omega_1}{d\omega_{dp}} + \frac{\partial \varphi(X_2; \omega_1)}{\partial X_2} \frac{dX_2}{d\omega_2} \frac{d\omega_2}{d\omega_{dp}} \right) \\ &+ n \frac{d\omega_2}{d\omega_{dp}} \left( \frac{\partial \varphi(X_2; \omega_2)}{\partial \omega_2} + \frac{\partial \varphi(X_2; \omega_2)}{\partial X_2} \frac{dX_2}{d\omega_2} \right), \end{aligned} \quad (\text{A7})$$

in which the first term is zero because  $\omega_1$  is constant ( $d\omega_1 = 0$ ) and the last two terms together are zero as a direct consequence of the frequency-shift invariance [Eq. (A1)]. This means that when using the  $f_2$ -sweep paradigm the forward delay is affected by a change in phase of the  $f_1$  component at the generation site caused by the variation of this site. The phase-place gradient can be translated into a phase-frequency gradient [Eq. (A6)]:

$$\tilde{D}_2 = -(n+1) \left( \frac{\partial \varphi(X_2; \omega_1)}{\partial \omega_1} \frac{\omega_1}{C} \frac{dX_2}{d\omega_2} \frac{d\omega_2}{d\omega_{dp}} \right). \quad (\text{A8})$$

When Eqs. (6) and (A5) are applied, the expression of this forward delay becomes

$$\tilde{D}_2 = \frac{(n+1)}{n} \frac{\omega_1}{\omega_2} \left( - \frac{\partial \varphi(X_2; \omega_1)}{\partial \omega_1} \right). \quad (\text{A9})$$

Similarly, the backward delay is written as

$$\tilde{D}_2 = - \frac{\partial \tilde{\varphi}(X_2; \omega_{dp})}{\partial \omega_{dp}} \frac{\partial \tilde{\varphi}(X_2; \omega_{dp})}{\partial X_2} \frac{dX_2}{d\omega_2} \frac{d\omega_2}{d\omega_{dp}}. \quad (\text{A10})$$

Again applying the relation between the changes in  $\omega_2$  and  $\omega_{dp}$  [Eq. (6)], using the frequency-shift invariance and the logarithmic distribution of the characteristic frequencies [Eqs. (A5) and (A6)]  $\tilde{D}_2$  is written as

$$\begin{aligned} \tilde{D}_2 &= - \frac{\partial \tilde{\varphi}(X_2; \omega_{dp})}{\partial \omega_{dp}} \left( \frac{\partial \tilde{\varphi}(X_2; \omega_{dp})}{\partial \omega_{dp}} \frac{\omega_{dp}}{C} \right) \frac{dX_2}{d\omega_2} \frac{(-1)}{n} \\ &= - \frac{\partial \tilde{\varphi}(X_2; \omega_{dp})}{\partial \omega_{dp}} \left( 1 + \frac{\omega_{dp}}{n\omega_2} \right). \end{aligned} \quad (\text{A11})$$

With use of the relation between  $\omega_1$ ,  $\omega_2$ , and  $\omega_{dp}$  [Eq. (2)] the backward delay  $\tilde{D}_2$  can be rewritten as

$$\tilde{D}_2 = \frac{(n+1)}{n} \frac{\omega_1}{\omega_2} \left( - \frac{\partial \tilde{\varphi}(X_2; \omega_{dp})}{\partial \omega_{dp}} \right). \quad (\text{A12})$$

Combination of the forward and the backward delay [Eqs. (A9) and (A12)] yields

$$D_2 = \frac{(n+1)}{n} \frac{\omega_1}{\omega_2} \left( - \frac{\partial \varphi(X_2; \omega_1)}{\partial \omega_1} - \frac{\partial \tilde{\varphi}(X_2; \omega_{dp})}{\partial \omega_{dp}} \right). \quad (\text{A13})$$

### b. USB DPOAEs with $x_g$ related to $X_{dp}$

Here we assume the prominent contribution to the upper sideband DPOAEs to come from  $x_g = X_{dp}$  that changes for both the  $f_1$ - and  $f_2$ -sweep paradigms.

$$\begin{aligned} \tilde{D} &= -(n+1) \left( \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial X_{dp}} \frac{dX_{dp}}{d\omega_{dp}} + \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} \frac{d\omega_1}{d\omega_{dp}} \right) \\ &+ n \left( \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial X_{dp}} \frac{dX_{dp}}{d\omega_{dp}} + \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} \frac{d\omega_2}{d\omega_{dp}} \right), \end{aligned} \quad (\text{A14a})$$

$$\tilde{D} = -\frac{\partial \tilde{\varphi}(X_{dp}; \omega_{dp})}{\partial \omega_{dp}} - \frac{\partial \tilde{\varphi}(X_{dp}; \omega_{dp})}{\partial X_{dp}} \frac{dX_{dp}}{d\omega_{dp}}. \quad (\text{A14b})$$

Using the relation of relative changes in the  $\omega_i$ 's [Eqs. (5) and (6)] in Eq. (A14a) we can describe the forward delays as

$$\begin{aligned} \tilde{D}_1 = & -(n+1) \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial X_{dp}} \frac{dX_{dp}}{d\omega_{dp}} - \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} \\ & + n \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial X_{dp}} \frac{dX_{dp}}{d\omega_{dp}}, \end{aligned} \quad (\text{A15a})$$

$$\begin{aligned} \tilde{D}_2 = & -(n+1) \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial X_{dp}} \frac{dX_{dp}}{d\omega_{dp}} - \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} \\ & + n \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial X_{dp}} \frac{dX_{dp}}{d\omega_{dp}}. \end{aligned} \quad (\text{A15b})$$

For simplification of these relations we used the frequency-shift invariance in combination with the logarithmic distribution of the characteristic frequencies over the basilar membrane [Eqs. (A5) and (A6)] as well as the relation between the  $d\omega_i$ 's [Eqs. (5) and (6)]:

$$\begin{aligned} \tilde{D}_1 = & (n+1) \frac{\omega_1}{\omega_{dp}} \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} - \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} \\ & - n \frac{\omega_2}{\omega_{dp}} \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2}, \end{aligned} \quad (\text{A16a})$$

$$\begin{aligned} \tilde{D}_2 = & (n+1) \frac{\omega_1}{\omega_{dp}} \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} - \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} \\ & - n \frac{\omega_2}{\omega_{dp}} \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2}, \end{aligned} \quad (\text{A16b})$$

which can be reduced to

$$\tilde{D}_1 = -n \frac{\omega_2}{\omega_{dp}} \left( \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} - \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} \right), \quad (\text{A17a})$$

$$\tilde{D}_2 = -(n+1) \frac{\omega_1}{\omega_{dp}} \left( \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} - \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} \right). \quad (\text{A17b})$$

Both the backward delays, which are expressed by Eq. (A14b), are zero according to the frequency-shift invariance expressed by Eq. (A1):

$$\tilde{D}_1 = \tilde{D}_2 = 0. \quad (\text{A18})$$

Therefore, the total group delays are equal to the forward delays:

$$D_1 = -n \frac{\omega_2}{\omega_{dp}} \left( \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} - \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} \right), \quad (\text{A19a})$$

$$D_2 = -(n+1) \frac{\omega_1}{\omega_{dp}} \left( \frac{\partial \varphi(X_{dp}; \omega_2)}{\partial \omega_2} - \frac{\partial \varphi(X_{dp}; \omega_1)}{\partial \omega_1} \right). \quad (\text{A19b})$$

### 3. LSB DPOAEs with OAE coming from reflection site $X_{dp}$

Here we assume the DP to be generated at  $x_g = X_2$  from which a retrograde wave towards the base and a forward wave towards the apex starts. The latter wave is reflected at  $X_{dp}$  from where a retrograde wave starts. The contribution of this wave to the DPOAE has the group delay

$$D_i^r = \tilde{D}_i(X_2; \omega_1, \omega_2) + \tilde{D}_i(X_2 \rightarrow X_{dp}; \omega_{dp}) + \tilde{D}_i(X_{dp}; \omega_{dp}), \quad (\text{A20})$$

where  $\tilde{D}_i(X_2; \omega_1, \omega_2)$  is the forward delay formulated earlier as  $\tilde{D}_i$ ,  $\tilde{D}_i(X_2 \rightarrow X_{dp}; \omega_{dp})$  is the forward delay for the generated DP traveling from  $X_2$  to  $X_{dp}$ , and  $\tilde{D}_i(X_{dp}; \omega_{dp})$  is the retrograde delay of the reflected wave. The second term of Eq. (A20) is written as

$$\tilde{D}_i(X_2 \rightarrow X_{dp}; \omega_{dp}) = \tilde{D}_i(X_{dp}; \omega_{dp}) - \tilde{D}_i(X_2; \omega_{dp}), \quad (\text{A21})$$

so that Eq. (A20) becomes

$$\begin{aligned} D_i^r = & \tilde{D}_i(X_2; \omega_1, \omega_2) + \tilde{D}_i(X_{dp}; \omega_{dp}) + \tilde{D}_i(X_{dp}; \omega_{dp}) \\ & - \tilde{D}_i(X_2; \omega_{dp}). \end{aligned} \quad (\text{A22})$$

According to assumption III or Eq. (A1) the second and third component are zero. The resulting terms, for  $f_1$ -sweep, give

$$\begin{aligned} D_1^r = & \tilde{D}_1(X_2; \omega_1, \omega_2) - \tilde{D}_1(X_2; \omega_{dp}) \\ = & -\frac{\partial \varphi(X_2; \omega_1)}{\partial \omega_1} - \frac{\partial \varphi(X_2; \omega_{dp})}{\partial \omega_{dp}}, \end{aligned} \quad (\text{A23a})$$

and for  $f_2$ -sweep

$$\begin{aligned} D_2^r = & \tilde{D}_2(X_2; \omega_1, \omega_2) - \tilde{D}_2(X_2; \omega_{dp}) \\ = & \frac{(n+1)}{n} \frac{\omega_1}{\omega_2} \left( -\frac{\partial \varphi(X_2; \omega_1)}{\partial \omega_1} - \frac{\partial \varphi(X_2; \omega_{dp})}{\partial \omega_{dp}} \right), \end{aligned} \quad (\text{A23b})$$

so that also for the DPOAE reflection component holds

$$\frac{D_2^r}{D_1^r} = \frac{(n+1)}{n} \frac{\omega_1}{\omega_2}. \quad (\text{A24})$$

- Anderson, D. J., Rose, J. E., Hind, J. E., and Brugge, J. F. (1971). "Temporal position of discharges of single auditory nerve fibers with the cycle of a sine-wave stimulus: frequency and intensity effects," *J. Acoust. Soc. Am.* **49**, 1131–1139.
- Bowman, D. M., Brown, D. K., Eggermont, J. J., and Kimberley, B. P. (1997). "The effect of sound intensity on  $f_1$ -sweep and  $f_2$ -sweep distortion product otoacoustic emissions phase delay estimates in human adults," *J. Acoust. Soc. Am.* **101**, 1550–1559.
- Bowman, D. M., Eggermont, J. J., Brown, D. K., and Kimberley, B. P. (1998). "Estimating cochlear filter response properties from distortion product otoacoustic emission (DPOAE) phase delay measurements in normal hearing adults," *Hear. Res.* **119**, 14–26.
- 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., Harris, F. P., and Beveridge, H. A. (1996). "Two sources of acoustic distortion products from the human cochlea," *J. Acoust. Soc. Am.* **100**, 3260–3267.
- Eggermont, J. J. (1979). "Narrow-band AP latencies in normal and recruiting human ears," *J. Acoust. Soc. Am.* **65**, 464–470.
- Fahey, P. F., and Allen, J. B. (1997). "Measurement of distortion product phase in the ear canal of the cat," *J. Acoust. Soc. Am.* **102**, 2880–2891.

- Goldstein, J. L., Baer, T., and Kiang, N. Y. S. (1971). "A theoretical treatment of latency, group delay and tuning characteristics for auditory nerve responses to clicks and tones," in *Physiology of the Auditory System*, edited by M. B. Sachs (National Educational Consultants, Baltimore), pp. 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.
- Kemp, D. T. (1986). "Otoacoustic emissions, travelling waves and cochlear mechanisms," *Hear. Res.* **22**, 95–104.
- Kimberley, B. P., Brown, D. K., and Eggermont, J. J. (1993). "Measuring human cochlear traveling wave delay using distortion product otoacoustic emission phase responses," *J. Acoust. Soc. Am.* **94**, 1343–1350.
- Knight, R. D., and Kemp, D. T. (2000). "Indications of different distortion product otoacoustic emission mechanisms from a detailed  $f_1$ ,  $f_2$  area study," *J. Acoust. Soc. Am.* **107**, 457–473.
- 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.
- Martin, G. K., Jassir, D., Stagner, B. B., Whitehead, M. L., and Lonsbury-Martin, B. L. (1998). "Locus of generation for the  $2f_1 - f_2$  vs  $2f_2 - f_1$  distortion-product otoacoustic emissions in normal-hearing humans revealed by suppression tuning, onset latencies, and amplitude correlations," *J. Acoust. Soc. Am.* **103**, 1957–1971.
- 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.
- Moulin, A., and Kemp, D. T. (1996a). "Multicomponent acoustic distortion product otoacoustic emission phase in humans. I. General characteristics," *J. Acoust. Soc. Am.* **100**, 1617–1639.
- Moulin, A., and Kemp, D. T. (1996b). "Multicomponent acoustic distortion product otoacoustic emission phase in humans. II. Implications for distortion product otoacoustic emission generation," *J. Acoust. Soc. Am.* **100**, 1640–1662.
- O'Mahoney, C. F., and Kemp, D. T. (1995). "Distortion product otoacoustic emission delay measurement in human ears," *J. Acoust. Soc. Am.* **97**, 3721–3735.
- Schneider, S., Prijs, V. F., and Schoonhoven, R. (1999). "Group delays of distortion product otoacoustic emissions in the guinea pig," *J. Acoust. Soc. Am.* **105**, 2722–2730.
- Schneider, S., Prijs, V. F., Schoonhoven, R., and van Hengel, P. W. J. (2000). " $f_1$ - versus  $f_2$ -sweep group delays of distortion product otoacoustic emissions in the guinea pig; experimental results and theoretical predictions," in *Recent Developments in Auditory Mechanics*, edited by H. Wada *et al.* (World Scientific, Singapore), pp. 360–366.
- Shera, C. A., and Guinan, J. J. (1999). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," *J. Acoust. Soc. Am.* **105**, 782–798.
- Sondhi, M. M. (1978). "Method for computing motion in a two-dimensional cochlear model," *J. Acoust. Soc. Am.* **63**, 1468–1477.
- Talmadge, C. L., Long, G. R., Tubis, A., and Dhar, S. (1999). "Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions," *J. Acoust. Soc. Am.* **105**, 275–292.
- Talmadge, C. L., Tubis, A., Long, G. R., and Piskorski, P. (1998). "Modeling otoacoustic emission and hearing threshold fine structures," *J. Acoust. Soc. Am.* **104**, 1517–1543.
- Talmadge, C. L., Tubis, A., Tong, C., Long, G. R., and Dhar, S. (2000). "Temporal aspects of otoacoustic emissions," in *Recent Developments in Auditory Mechanics*, edited by H. Wada *et al.* (World Scientific, Singapore), pp. 353–359.
- 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," *Hear. Res.* **59**, 138–156.
- Viergever, M. A. (1980). "Mechanics of the ear—A mathematical approach," Ph.D. dissertation, Delft University of Technology (Delft University Press, Delft, The Netherlands).
- Wable, J., Collet, L., and Chéry-Croze, S. (1996). "Phase delay measurements of distortion product otoacoustic emissions at  $2f_1 - f_2$  and  $2f_2 - f_1$  in human ears," *J. Acoust. Soc. Am.* **100**, 2228–2235.
- Whitehead, M. L., Stagner, B. B., Martin, G. K., and Lonsbury-Martin, B. L. (1996). "Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency," *J. Acoust. Soc. Am.* **100**, 1663–1679.
- Zweig, G., and Shera, C. A. (1995). "The origins of periodicity in the spectrum of otoacoustic emissions," *J. Acoust. Soc. Am.* **98**, 2018–2047.

# The behavior of spontaneous otoacoustic emissions during and after postural changes

Emile de Kleine<sup>a)</sup> and Hero P. Wit

Department of Otorhinolaryngology, University Hospital Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands

Pim van Dijk

Department of Otorhinolaryngology and Head & Neck Surgery, University Hospital Maastricht, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands

Paul Avan

Biophysics Department, Faculty of Medicine, P.O. Box 38, 63001 Clermont-Ferrand, France

(Received 23 November 1999; revised 25 February 2000; accepted 8 March 2000)

Spontaneous otoacoustic emissions (SOAEs) were studied in humans during and after postural changes. The subjects were tilted from upright to a recumbent position (head down 30 degrees) and upright again in a 6-min period. The SOAEs were recorded continuously and analyzed off-line. The tilting caused a change in the SOAE spectrum for all subjects. Frequency shifts of 10 Hz, together with changes of amplitude (5 dB) and width (5 Hz), were typically observed. However, these changes were observed in both directions (including the appearance and disappearance of emission peaks). The most substantial changes occurred in the frequency region below 2 kHz. An increase of the intracranial pressure, and consequently of the intracochlear fluid pressure, is thought to result in an increased stiffness of the cochlear windows, which is probably mainly responsible for the SOAE changes observed after the downward turn. The time for the spectrum to regain stability after a postural change differed between the two maneuvers: 1 min for the downward and less than 10 s for the upward turn. © 2000 Acoustical Society of America. [S0001-4966(00)04306-X]

PACS numbers: 43.64.Jb, 43.64.Bt, 43.64.Kc [BLM]

## INTRODUCTION

Otoacoustic emissions (OAEs) are acoustic signals generated in the inner ear and measured in the external ear canal. The generation of OAEs is found to be linked to the presence of an intact inner ear. Therefore, OAEs are now widely used as a clinical tool for the screening of hearing (for a review, see Probst *et al.*, 1991). Because the generation site of the OAEs and the location of measurement do not coincide, attention has to be paid to the boundary conditions of the OAE measurement.

One factor influencing the OAE measurements is the subject's posture. Measurements of spontaneous OAEs (SOAEs) in standing versus supine position have shown an upward shift in the SOAE frequency (Bell, 1992). The postural dependence of evoked OAEs (EOAEs) has also been studied and has shown more complex changes (e.g., Antonelli and Grandori, 1986; Büki *et al.*, 1996; Ferguson *et al.*, 1998). In addition, postural changes have been shown to affect the audiogram fine structure, which is thought to be closely related to OAEs (Wilson, 1980).

Postural changes affect the intracranial pressure (ICP) in a regular manner (Chapman *et al.*, 1990). Because the intracranial space and the cochlea are connected by the cochlear aqueduct, the pressures in the intracranial and the intracochlear fluid compartments are closely related. In fact, several authors have studied ICP changes by means of measurements performed in the ear canal (e.g., Densert *et al.*, 1977;

Magnæs, 1978; Casselbrant, 1979; Marchbanks and Reid, 1990; Büki *et al.*, 1996). Increase of the inner ear fluid pressure is thought to lead to a slight bulge in the cochlear windows, and hence to increase the stiffness of the windows. Therefore, the changes due to the altered fluid pressure are typically described in terms of an impedance change of the middle ear (Büki *et al.*, 1996).

Presumably, the relation between OAE properties and posture is due to changes in the intracochlear pressure. Thus, these OAE measurements may possibly reflect pathological conditions in cochlear fluid regulation, as is hypothesized to be the case in Menière's disease (for a review, see Horner, 1993).

In this report, the behavior of SOAEs is studied before, during, and after postural changes. Part A of the results section describes the dynamics of the changes in the SOAEs resulting from the postural changes (from upright to lying down and vice versa). In particular, time constants of the SOAE changes are determined. These may reflect properties of the inner ear fluid control. Part B discusses the influence of the altered boundary conditions on SOAEs. Stationary properties of the SOAEs in standing and recumbent position are compared. The peaks in the SOAE spectrum are analyzed and characterized by their height, width, and center frequency.

## I. MATERIALS AND METHODS

Spontaneous otoacoustic emissions (SOAEs) were recorded from 14 normal-hearing ears. The ears were screened

<sup>a)</sup>Electronic mail: E.de.Kleine@med.rug.nl

beforehand for the presence of at least one detectable SOAE. An ER-10C microphone system from Etymotic Research was used with 40-dB gain. The microphone was connected to the subject's ear canal with a foam eartip. The microphone signal was filtered and amplified with a Krohn-Hite 3550 filter (0.7–10-kHz band-pass) and a Stanford Research Systems SR560 preamplifier (300-Hz high-pass and 20-dB gain). The microphone system was calibrated in a Zwislocki coupler. All SOAE recordings were performed in a sound-proofed chamber.

For the screening measurement, the microphone signal was analyzed on-line with a Stanford Research Systems SR 760 Spectrum Analyzer. This measurement was performed with the subject sitting comfortably in a chair. After the screening measurement, the subject was positioned on a reclining bed, standing upright. The eartip was inserted in the external auditory canal and the recording of the microphone signal was begun. The microphone signal was stored continuously on a Denon DAT recorder with a 48-kHz sampling rate. After 1 min, the subject was tilted, within 3 s, to a head-down position, face up (−30 degrees with respect to the horizontal plane). About 2.5 min later, the reverse procedure was carried out (subject standing upright again). After another 2.5 min, the recording was stopped, resulting in a total recording of 6 min. Off-line, the recorded signals were digitally transferred to a computer disk using a Singular Solutions A/D64x connected to a NeXT computer.

Averaged spectra were computed of subsequent segments of the resulting sound file (typically 32 768 points). Time windows overlapping 50% were used to calculate the averaged spectra. A simple level-crossing artifact rejection method was used to exclude windows which were contaminated by (strong) unwanted sounds, such as swallowing and coughing of the subject.

For the averaged spectra, each spectral SOAE peak was modelled with a Lorentzian curve

$$L(f_c, h, w; f) = \frac{h}{1 + 4[(f - f_c)/w]^2} + C, \quad (1)$$

where  $f_c$ ,  $h$ , and  $w$  represent the peak's center frequency, height, and width, and  $C$  is the (local) noise background. Only peaks that exceeded the local background by 3 dB were included in the analyses to be reported. For a spectrum containing multiple SOAE peaks, each peak was fitted separately.

In addition to these spectral parameters, the instantaneous frequency and amplitude of SOAE peaks were calculated using the adaptive least-squares fit time-domain filtering described by Long and Talmadge (1997).

## II. RESULTS

Experiments were performed on 14 ears, each with at least one SOAE peak present. All ears demonstrated an SOAE alteration during the postural changes, although some individual peaks did not. In part A of this section we focus on the dynamics of the SOAE changes, that is, the process of change from one position into another. In part B we discuss

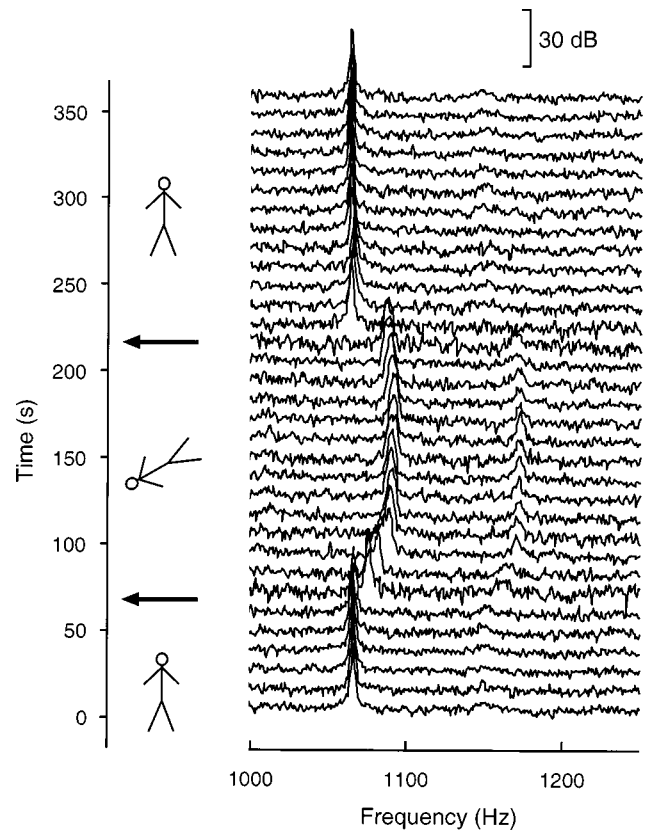


FIG. 1. Waterfall plot of successive SOAE spectra during a 6-min period. The subject's body position is indicated alongside the time axis; the arrows mark the moments at which the subject was tilted. Each spectrum represents the average of a period of 10.9 s and was shifted by 10 dB. Initially, the strongest emission was at 1065 Hz, 15 dB SPL.

the eventual effects on SOAE spectrum after a postural change. In other words, the steady states of the SOAE spectrum in the distinct positions are compared.

### A. Dynamic aspects

Figure 1 displays a typical example of the changes in the SOAE spectrum in one ear during the 6-min period of the experiment. During this period, the posture of the subject was altered twice, as indicated alongside the ordinate. Each individual curve in the waterfall plot represents the average spectrum of the recorded microphone signal of a duration of 10.9 s.

During the first 60 s, with the subject standing upright, the center frequency of the main SOAE peak was 1065 Hz. After tilting the subject backward (at  $t=60$  s) to the recumbent position, the center frequency gradually shifted to 1090 Hz. Furthermore, the width of the peak could be seen to increase, while its height decreased. At  $t=210$  s, the subject was turned upright again. The center frequency returned to 1065 Hz (the same value as for  $t < 60$  s) within one period of averaging and remained at this frequency. The width of the peak decreased again after the upward turn.

Besides the changes in the main peak at 1065 Hz, a second peak appeared after the downward tilt at about 1170 Hz. A more detailed inspection revealed that this smaller peak started at 1150 Hz for the upright position and shifted to 1173 Hz after the downward turn (cf. Fig. 2). The height



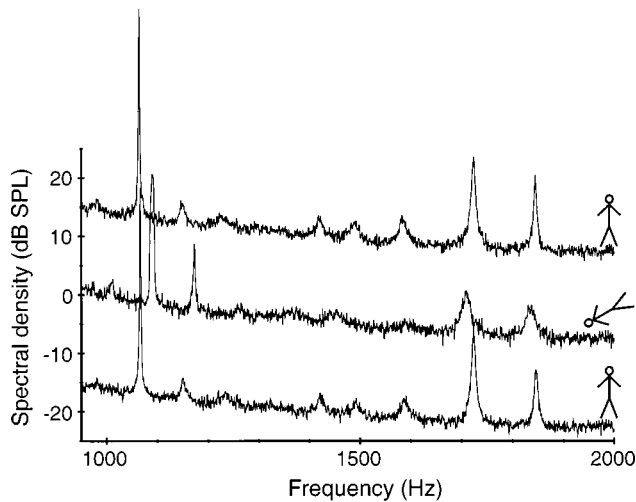


FIG. 2. The SOAE spectra of the final minute of each of the three stages of one experiment in which the spectrum was stable. Lower trace: upright, 0–60 s; middle trace: recumbent, head down 30 degrees, 150–210 s; upper trace: upright, 300–360 s. The upper two traces were shifted by 15 and 30 dB, respectively. Same measurements as Fig. 1.

of this peak increased and its width decreased in the supine position. As a result of a return to the upright position at  $t = 210$  s, the peak height decreased again, and also the center frequency and width of the peak returned to their original values (i.e., the values for the initial upright position,  $t < 60$  s).

Regarding the time course of the changes in this waterfall plot, it should be noted that the time for the spectrum to regain stability after a postural change differed strongly between the downward and upward turn (Fig. 1). After the first postural change (from +90 degrees to -30 degrees,  $t = 60$  s), the frequency shift took approximately 50 s whereas after the second change (from -30 degrees back to +90 degrees,  $t = 210$  s) it took maximally 10 s. In Fig. 3, the time course of the relative frequency shift of four SOAE peaks from one ear is shown (dashed lines; same ear as Fig. 1). Although the direction of frequency shift differed for differ-

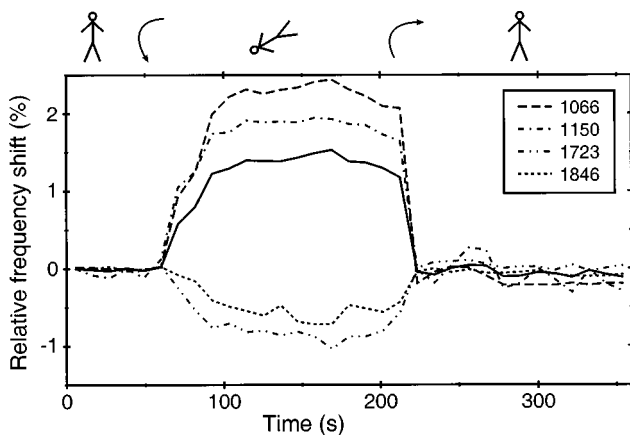


FIG. 3. The relative SOAE frequency shift throughout one experiment. The relative shift of the center frequency of four SOAE peaks from one ear (dashed lines), and the average of the absolute value of these four curves (solid line). The subject's posture was altered at about  $t = 60$  s and  $t = 210$  s, as indicated. The average frequency of the first 60 s was defined as 100%, and is denoted in the legend. Same measurements as Fig. 1.

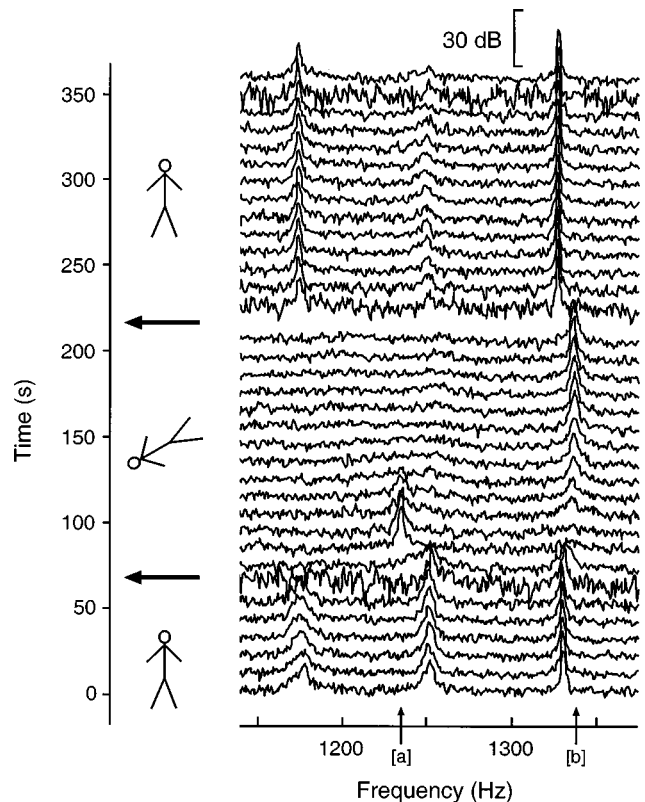


FIG. 4. Waterfall plot, as in Fig. 1. A transient SOAE peak arose 20 s after the downward turn at 1234 Hz, but disappeared again within 1 min (arrow [a]). The peak at 1337 Hz disappeared simultaneously with the appearance of the transient peak (arrow [b]). One trace was omitted; the corresponding spectrum was excessively contaminated by noise during the supine-to-upright rotation.

ent SOAEs, all SOAEs shifted in frequency with a similar time course: a slow adjustment after body position was changed to supine, and a rapid readjustment after it was returned upright. The solid line represents the average of the absolute values of the other four lines. The time constant of the average slow modification was 20 s, whereas the time constants for the individual peaks ranged from 15 to 31 s (fitted with  $e^{-t/\tau}$ , where  $t$  denotes time and  $\tau$  is the time constant). Similar time courses were also observed for other ears, especially for strong emission peaks. We did not observe any changes caused by an upward turn taking more than 10 s.

In addition to the appearance of SOAE peaks, we also observed the disappearance of some peaks after the downward turn. This, of course, can be regarded as an amplitude change. These peaks reappeared after the subject was tilted upright. In one case, a transient SOAE peak was observed after the posture change (Fig. 4): about 20 s after the downward postural turn, a new SOAE peak appeared at 1234 Hz (see arrow [a]), though within 50 s, it disappeared again. Simultaneous with the appearance or disappearance of the transient SOAE peak, the peak at 1337 Hz (arrow [b]) disappeared or appeared, respectively, suggesting a strong interaction between these SOAE peaks.

## B. Stationary changes

The previous section described how the changes in the SOAE spectrum occurred within 1 min after tilting the sub-

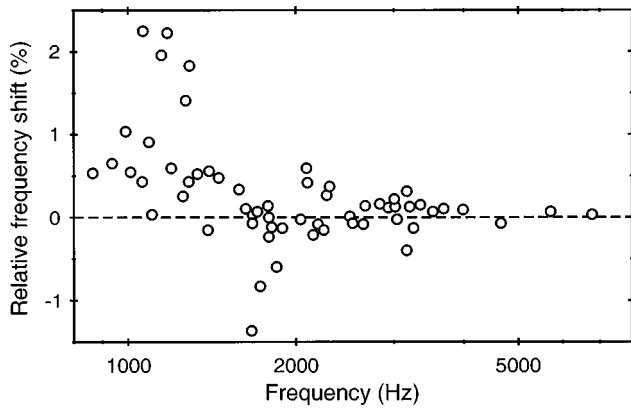


FIG. 5. Frequency shift of SOAE peaks resulting from a postural change (tilt from +90 degrees to -30 degrees). The relative frequency shift  $[(f_2 - f_1)/f_1]$  versus the initial center frequency of the SOAE peak (60 peaks from 13 ears).

ject (see Fig. 1). To make a detailed comparison between the stationary properties corresponding to the different body positions, we calculated the frequency spectra of the final minute of each of the three phases of the experiment (i.e., 0–60 s, 150–210 s, 300–360 s) in which the spectra were stable. Figure 2 shows the average SOAE spectra from one ear for each of the three stationary phases of the experiment. The spectrum for the subject lying down clearly differs from the spectra corresponding to the upright position. The latter two virtually coincide.

In order to quantify these changes, we fitted the major peaks of the three respective spectra with a Lorentzian curve [Eq. (1)]. Thus, three parameters were obtained for all peaks: the peak height, width, and center frequency. Typically, SOAE peaks displayed a reversed frequency shift during the experiment [i.e.,  $\text{sign}(f_2 - f_1) = \text{sign}(f_2 - f_3)$ , where  $f_i$  denotes the average center frequency of part  $i$  of the experiment]. Across all ears, 73 SOAE peaks were monitored. Of these 73 peaks, 6 SOAE peaks did not display a reversed shift, whereas 4 peaks only existed in the upright position, and 3 peaks only in the supine position. For the remaining 60 peaks for which we did observe a reversed shift, the relative frequency shift  $[(f_2 - f_1)/f_1]$  was computed. Figure 5 shows the relative frequency shift of these peaks as a function of their initial center frequency  $f_1$ . Figure 5 demonstrates that a change in posture from upright to supine resulted in positive as well as negative shifts of the center frequency of the SOAE peaks. Even within one ear, both positive and negative shifts could be observed (see Fig. 2). No intrasubject correlations were found. Frequency shifts ranged from -22.8 to +26.2 Hz, with an average of +3.4 Hz. The largest changes in center frequency were observed for SOAE peaks with center frequencies below 2 kHz. Of all 60 peaks, 70% showed a positive frequency shift.

For all of the 60 SOAE peaks considered above, we also calculated the change in the peak height ( $h_2 - h_1$ ) and width ( $w_2 - w_1$ ), where  $h_i$  denotes the peak height in part  $i$  of the experiment and  $w_i$  the peak width. Figure 6 shows the height change versus the width change for all peaks after the postural change. The average change in height equals -2.3 dB. Here 33% of the peaks showed a height increase and 67% a

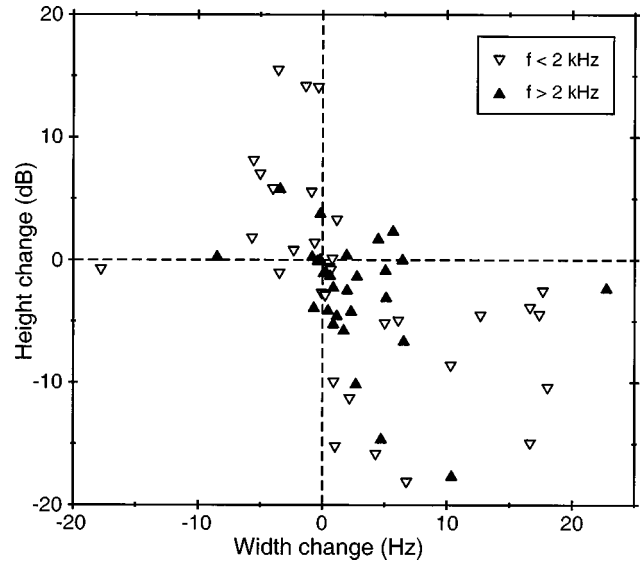


FIG. 6. Change of SOAE peak height and width resulting from a postural change (tilt from +90 degrees to -30 degrees). The height difference ( $h_2 - h_1$ ) versus the width difference ( $w_2 - w_1$ ) for each SOAE peak (60 peaks from 13 ears). The open and closed symbols correspond to low- and high-frequency emissions, respectively.

decrease. The average width change equals +2.7 Hz. Here 67% of the peaks showed a width increase and 33% a decrease. The largest changes occurred for frequencies below 2 kHz. For 50 of the 60 peaks, an increase of peak width went with a decrease of peak height (35), or vice versa, a decrease of peak width went with an increase of peak height (15). No obvious dependence on center frequency was present.

We further examined the nature of the changes in the SOAE peaks by computing a spectrum with a high frequency resolution: a time window of 131 071 points was used, yielding a frequency resolution of 0.18 Hz. Figure 7 displays the spectra obtained from parts 2 and 3 of the experiment (cf. Fig. 2). The SOAE spectrum related to the upright position shows a narrow peak whereas the peak associated with the supine position is much broader (note the 1-Hz side bands in the first spectrum, indicating a 1-Hz modulation). The origin

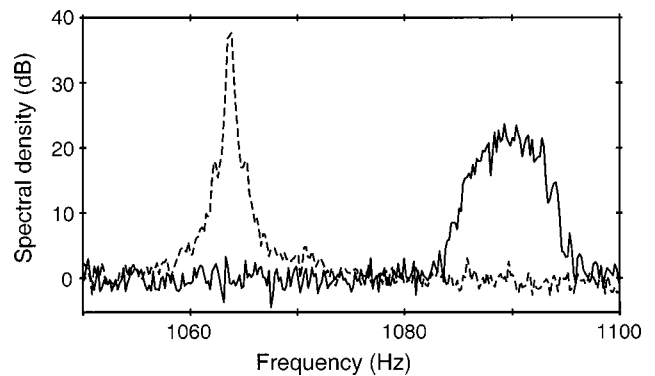


FIG. 7. Detailed spectrum of one SOAE peak (cf. Figs. 1 and 2). The dashed line was calculated from the recording in upright position (+90 degrees), whereas the solid line represents the (shifted) peak corresponding to supine position (-30 degrees). The broadening of the latter peak was caused by an increased frequency fluctuation (see Fig. 8). The peak for upright posture clearly shows side bands at about 1.5 Hz. A time window of 131 072 points (5.5 s) was used, yielding a frequency resolution of 0.18 Hz.

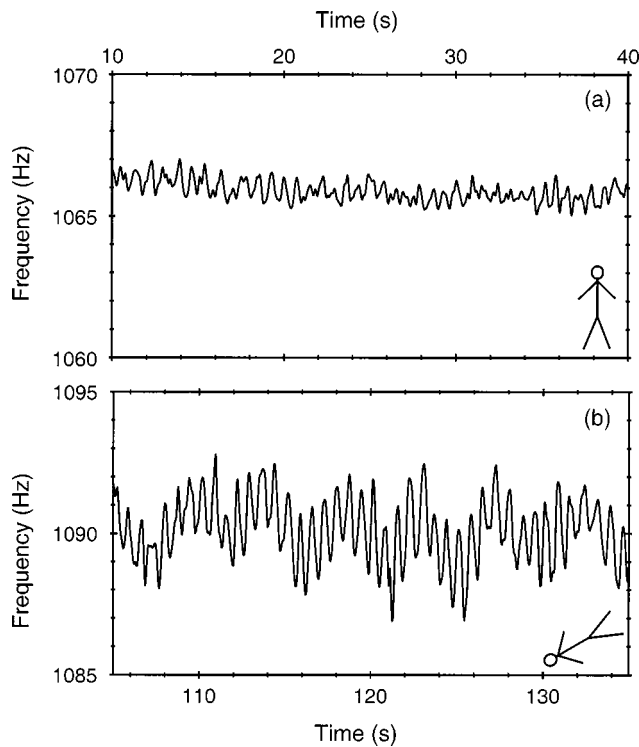


FIG. 8. The instantaneous frequency of one SOAE from one recording, during two separate periods of 30 s. The subject was standing upright in (a), whereas she was in a recumbent position with head down 30 degrees in (b). The tilting was performed at  $t = 60$  s. In panel (b), the frequency fluctuation can be seen to have two components, with a periodicity of about 1 and 5 s. (Same SOAE as Fig. 7.)

of this increase of width becomes clear from Fig. 8, in which the instantaneous frequency of this SOAE peak is plotted. The instantaneous frequency and amplitude were computed using the adaptive least-squares fit time-domain filtering technique as described in Long and Talmadge (1997), with a window length of 0.34 s (computation of the instantaneous frequency from the Hilbert transform gave the same results). The peak's center frequency could be seen to fluctuate during the recording. Substantial differences between the two body positions were present. In the case of the subject standing upright [panel (a)], the frequency fluctuated over approximately 1 Hz with a period of about 1 s (cf. the 1-Hz side bands in Fig. 7). In the supine position [panel (b)], the frequency fluctuated over about 5 Hz with two periodicities: one with a period of about 1 s and one with a period of 4 s. For the supine position, a corresponding fluctuation of the emission amplitude was present: both instantaneous amplitude and frequency increased in parallel [see Fig. 9(a)]. Amplitude modulation could not be detected in the upright position. The Fourier transform of the instantaneous SOAE frequency [Fig. 9(b)] clearly shows the two peaks of the modulation at 0.2 and 1.2 Hz. These data show that both the amplitude and frequency modulation were stronger in the recumbent body position than in the upright position.

### III. DISCUSSION

Posture affects hearing in various ways. We observed that posture changes caused alterations in the SOAE spectrum. Emission peaks showed both upward and downward

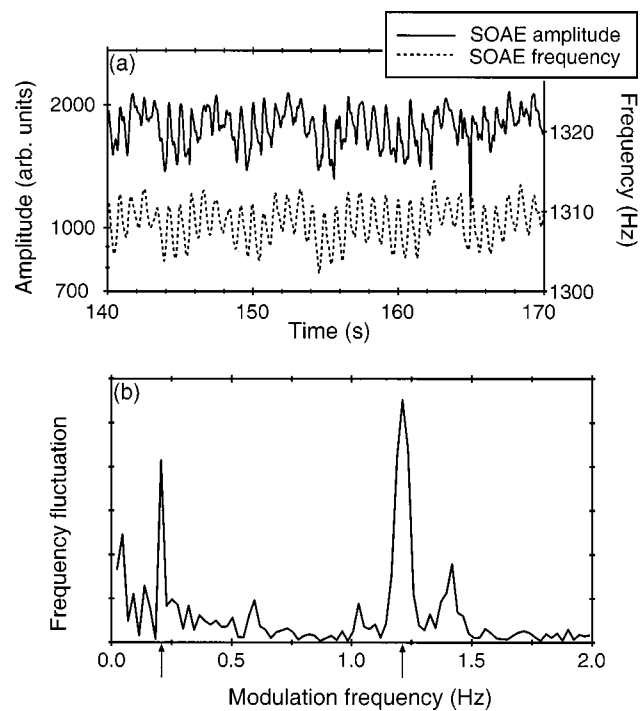


FIG. 9. (a) The instantaneous amplitude and frequency of one SOAE, during a 30-s period with the subject in supine position ( $-30$  degrees). (b) Fourier transform of the instantaneous frequency from (a), showing two peaks at 0.2 and 1.2 Hz (indicated by the arrows), probably representing modulation by breathing and heartbeat, respectively.

shifts in amplitude, width, and center frequency, and these shifts could be different for different SOAEs in the same ear (see, for example, Fig. 2). Frequency shifts in SOAEs have also been studied by Bell (1992), who showed a rise in frequency for three SOAE peaks after a body tilt from horizontal to  $-30$  degrees (about 5 Hz at 1.5 kHz). The postural dependence of click-evoked OAEs (CEOAEs) was first described in Johnsen and Elberling (1982), and showed no significant differences between sitting and lying down flat. Also, Froelich *et al.* (1994) and Ferguson *et al.* (1998) found no overall effect of posture on OAE amplitude. Antonelli and Grandori (1986) observed decreasing latency and amplitude for click evoked and tone-burst evoked OAEs, resulting from a downward body tilt from sitting to lying ( $-40$  degrees), whereas others found systematic changes in latency only (Phillips and Farrell, 1992; Büki *et al.*, 1996). The experiments presented by Büki *et al.* (1996) showed that posture affected CEOAEs in a fairly systematic way. A postural tilt (from  $+90$  degrees to  $-30$  degrees) effected a phase lead (i.e., shorter latency), mainly for frequency components below 2 kHz. The effect of posture on distortion product OAEs (DPOAEs) consisted of a phase lead and a decrease of level at low frequencies (Büki *et al.*, 2000). Furthermore, for inverted position, the acoustic impedance and the auditory threshold increased (Macrae, 1972), and the audiogram fine structure showed an interchange of the peaks and valleys (Wilson, 1980). Horst and Ritsma (1984), however, found no audiogram differences between sitting and lying horizontally.

Posture is believed to modulate the pressure of the in-

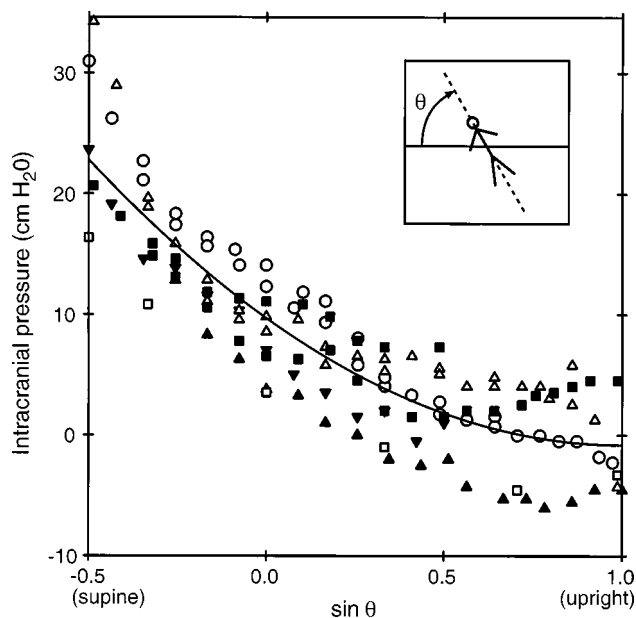


FIG. 10. The intracranial pressure as a function of the body position, from supine ( $-30$  degrees) to upright ( $+90$  degrees); derived from Chapman *et al.* (1990). Different symbols represent different subjects. The solid line denotes the least-squares fit of all points:  $ICP = a(\sin \theta - 1)^2 - b$ , with  $a = 10.5 \text{ cm H}_2\text{O}$ , and  $b = 0.81 \text{ cm H}_2\text{O}$ .

tracochlear fluids. The intracochlear space is connected to the intracranial space via the cochlear aqueduct and several minor pathways (e.g., Carlborg *et al.*, 1982). Because the cochlear aqueduct is connected to the perilymphatic space, the perilymphatic pressure will be altered primarily. However, in normal (guinea pig) ears no difference can be measured between the perilymphatic and the endolymphatic pressure (e.g., Andrews *et al.*, 1991), even after infusion of artificial endolymph (Takeuchi *et al.*, 1991). Thalen *et al.* (1998) showed that modulation of intracranial fluid pressure in the guinea pig gave a corresponding pressure change in the intracochlear fluid.

From observations during neurosurgical procedures we know that posture changes result in a modulation of the intracranial fluid pressure (Magnæs, 1976, 1978; Chapman *et al.*, 1990). Thus, given the connection between the cochlear and cranial fluids, posture influences the intracochlear fluid pressure in a straightforward way.

Figure 10 shows intracranial pressure data from five subjects, taken from Chapman *et al.* (1990). The relation between the ICP (in  $\text{cm H}_2\text{O}$ ) and the angle of rotation  $\theta$  can be described by

$$ICP = 10.5(\sin \theta - 1)^2 - 0.81, \quad (2)$$

and is plotted as the solid line in Fig. 10 (the parameters were determined by a least squares fit). Note that the largest increase of the ICP occurs for  $\theta < 0$  and this is why we compared upright and recumbent position, head down ( $\theta = -30$  degrees). Presumably, the relation between posture and the steady state of the intracochlear fluid pressure shows a great resemblance to Fig. 10.

The exact mechanisms of the effects of the intracochlear fluid pressure on hearing are unknown. Most authors suggest that the increased fluid pressure causes a slight bulge in the

cochlear windows, and hence increases the stiffness of the windows (e.g., Wilson, 1980; Büki *et al.*, 1996). In terms of a mechanical cochlear model, this corresponds to a modification of the cochlear boundary conditions. In agreement with this suggestion, the observed changes for click-evoked OAEs are consistent with predictions from a simple model with altered stiffness of the annular ligament of the stapes (Büki *et al.*, 1996). In general, postural changes are accompanied by a small increase in middle-ear pressure (Büki *et al.*, 1996; Ferguson *et al.*, 1998). Pressure changes of ambient air are known to alter OAEs in a comparable manner (e.g., Hauser *et al.*, 1993). The latter authors, however, used higher pressures ( $\sim 400 \text{ daPa}$ ) than those observed with postural changes ( $\sim 40 \text{ daPa}$ ). Therefore, middle-ear pressure effects alone cannot explain the OAE changes. Since the exact origin of spontaneous OAEs is not fully understood, it is difficult to determine the mechanisms involved in the observed SOAE changes. The OAEs are thought to be generated by linear (coherent) reflection and nonlinear distortion in the cochlea (Shera and Guinan, 1999). The location of the individual SOAE peaks shows close resemblance to the fine structure observed in the audiogram (and evoked OAEs), and probably finds its origin in wave reflections via distributed spatial cochlear inhomogeneities (Talmadge *et al.*, 1998). Considering the alterations of SOAEs, we have to look at the complex way in which the individual peaks in the spectrum interact. The SOAE peaks are influenced by external stimuli (e.g., suppression and synchronization), but also by each other (Burns *et al.*, 1984; Wit, 1990; Long *et al.*, 1993; Murphy *et al.*, 1995; Van Dijk and Wit, 1998). It seems clear that a SOAE peak is not an isolated phenomenon, which is confirmed by our observations (see Figs. 2 and 4). Given the complexity of generation and interactions, an increase of the stiffness of the cochlear windows may well be responsible, at least in part, for the phenomena observed.

Modification of the outer- or middle-ear pressure also induces changes in both amplitude and frequency of OAEs. These pressure changes are realized by manipulation of the atmospheric or closed ear canal pressure, or by Valsalva maneuvers. Generally, for both negative and positive pressure changes, the SOAE amplitude decreases, whereas the center frequency increases (Kemp, 1981; Wilson and Sutton, 1981; Schloth and Zwicker, 1983; Hauser *et al.*, 1993). In some cases, an increase of amplitude or a decrease of frequency has also been detected. Furthermore, the appearance of new SOAE peaks has been observed during pressure changes. The SOAEs are affected most at the lower frequencies ( $< 2 \text{ kHz}$ ). The amplitude of click-evoked OAEs is influenced by both positive and negative ear canal (and atmospheric) pressure in a similar manner (Naeve *et al.*, 1992; Hauser *et al.*, 1993). The amplitude is greatest at ambient pressure and decreases for deviating pressures, again predominantly for the lower frequencies. Wilson (1980) reported an interchange of peaks and valleys in the audiogram after a Valsalva maneuver, similar to that due to a postural change (see above). This interchange was observed for both positive and negative middle-ear pressures, too. In conclusion, direct manipulation of boundary conditions through outer ear pressure yields results comparable to those induced by posture changes. Thus,

it is likely that boundary condition manipulation is an important factor in the phenomena we observed.

We observed strong SOAEs with 1-Hz sidebands (Fig. 7), as previously noted in adults by Bell (1992) and Van Dijk *et al.* (1994). The involvement of heartbeat was suggested in this fluctuation ( $\sim 60$  beats/min = 1 Hz). Long and Talmadge (1997) performed a detailed analysis of SOAEs and found that all strong SOAEs have sidebands at multiples of approximately 1 Hz. They showed the sidebands to stem from frequency modulation correlated to the subject's heartbeat. Emission amplitude modulation was not detected. In Fig. 8 we show that frequency modulation increases substantially after a body tilt from upright to supine posture (head down). Two periodicities can be observed clearly in the latter case. In Fig. 9 the instantaneous amplitude and frequency of another SOAE are plotted for recumbent body position. In this position, amplitude modulation as well as frequency modulation can be detected. Both modulations are in phase; in other words, increasing amplitude goes with increasing frequency. From the Fourier transform of the instantaneous frequency, the two components of the modulation can be clearly discerned: a 0.2- and a 1.2-Hz periodicity, most probably related to breathing and heartbeat. Long and Talmadge (1997) and Ren *et al.* (1995) suggested several mechanisms which could lead to the modulation of emissions by heartbeat (and respiration): (1) acoustic interaction of the pulse beat with SOAEs near the ear canal; (2) modulation of the pressure of the cochlear fluids, which are connected to the cerebrospinal fluid (CSF) through the cochlear aqueduct, with the CSF pressure being known to be modulated by breathing and heartbeat (Marchbanks and Reid, 1990); (3) blood flow through the vessel of the basilar membrane, which could modulate the mass of the basilar membrane; (4) electrical activity in the cochlea generated by heartbeat; (5) modulation of the cochlear pressure by blood flow through the stria vascularis; and (6) modulation of the nonlinear stiffness, a nonphysiological parameter from the model as described by Long and Talmadge (1997). The latter authors conclude that the most likely origin of the heartbeat modulation is a change in the mass of the organ of Corti, probably through blood flow through the vessel of the basilar membrane. In addition to reproducing their measurements, not only have we observed heartbeat modulation, but also modulation by respiration. The CSF pressure is modulated by breathing even more strongly than by heartbeat (Marchbanks and Reid, 1990). However, the transmission of transient pressure changes to the cochlea is doubtful because of the characteristics of the cochlear aqueduct. On the other hand, the inner ear pressure in cats and guinea pigs clearly shows modulation by respiration and heartbeat (Beentjes, 1972; Wit, 1999). Moreover, Beentjes (1972) reported this modulation to be absent after closing the cochlear aqueduct. To summarize, we expect SOAE modulation via the CSF pressure to also play an important role.

We have also recorded the time course of alterations of the SOAE spectrum caused by postural changes. Typical examples of the time course of these alterations are depicted in Figs. 1 and 3. The time for the spectrum to regain stability after a downward postural change is of the order of 1 min

( $\tau = 20$  s), whereas this takes less than 10 s after an upward turn. It is remarkable that the characteristic times of the SOAE changes after the upward and downward change of position differ this much. Note that our subjects were in supine position for only 2.5 min, whereas they were in the upright position for a long time prior to the experiment; this fact may be reflected in the observed time courses. Possibly, the changes would have continued further after 2.5 min in the supine position. Looking at Figs. 1 and 3, however, it seems likely that the main changes occur within the periods we studied. If the observed changes are effected by changes in the inner ear pressure, via an alteration in the intracranial pressure, the characteristic time of these changes will reflect properties of the pressure relation of the two compartments. Therefore, some authors (e.g., Marchbanks and Reid, 1990; Rosingh *et al.*, 1998) have associated this characteristic time with the patency of the cochlear aqueduct (i.e., larger patency yields a smaller time constant). The patency of the cochlear aqueduct was observed to be dependent on the flow direction through the aqueduct, for cats (Beentjes, 1972; Densert *et al.*, 1986) and guinea pigs (Wit *et al.*, 1999). However, the directional dependence of the patency described in these studies differed, and was not as strong as the one we observed. On the other hand, the ICP changes were not instantaneous, but showed a complex kind of adaptation to the altered body position. Magnæs (1978) reported the pressure changes of the ventricular CSF (i.e., ICP) as having a rapid ( $\sim 2$  s) and a slow secondary component (after  $\sim 10$  s) for both sitting up and lying down. Furthermore, our subjects' facial color also showed adaptation: within 10 s after the downward turn, it had become rather red, whereas about 20 s later the color was paler again; after the upward turn no color changes were obvious. This probably reflects the complex response of the venous system on the posture changes (Magnæs, 1978). Thus, the remarkable asymmetry observed for the time course of the SOAE changes is possibly caused (1) by asymmetric properties of the cochlear aqueduct, or (2) by an asymmetric behavior of the ICP changes, or (3) by a combination of these two.

In view of the possible relation with the patency of the cochlear aqueduct, as mentioned earlier, postural experiments have been performed to study this patency and its possible pathologies (e.g., Rosingh *et al.*, 1998). The patency of the cochlear aqueduct has been suggested to be an important factor in intracochlear hydromechanics (Marchbanks and Reid, 1990) and was found to decrease with increasing age (Włodyka, 1978), although this finding could not be confirmed in a recent anatomical study (Gopen *et al.*, 1997). Disturbed patency has also been hypothesized to be related to endolymphatic hydrops (i.e., an increased volume of the scala media), which is assumed to be a primary pathology in Menière's disease (for a review, see Horner, 1993). We think that postural experiments, similar to those described in this article, may be useful in revealing pathological conditions of cochlear fluid regulation.

In conclusion, we observed nonuniform changes in the SOAE spectrum (that is, various alterations of height, width, and center frequency) after a posture change. Altered stiffness of the cochlear windows as a result of altered intraco-

chlear pressure may be responsible, at least partly, for the phenomena observed. In accordance with the model predictions of Büki *et al.* (1996) for evoked OAEs, the major changes occurred for the lower emission frequencies. The remarkable asymmetric behavior of the SOAE changes in time is possibly due to a combination of the patency of the cochlear aqueduct and an asymmetric behavior of the intracranial pressure.

## ACKNOWLEDGMENTS

We thank Carrick Talmadge for providing the adaptive filtering software.

This work was supported by the Netherlands Council for Earth and Life Sciences (ALW), part of the Netherlands Organization for Scientific Research (NWO). This study was also supported by the Heinsius Houbolt Foundation and is part of the research program of our department: Communication through Hearing and Speech. The program is incorporated in the Sensory Systems Group of the Groningen Graduate School for Behavioral and Cognitive Neurosciences (BCN).

Andrews, J. C., Böhmer, A., and Hoffman, L. F. (1991). "The measurement and manipulation of endolymphatic pressure in experimental endolymphatic hydrops," *Laryngoscope* **101**, 661–668.

Antonelli, A., and Grandori, F. (1986). "Long term stability, influence of the head position and modelling considerations for evoked otoacoustic emissions," *Scand. Audiol. Suppl.* **25**, 97–108.

Beentjes, B. I. J. (1972). "The cochlear aqueduct and the pressure of cerebrospinal and endolabyrinthine fluids," *Acta Otolaryngol. (Stockh.)* **73**, 112–120.

Bell, A. (1992). "Circadian and menstrual rhythms in frequency variations of spontaneous otoacoustic emissions from human ears," *Hear. Res.* **58**, 91–100.

Büki, B., Avan, P., Lemaire, J. J., Dordain, M., Chazál, J., and Ribári, O. (1996). "Otoacoustic emissions: a new tool for monitoring intracranial pressure changes through stapes displacements," *Hear. Res.* **94**, 125–139.

Büki, B., Chomiccki, A., Dordain, M., Lemaire, J.-J., Wit, H. P., Chazal, J., and Avan, P. (2000). "Middle ear influence on otoacoustic emissions. II: Contributions of posture and intracranial pressure," *Hear. Res.* **140**, 202–211.

Burns, E. M., Strickland, E. A., Tubis, A., and Jones, K. (1984). "Interactions among spontaneous otoacoustic emissions. I. Distortion products and linked emissions," *Hear. Res.* **16**, 271–278.

Carlborg, B., Densert, B., and Densert, O. (1982). "Functional patency of the cochlear aqueduct," *Ann. Otol. Rhinol. Laryngol.* **91**, 209–215.

Casselbrant, M., (1979). "Indirect determination of variations in the inner ear pressure in man, an experimental study," *Acta Otolaryngol. (Stockh.)* **362**, 3–57.

Chapman, P. H., Cosman, E. R., and Arnold, M. A. (1990). "The relationship between ventricular fluid pressure and body position in normal subjects and subjects with shunts: A telemetric study," *Neurosurgery* **26**, 181–189.

Densert, O., Ivarsson, A., and Pedersen, K. (1977). "The influence of perilymphatic pressure on the displacement of the tympanic membrane," *Acta Otolaryngol. (Stockh.)* **84**, 220–226.

Densert, B., Densert, O., Erlandsson, B., and Sheppard, H. (1986). "Transmission of square wave pressure pulses through the perilymphatic fluid in cats," *Acta Otolaryngol. (Stockh.)* **102**, 186–193.

Ferguson, M. A., Davis, A. C., and Lovell, E. A. (1998). "CSF and posture effects on transient evoked otoacoustic emissions—pilot study," in *Intracranial and Inner Ear Physiology and Pathophysiology*, edited by A. Reid, R. J. Marchbanks, and A. Ernst (Whurr, London), pp. 35–43.

Froelich, P., Ferber, C., Remond, J., Jaboulay, J. M., Duclaux, R., and Collet, L. (1994). "Lack of association between transiently evoked otoacoustic emission amplitude and experimentation linked-factors (repeated acoustic stimulation, cerebrospinal fluid pressure, supine and sitting positions, alertness level)," *Hear. Res.* **75**, 184–190.

Gopen, Q., Rosowski, J. J., and Merchant, S. N. (1997). "Anatomy of the normal human aqueduct with functional implications," *Hear. Res.* **107**, 9–22.

Hauser, R., Probst, R., and Harris, F. P. (1993). "Effects of atmospheric pressure variation on spontaneous, transiently evoked, and distortion product otoacoustic emissions in normal human ears," *Hear. Res.* **69**, 133–145.

Horner, K. C. (1993). "Functional changes associated with experimentally induced endolymphatic hydrops," *Hear. Res.* **68**, 1–18.

Horst, J. W., and Ritsma, R. J. (1984). "Audiogram properties of sound-emitting ears," 11th ICA Santa Barbara.

Johnsen, N. J., and Elberling, C. (1982). "Evoked acoustic emissions from the human ear. II. Normative data in young adults and influence of posture," *Scand. Audiol.* **11**, 69–77.

Kemp, D. T. (1981). "Physiologically active cochlear micromechanics—one source of tinnitus," in *Tinnitus*, edited by D. Evered and G. Lawrenson (Pitman, London), pp. 54–81.

Long, G. R., and Talmadge, C. L. (1997). "Spontaneous otoacoustic emission frequency is modulated by heartbeat," *J. Acoust. Soc. Am.* **102**, 2831–2848.

Long, G. R., Sun, C., and Talmadge, C. L. (1993). "Interactions between spontaneous emissions and external tones: suppression, frequency shifting and distortion product generation," in *Biophysics of Hair Cell Sensory Systems*, edited by H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 40–46.

Macrae, J. H. (1972). "Effects of body position on the auditory system," *J. Speech Hear. Res.* **15**, 330–339.

Magnæs, B. (1976). "Body position and cerebrospinal fluid pressure. Part I: Clinical studies on the effect of rapid postural changes," *J. Neurosurg.* **44**, 687–697.

Magnæs, B. (1978). "Movement of cerebrospinal fluid within the craniospinal space when sitting up and lying down," *Surg. Neurol.* **10**, 45–49.

Marchbanks, R. J., and Reid, A. (1990). "Cochlear and cerebrospinal fluid pressure: their inter-relationship and control mechanisms," *Br. J. Audiol.* **24**, 179–187.

Murphy, W. J., Talmadge, C. L., Tubis, A., and Long, G. R. (1995). "Relaxation dynamics of spontaneous otoacoustic emissions perturbed by external tones. II. Suppression of interacting emissions," *J. Acoust. Soc. Am.* **97**, 3711–3720.

Naeve, S. L., Margolis, R. H., Levine, S. C., and Fournier, E. M. (1992). "Effect of ear-canal air pressure on evoked otoacoustic emissions," *J. Acoust. Soc. Am.* **91**, 2091–2095.

Phillips, A. J., and Farrell, G. (1992). "The effect of posture on three objective audiological measures," *Br. J. Audiol.* **26**, 339–345.

Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (1991). "A review of otoacoustic emissions," *J. Acoust. Soc. Am.* **89**, 2027–2067.

Ren, T., Zhang, M., Nuttall, A. L., and Miller, J. M. (1995). "Heart beat modulation of spontaneous otoacoustic emissions in guinea pig," *Acta Otolaryngol. (Stockh.)* **115**, 725–731.

Rosingh, H. J., Wit, H. P., and Albers, F. W. J. (1998). "Perilymphatic pressure dynamics following posture change in patients with Menière's disease and in normal hearing subjects," *Acta Otolaryngol. (Stockh.)* **118**, 1–5.

Schloth, E., and Zwicker, E. (1983). "Mechanical and acoustical influences on spontaneous oto-acoustic emissions," *Hear. Res.* **11**, 285–293.

Shera, C. A., and Guinan, J. J. (1999). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," *J. Acoust. Soc. Am.* **105**, 782–798.

Takeuchi, S., Takeda, T., and Saito, H. (1991). "Pressure relationship between perilymph and endolymph associated with endolymphatic infusion," *Ann. Otol. Rhinol. Laryngol.* **100**, 244–248.

Talmadge, C. L., Tubis, A., Long, G. R., and Piskorski, P. (1998). "Modeling otoacoustic emission and hearing threshold fine structure," *J. Acoust. Soc. Am.* **104**, 1517–1543.

Thalen, E. O., Wit, H. P., Segenhout, J. M., and Albers, F. W. J. (1998). "Dynamics of inner ear pressure change caused by cerebrospinal fluid pressure manipulation in guinea pigs," in *Intracranial and Inner Ear Physiology and Pathophysiology*, edited by A. Reid, R. J. Marchbanks, and A. Ernst (Whurr, London), pp. 29–34.

- Van Dijk, P., and Wit, H. P. (1998). "Correlated amplitude fluctuations of spontaneous otoacoustic emissions," *J. Acoust. Soc. Am.* **104**, 336–343.
- Van Dijk, P., Wit, H. P., Tubis, A., Talmadge, C. L., and Long, G. R. (1994). "Correlation between amplitude and frequency fluctuations of spontaneous otoacoustic emissions," *J. Acoust. Soc. Am.* **96**, 163–169.
- Wilson, J. P. (1980). "Evidence for a cochlear origin for acoustic re-emissions, threshold fine-structure and tonal tinnitus," *Hear. Res.* **2**, 233–252.
- Wilson, J. P., and Sutton, G. J. (1981). "Acoustic correlates of tonal tinnitus," in *Tinnitus*, edited by D. Evered and G. Lawrenson (Pitman, London), pp. 82–107.
- Wit, H. P. (1990). "Spontaneous otoacoustic emission generators behave like coupled oscillators," in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. Geisler, J. Matthews, M. Ruggero, and C. Steele (Springer-Verlag, Berlin), pp. 259–266.
- Wit, H. P. (1999). Unpublished work.
- Wit, H. P., Thalen, E. O., and Albers, F. W. J. (1999). "Dynamics of inner ear pressure release, measured with a double-barreled micropipette in the guinea pig," *Hear. Res.* **132**, 131–139.
- Włodyka, J. (1978). "Studies on cochlear aqueduct patency," *Ann. Otol. Rhinol. Laryngol.* **87**, 22–28.

# Study of mechanical motions in the basal region of the chinchilla cochlea

William S. Rhode<sup>a)</sup> and Alberto Recio

*Department of Physiology, University of Wisconsin, Madison, Wisconsin 53706*

(Received 12 December 1999; accepted for publication 28 February 2000)

Measurements from the 1–4-mm basal region of the chinchilla cochlea indicate the basilar membrane in the hook region (12–18 kHz) vibrates essentially as it does more apically, in the 5–9-kHz region. That is, a compressive nonlinearity in the region of the characteristic frequency, amplitude-dependent phase changes, and a gain relative to stapes motion that can attain nearly 10 000 at low levels. The displacement at threshold for auditory-nerve fibers in this region (20 dB SPL) was  $\sim 2$  nm. Measurements were made at several locations in individual animals in the longitudinal and radial directions. The results indicate that there is little variability in the phase of motion radially and no indication of higher-order modes of vibration. The data from the longitudinal studies indicate that there is a shift in the location of the maximum with increasing stimulus levels toward the base. The cochlear amplifier extends over a 2–3-mm region around the location of the characteristic frequency. © 2000 Acoustical Society of America. [S0001-4966(00)01506-X]

PACS numbers: 43.64.Kc [LHC]

## INTRODUCTION

Studies of cochlear mechanics have shown there is a compressive nonlinearity in the vibration of the basilar membrane (e.g., Rhode, 1971; Sellick *et al.*, 1982; Robles *et al.*, 1986). These studies have, in large part, avoided measurements in the hook region of the cochlea, with a few exceptions (Khanna and Leonard, 1982; Cooper and Rhode, 1992; Xue *et al.*, 1995; Narayan and Ruggero, 1999). This avoidance appears to be due to the fact that the highest-frequency region of the cochlea is more sensitive to insult and to date has not yielded measurements of comparable sensitivity and nonlinear compression to those just a few millimeters apical to the round window region. Lack of a compressive nonlinearity in the vibration of the basilar membrane is generally thought to imply the cochlea was damaged before measurements were made.

There have also been few systematic observations of the spatial distribution of the vibration of the basilar membrane for a set frequency. An exception is the study by Russell and Nilsen (1997) that showed the spatial pattern of a 15-kHz tone in the guinea pig cochlea. They concluded that there was no movement of the position of the maximum vibration along the cochlear length with increasing stimulus level, even though a compressive nonlinearity was obtained. This result is at odds with the results of noise damage studies of the cochlea, where it is known that the maximum damage occurs at a location tuned  $\sim 0.5$  octave higher than the exposure frequencies (Cody and Johnstone, 1981; McFadden, 1986; Ruggero *et al.*, 1996). In addition, everything known about the displacement transfer curves for a given location in the cochlea indicates that there is a shift to lower frequencies of the maximum in the transfer function as level is increased, a result also found by Nilsen and Russell. This frequency shift should translate into a longitudinal shift of the location of the maximum vibration with level. The shift is attributed

to the compressive nonlinearity in the region of the resonant frequency for a region that only extends in frequency about 0.5 octave below the characteristic frequency (CF).

Another contentious issue is the radial pattern of vibration of the basilar membrane. The basilar membrane is far from a homogeneous mechanical structure and much evidence indicates that outer hair cells (OHCs) induce motion (Brownell *et al.*, 1985). As such, there have been attempts to determine whether there are higher-order modes of vibration in the basilar membrane. Direct measurements in the 15-kHz region of the guinea pig cochlea indicate that there is a 180° phase difference between regions of the basilar membrane (Nilsen and Russell, 1999) supporting the existence of higher-order modes. Some cochlear model and hair cell studies have come to a similar conclusion (Kolston, 1999; de Boer, 1990; Hubbard, 1993; Mountain and Cody, 1999). In contrast, Cooper (1999a) found no higher-order modes of vibration in the 30-kHz region of the guinea pig cochlea. The octave difference in the regions investigated in the two sets of measurements (15- vs 30-kHz tuning) could be a factor in the different outcomes.

The present set of studies was undertaken to address these issues using the basal region of the chinchilla cochlea as the venue. As will be shown, our results indicate that the basal region of the cochlea behaves similarly to the more apical region (9 kHz) that has been extensively studied (e.g., Robles *et al.*, 1986; Ruggero *et al.*, 1997). Further, there is a basal shift in the location of the maximum vibration with level increases. Finally, there are no 180° phase shifts in the radial direction of the basilar membrane over the range of levels that was used in the present series of experiments, suggesting that only a simple beam model of these vibrations of the basilar membrane in the radial direction is adequate.

## I. METHODS

Methods have been previously described in greater detail than here (Cooper and Rhode, 1992). Seventeen chinchilla cochleas were studied at several locations that were

<sup>a)</sup>Electronic mail: rhode@physiology.wisc.edu



within 4 mm of the basal origin of the cochlear partition. All procedures were approved by the Animal Care and Use Committee of the University of Wisconsin-Madison.

Each animal was initially anesthetized using a dose rate of 75 mg/kg of pentobarbital. Additional doses were administered to maintain the animal in a deeply areflexive state. All anesthetics were administered intraperitoneally. A tracheotomy was performed to ensure an open airway. After the ear was surgically removed, four screws were implanted in the skull and cemented in to form a rigid base. A bolt was also cemented to the base to provide for a stable fixation of the skull to a headholder with six degrees of freedom for positioning the cochlea under the microscope objective.

The bulla was opened widely and a silver ball electrode was positioned so as to touch the round window for the purpose of recording the compound action potential (CAP) of the auditory nerve in response to short duration tones (16 ms) for each animal. Since we recorded only in the high-frequency region of the cochlea, the stimulus was stepped in 2-kHz increments from 2 to 20 kHz. At each frequency a visual detection threshold for  $N_1$  was determined by viewing an average of 20 repetitions as the stimulus level was varied in 1-dB steps. If the thresholds were above our best threshold curve by more than 30 dB, no data were collected. High CAP thresholds equated to little or no compression in the hook region (Sellick *et al.*, 1982). CAPs were not typically recorded after mechanical measurements were initiated except to verify that they increased whenever mechanical sensitivity decreased.

In the hook region, surgery involved either cutting the round window membrane with a tungsten microelectrode that had a curved tip, or if a more apical location (3–4 mm from the base) was to be included in the study, the overlying cochlear bone was shaved down using a microchisel until the remaining tissue and/or bone debris could be removed with the microelectrode pick. Gold-coated polystyrene beads 25  $\mu\text{m}$  in diameter served as retroreflectors. They were placed in the perilymph and allowed to sink to the basilar membrane. They have a specific gravity of 1.05 that should minimize any loading of the basilar membrane. In the more apical locations, a glass cover slip was then placed over the opening with no hydromechanical seal. The cover glass served to avoid the problem of an unstable air–fluid interface. When measurements were made in the hook region, a glass cover was laid over the bony cochlea and extended out over the round window. Perilymph wicked up to the glass and the beads could then be imaged without the detrimental effects on the measurements that occur through an unstable interface between the perilymph and air.

An opening in the bony ear canal, immediately over the tympanic membrane, was made so that the probe tube of a 1/2-in. Bruel&Kjaer condenser microphone could be visualized as it was positioned parallel to the tympanic membrane within 1 mm of the tip of the malleus. The opening was sealed with a glass cover after a 45- $\mu\text{m}$  bead was placed on the tympanic membrane at the tip of the malleus (or umbo). The bead was used as a retroreflector for the interferometer and allowed the measurement of the transfer function of the malleus. A probe tube was inserted through the ear canal

along with the sound source and fixed in place with dental cement forming a closed-field acoustic system. The sound source was a reverse-driven condenser microphone cartridge with digital compensation for the nonlinearity resulting in all harmonics being  $< -50$  dB relative to the fundamental for stimulus levels between 80 and 100 dB SPL and  $< -70$  dB *re* the fundamental for levels  $< 80$  dB SPL. The stimuli were presented by a 16-bit D/A at a 200-kHz rate (TDT<sup>®</sup>).

The temperature of the cochlea can have an effect on the operation of the cochlea, especially in the hook region. Cooling the guinea pig cochlea to 36° caused CAP thresholds to increase for frequencies above 24 kHz (Brown *et al.*, 1983). Thresholds of auditory-nerve fibers (ANFs) are usually increased by 20 to 30 dB when the temperature of the cochlea is reduced by  $\sim 5$  °C to 29–32 °C (Ohlemiller and Siegel, 1994). In order to be certain that results were not being compromised by cochlear temperature, it was measured with a microthermister in several experiments. The chinchilla's core temperature was  $\sim 35.6$  °C when measured in a resting (sleep) state. The temperature was manipulated in a few experiments to determine if normal variation in temperature was a cause for concern. While cooling the cochlea down to 31 °C significantly increased the CAP thresholds, temperatures around 35 °C had little effect on CAPs or vibration amplitudes ( $< 2$  dB based on two experiments). Nevertheless, the homeothermic heating blanket used to maintain the chinchilla's core temperature at 37.5 °C was wrapped completely around the chinchilla with an edge positioned near the bulla.

Mechanical responses were measured using a custom-built displacement-sensitive heterodyne laser interferometer (Cooper and Rhode, 1992). The laser was coupled to the preparation using a long working distance lens (Nikon SLWD 5X, NA 0.1). The laser was focused to a spot  $\sim 5$   $\mu\text{m}$  on the reflective beads. The interferometer was not sensitive enough to measure basilar-membrane vibration without the gold-coated beads. Instantaneous phase was measured using two single-cycle phasemeters that worked in quadrature. The phasemeter outputs were sampled at 250 kHz and the phase was unwrapped using software. Response amplitudes were corrected for the frequency response of the recording system. The noise floor was  $< 5$  pm/ $\sqrt{\text{Hz}}$ .

Basilar-membrane input–output (I/O) functions were determined using 30-ms tone bursts with 1-ms raised-cosine rise and fall times and presented every 100 ms. The stimulus level covered a 100-dB SPL range in 5-dB SPL steps. A minimum of eight responses was averaged. Analysis consisted of Fourier decomposition of the steady-state portion of the averaged response at the stimulus frequency. Vibration of the ossicles was measured at the tip of the manubrium (umbo) or at the incudo–stapedial joint or both locations either before or after (sometimes both before and after) the basilar-membrane measurements.

*Abbreviations:* ANF: auditory nerve fibers; dB SPL: decibels *re* 0.0002 dynes/centimeter squared sound-pressure

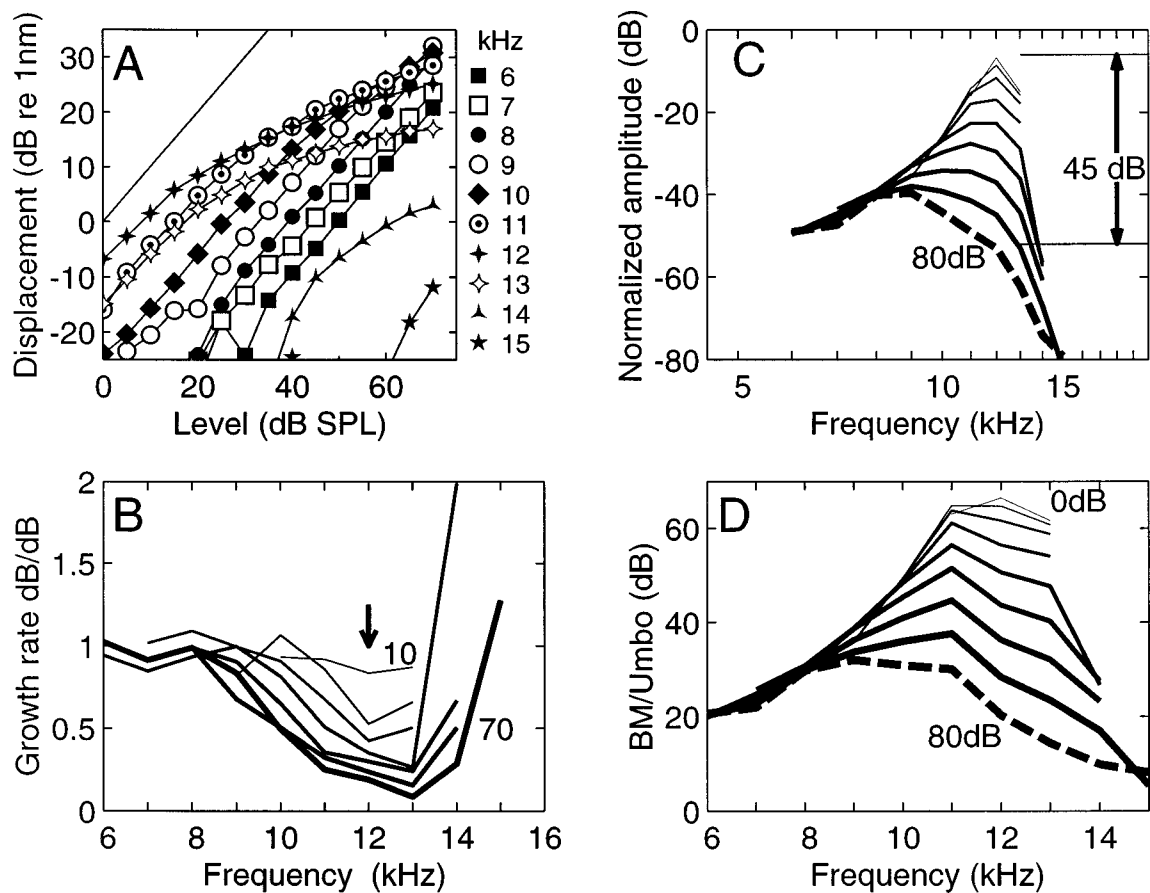


FIG. 1. Data shown are for a location in the region of the round window. (A) I/O functions at the frequencies indicated in the legend. (B) Growth rate is measured as the slope of the I/O functions in panel (A) in dB/dB (dB nm/dB SPL). Slopes of:  $<1$  imply compression,  $=1$  means linear, and  $>1$  expansive nonlinearity. CF=12 kHz is marked by the arrow. (C) The BM transfer curves are normalized by stimulus level from 0 to 80 dB SPL (10-dB increments). Increasing linewidth indicates increasing stimulus level. (D) The basilar membrane to umbo transfer ratios for 0 to 80 dB SPL. Increasing linewidth indicates increasing level where the level step size is 10 dB. The dashed line for 80 dB SPL indicates that the data were extrapolated from the I/O functions using the final slope. Chinchilla cb04.

level;  $f$ : frequency; OHC: outer hair cells; I/O: input/output functions.

## II. RESULTS

### A. Amplitude functions

One of the fundamental properties of vibration of cochlear mechanics is the compressive behavior of basilar-membrane vibration that has been quantified best in the basal region of the cochlea. The compressive nonlinearity in the vicinity of the characteristic frequency is now attributed to the outer hair cells (OHCs). Cochlear mechanical studies to date have been principally in the basal region of chinchilla and guinea pig cochleas, though the most basal portion, often referred to as the hook region, has not been extensively studied (Robles *et al.*, 1986; Sellick *et al.*, 1982; Cooper, 1999a; Nuttall and Dolan, 1996; Khanna and Leonard, 1982). Present studies cover the most basal 4 mm of the cochlea including the hook region ( $5500 \text{ Hz} < \text{CF} < 18000 \text{ Hz}$ ).

Amplitude input/output (I/O) functions are shown in Fig. 1(A) for a location along the basilar membrane with a CF of 12 kHz that was accessed by removing the round window membrane. A displacement of 0.5 nm at 0 dB SPL indicates this cochlea exhibited excellent sensitivity. The deviation of I/O functions from the straight line (upper left-

hand corner) indicates nonlinear compression of the amplitude response. The slope of the I/O functions, dB nm/dB SPL and abbreviated dB/dB, is termed growth rate: 0 dB/dB implies maximum compression (no growth), 1 dB/dB implies linearity or no compression, and values  $>1$  imply an expansive nonlinearity. At CF, some compression remains even at 0 dB SPL. Well below CF the I/O functions are linear, while above CF the compression first increases and then decreases abruptly until it disappears completely, as seen for 15 kHz.

Isoslevel growth-rate functions derived from the I/O functions in Fig. 1(A) illustrate compression is a continuously varying function of both frequency and level [Fig. 1(B)]. At CF (indicated by the arrow), growth rate is 0.85 dB/dB at 10 dB SPL, indicating compression that increases in strength until growth rate reaches 0.2 dB/dB at 70 dB SPL. Maximum compression occurs above CF at 13 kHz. There is a transition region where I/O curves become expansive (slope  $>1$  dB/dB) for 15 kHz at the higher levels and for 14 kHz between 35 and 45 dB SPL. For frequencies  $>\text{CF}$ , the I/O relations begin to become linear as compression decreases. Below CF the compression decreases with frequency until the I/O curves become linear at 8 kHz and remain so for all lower frequencies. The growth-rate func-

tions indicate that there is a “threshold” for the start of compression that continuously increases as the frequency deviates from the region of maximum compression. These general features of the compressive behavior hold across the entire basal region.

The I/O functions are used to construct frequency-transfer functions for the basilar membrane at this cochlear location. As the stimulus level is increased, compression causes the peak frequency to decrease, and at the same time the peak to broaden [Fig. 1(C)]. The maximum shift in the peak is  $\sim 0.5$  octave. The increase in amplitude of vibration over an 80-dB input range at CF is 45 dB less due to compression. If the data were extrapolated to 100 dB SPL, the estimate of the decrease in gain due to compression would be 62 dB. However, it will be shown that this is not entirely justified based on measurements made at the 90–100-dB SPL levels in other experiments.

Frequency-transfer functions at each level were normalized by the vibration of the umbo at the same levels to remove the level contribution to the resulting transfer functions along with the middle ear vs frequency relation [Fig. 1(D)]. For the linear part of the response ( $f < 8$  kHz) all isolevel functions superimpose, while for  $f > 8$  kHz they increasingly deviate from each other. The 0-dB curve shows a maximum gain of 65 dB over umbo vibration. For purposes of comparison with other data in the literature where the basilar membrane to stapes motion is compared (Ruggero *et al.*, 1997), the basilar membrane to malleus ratio has to be increased at least 10 dB to obtain the basilar membrane to stapes ratio; therefore, the gain would be  $> 75$  dB. The broadening of the mechanical filter with increasing level is obvious (note that the 80-dB dashed curve is an extrapolation of the data based on the slope of the I/O curves at 70 dB SPL). For most of these studies, I/O functions were measured at 20 to 40 frequencies for levels that sometimes reached 100 dB SPL. If all of the data from one spot were displayed together, the result would be too complicated, so only a portion of the data are displayed in Fig. 2(A) and (B). The I/O level functions are compressive in the vicinity of CF = 7250 Hz [Fig. 2(A)]. As the frequency increased beyond CF (e.g., 8500 Hz) the compression increased and there is a gradual tendency for the I/O curves to trend towards a linear relation at levels  $> 90$  dB SPL. This was a consistent feature whenever the data were collected at sufficiently high levels. For some frequencies  $> CF$ , a nonmonotonicity is seen in the I/O curves where the amplitude decreases with increasing level. This feature, usually referred to as a notch, was shown previously in the apex where it was concluded that it was due to opening the cochlea (Cooper and Rhode, 1996). This notch in the base appears more complicated in that it may actually be two (or more) notches that occur at different frequencies and levels (see Fig. 11).

The normalized amplitude functions in Fig. 2(B) are similar to those in Fig. 1(C), except for lower gain. In contrast, the growth-rate functions illustrate some additional features [Fig. 2(C)]. First, as the stimulus level is increased to 80–90 dB SPL, compression decreases for much of the frequency range including CF. Second, the I/O curves tend back toward a linear relation. Third, in the region of the I/O

notches, the compression curves become complex. The slope of these curves becomes negative. For some frequencies near the notch frequency the I/O curves become expansive, as seen for the 90-dB SPL level curve that has a growth rate  $> 1$  dB/dB.

The compressive behavior shown in Fig. 2 is a general feature of basilar-membrane vibration present across the range of CFs that was studied. Figure 3 illustrates this for CFs of 5500, 9750, and 14 000 Hz in three separate experiments. The sensitivity varied somewhat across these studies with a significant portion of the variability being due to the placement of the gold bead retroreflector. Since these beads were allowed to sink through the perilymph, there was no control over where they would land or whether they would reflect well once they were in position. Beads near the center of the basilar membrane consistently had the higher sensitivity. The sets of I/O curves shown were selected because they represent the extremes in CFs (5500 Hz was the lowest CF and 14 kHz was near the highest CF) not because they had the best sensitivity or largest compression.

Compression decreased with increasing level for  $f > CF$  where the slopes of the I/O functions are seen to increase for levels  $> 85$  dB SPL [Fig. 3(A)]. This feature is also seen in the growth-rate functions in Fig. 3(C) where the highest level function, indicated by the thickest line, is closer to 1.0 (no compression) than the immediately lower levels. This relation holds across all frequencies. At the highest frequencies the growth rate goes above the 1-dB/dB value, implying expansion. The I/O functions for the 9750-Hz location exhibit the same pattern versus frequency relative to CF as the lower-frequency location [Fig. 3(D)]. The amplitude curves in panel (E) show that the compression at CF is  $\sim 55$  dB over the 100-dB SPL range. Again, growth rate decreases at high levels as seen best in panel (F) where the thickest line (highest level) is closer to 1 dB/dB than the immediately lower levels. In addition, the I/O functions are expansive around 12 kHz. For a location accessed through the round window with a CF = 14 kHz, most of the features of the I/O and growth-rate functions are similar to those in the lower CF region, implying that there is little change in the mechanical vibration of the basilar membrane in the hook region of the cochlea [Fig. 3(G),(H),(I)]. Compression is similar, exhibiting a return to linearity at high frequencies (e.g., 20 kHz) with an expansive nonlinearity at high levels and high frequencies and a  $\sim 0.5$ -octave shift at the highest levels.

## B. Phase functions

Phase of firing of auditory-nerve fibers exhibits a lag/lead relation for frequencies below/above CF with increasing stimulus level and no change at CF (Anderson *et al.*, 1971). This basic relation appears to have a basis in the motion of the basilar membrane (Geisler and Rhode, 1984; Ruggero *et al.*, 1997). The phase data corresponding to the amplitude data in Fig. 1 are shown in Fig. 4 in two forms. Panel (A) shows the phase versus frequency relation at several stimulus levels after it has been corrected for the umbo phase (which also removes the phase associated with the acoustic characteristic of the phone). The dB level is indicated by symbols

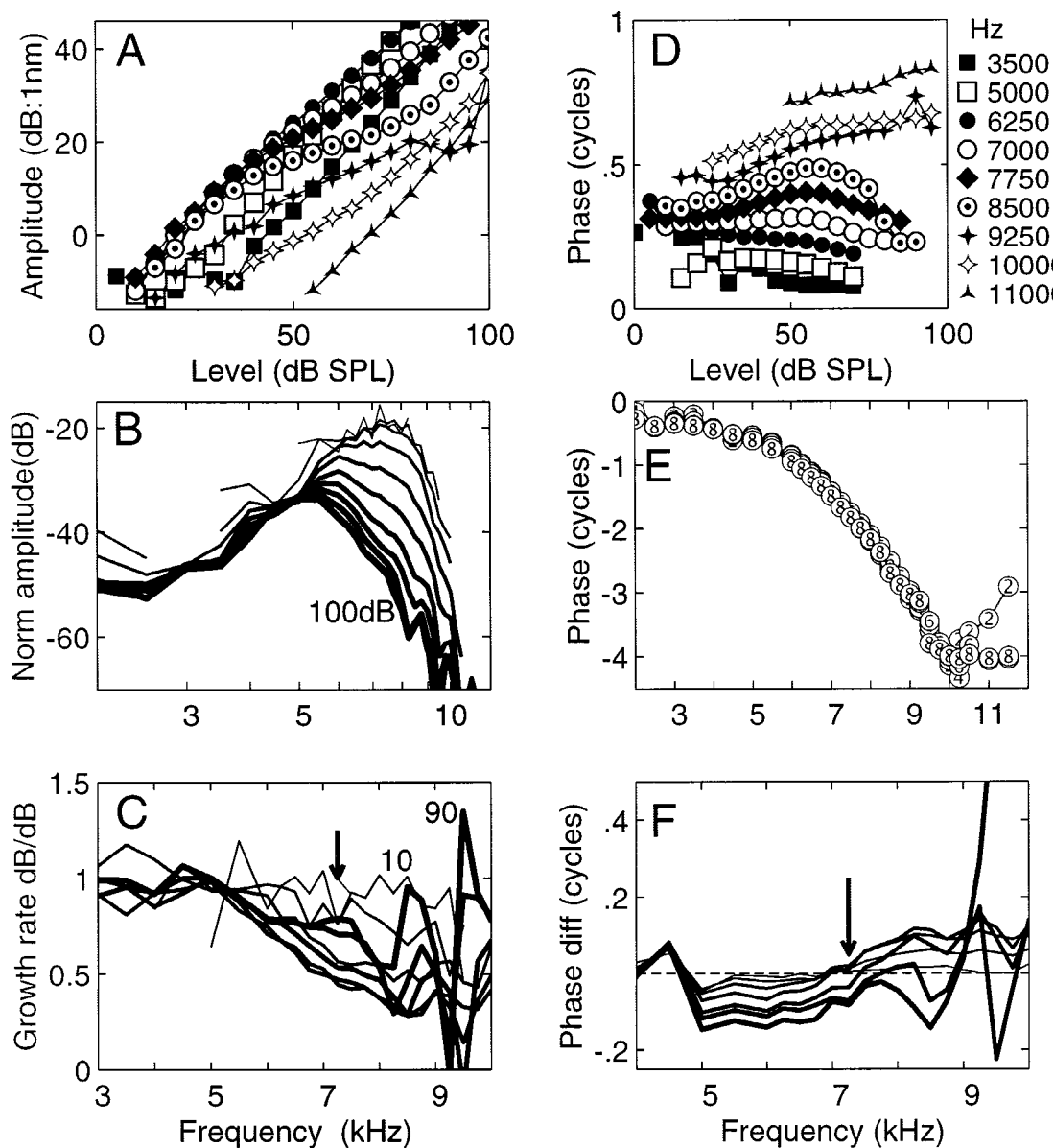


FIG. 2. The transfer functions for the 7.25-kHz region of a chinchilla cochlea. (A) Amplitude I/O functions at the frequencies listed in panel (D). In panels (A) and (D), only every third I/O function is shown. Also, all I/O function for frequencies below 3500 Hz are not shown since they are linear and the phase functions are constant. (B) The normalized amplitude functions based on the I/O functions shown in part (A). Heavier lines indicate the higher levels (0 to 100 dB SPL in 10-dB steps in this panel and all others). (C) The compression curves based on the data in part (A). (D) The phase I/O curves at the indicated frequencies. (E) The phase-transfer curves corresponding to the amplitude transfer curves shown in panel (B). The symbols indicate one-tenth of the stimulus level in dB SPL (a 20-dB increment was used). The fact that all the symbols do not superimpose implies a nonlinear process. (F) The phase difference between the 30-dB SPL phase-transfer curve and higher-level phase data is plotted based on the phase relation shown in panel (E) (cb61).

where the number circled is the first digit of the stimulus level (e.g., ③=30 dB SPL). At CF=12 kHz there is little change in phase with level while below CF the phase decreases (phase lag) and above CF the phase increases (phase lead). At sufficiently low frequencies the I/O function becomes linear, and phase changes with level are absent. In summary, the mechanics appear to underlie the auditory-nerve phase behavior.

Another way of displaying the same data is to plot phase versus level, analogous to amplitude I/O functions [Fig. 4(B)]. At the lowest frequency ( $f=6$  kHz), which had a linear I/O function, there is little phase change with level. As the frequency is increased up to the CF of the region and the I/O functions begin to show compression, the phase lags

with increasing level. At 12 kHz (=CF) there is no phase change between 20 and 40 dB SPL but a phase lag does exist at the highest level. At 13 kHz ( $f>CF$ ) there is an increasing phase lead up to a level of 65 dB SPL. This basic pattern was repeated in all the data collected for which compression was present.

A more complete set of phase data than shown in Fig. 4 consists of data for a wider frequency range and stimulus levels as high as 100 dB SPL [Fig. 2(D),(E)]. Below CF, the phase lags with increasing level. At CF below 50 dB SPL there is virtually no phase change, while above this level there is a phase lag. Above CF there is an increasing phase lead (relative to lower levels) up to 55 dB SPL. For levels >55 dB SPL, there is always a phase lag except at frequen-

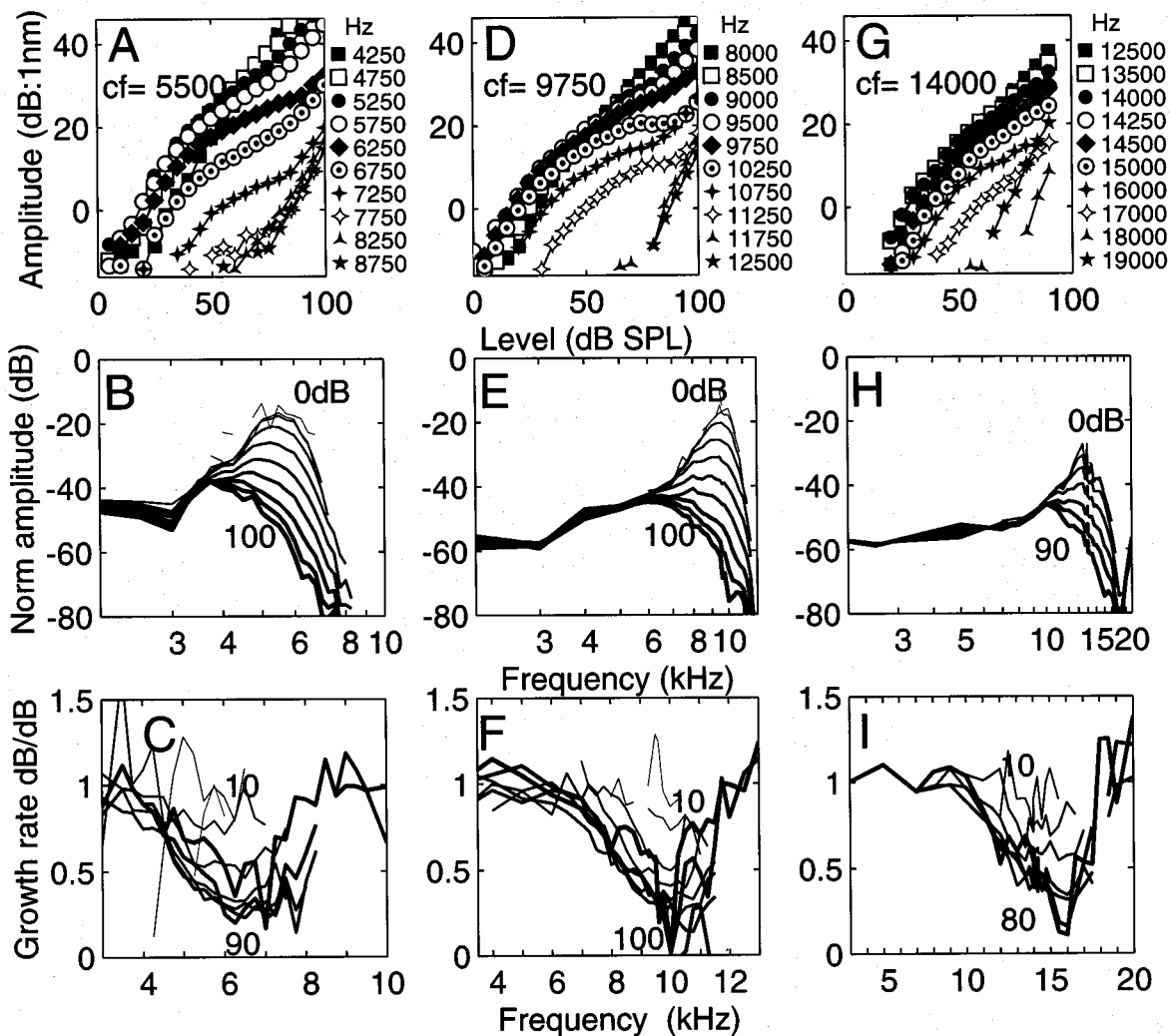


FIG. 3. I/O, transfer, and compression functions are illustrated for three sets of responses. (A)–(C) Data for chinchilla cb58. CF=5.5 kHz. (D)–(F) Data for cb24. CF=9.75 kHz. (G)–(I) Data for chinchilla cb21. CF=14 kHz. In all cases heavier lines correspond to higher stimulus levels. All levels shown are in 10-dB steps. Data were collected in 5-dB steps as shown in top panels. Every other I/O curve is shown with all the low-frequency I/O curves not shown. The labels in panels (B), (C), (E), (F), (H), and (I) are the stimulus levels in dB SPL and are positioned nearest the appropriate curve.

cies corresponding to notches (dips) in the I/O functions. At notch frequencies, phase can be somewhat unpredictable, either increasing or decreasing by a variable amount up to 180°.

An alternate way of displaying the data in Fig. 2(E) is to compute the phase difference between each phase-transfer curve and the 30 dB SPL phase curve; phase versus level and phase versus frequency relations are thereby emphasized [Fig. 2(F)]. The typical phase lag/lead relation below/above CF is present. At the highest stimulus levels (shown with the thickest lines), for  $f > CF$  and levels  $>60$ – $70$ -dB SPL, the phase lags. At a notch frequency of 9.25 kHz, a rapid phase change of  $\sim 0.5$  cycle occurs.

### C. Longitudinal spatial patterns

One aspect of cochlear mechanical studies that has received little attention is the spatial distribution of basilar-membrane motion for individual frequencies. This has been largely due to the difficulty of making the required measurements. A major problem is that making the required large

opening in the cochlea increases the possibility of damage to the cochlea. One study of the spatial distribution of a 15-kHz tone in the base of the guinea pig cochlea is due to Russell and Nielsen (1997). The measurement approach used here differs from theirs in that a complete set of I/O functions was collected at several locations along the basal region of the cochlea. From these I/O functions any spatial measure can be derived, e.g., the amplitude and phase-transfer functions at each location, as is illustrated for seven locations in Fig. 5(A) and (B). The CFs associated with the most basal and apical points are 13 and 6.75 kHz, respectively. Based on Greenwood's formula for the chinchilla, the cochlear distance between these frequencies is 2.46 mm (Greenwood, 1990). The measured distance was 2.3 mm. Based on this agreement, the formula was used to locate the 10-kHz point at 1.25 mm and is indicated by the arrow in Fig. 5(C) (see the caption). The amplitude- and phase-spatial distribution produced by 10-kHz stimuli are shown in Fig. 5(C) and (D). The lowest level in the amplitude curve is for 20 dB SPL and at this level the peak amplitude occurs at 1.25 mm and

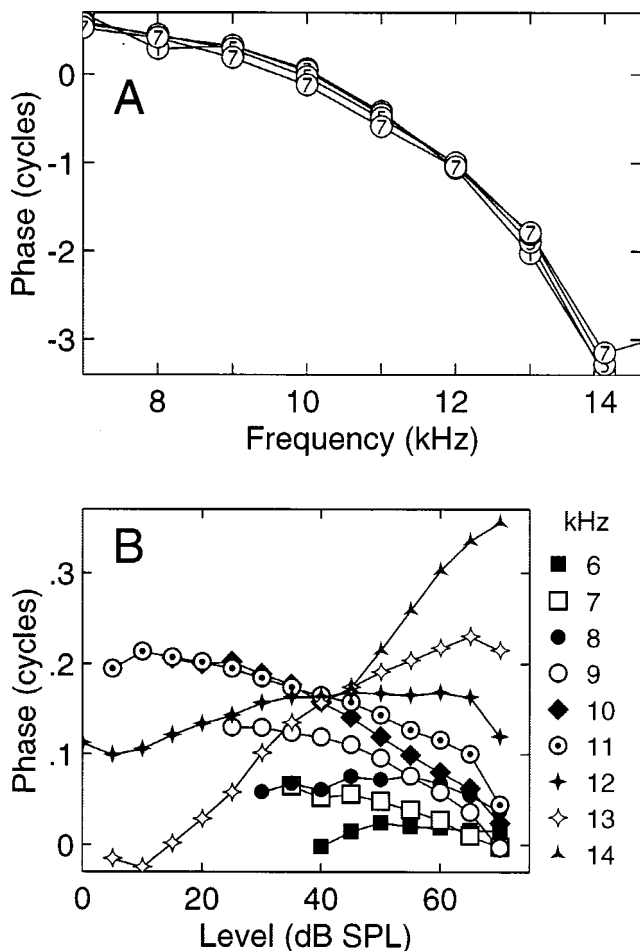


FIG. 4. The phase relations for the transfer- and I/O functions shown in Fig. 1. Level code in panel (A) as in Fig. 2(E). CF=12 kHz.

gradually shifts to the 0.8-mm location with increasing level and towards the 0.2-mm location at the highest level. The frequency shift for this spatial shift is from the 8250-Hz to the 12 500-Hz region (recognizing that the spatial distribution is fairly crudely sampled). This shift is in contrast to the results of Russell and Nilsen (1997), who report that for 15 kHz in the guinea pig that there was no spatial shift of the peak amplitude.

A second spatial study is illustrated in Fig. 6 for five locations spanning 1.2 mm with a CF frequency range between 10.5 and 15.7 kHz. The results are similar to those shown in Fig. 5. In particular, there is a shift in the spatial location for the maximum toward the base with increasing level [Fig. 6(C)]. There is also a corresponding change in the phase relation as a function of increasing level that is apparent since all the phase curves do not superimpose [Fig. 6(D)]. At the highest level there is a phase lag at all locations indicated by the thickest line, while at lower levels there is a phase lag/lead relation for locations basal/apical to the CF location. Three other studies of the longitudinal pattern yielded results consistent with those shown.

Collecting complete frequency transfer functions at several locations permits construction of the spatial distributions produced by pure tones over the region studied. Figure 7 shows the distributions for five frequencies (9–13 kHz) based on the data collected at seven locations (Fig. 5). The

stimulus level was set to 50 dB SPL, a value in the compressive region of the I/O functions. The apical extent of the spatial vibration pattern is seen to move systematically toward the base as the CF is increased in 1-kHz steps. The apical slopes of the amplitude pattern are all similar, with the attenuation rate for the 12-kHz curve equal to  $-47$  dB/mm. Clearly, finer spatial sampling increments are desirable but difficult to achieve when using beads.

These data can also be used to determine a number of other features of the traveling waves. For example, the traveling wave velocity versus location for any frequency can be determined from the phase relations, as is shown for two experiments in Fig. 8. For frequencies low relative to CF, the velocity can be as high as 100 m/s [4.5 kHz at 0.2 mm, off scale in Fig. 8(C)]. For waves due to frequencies corresponding to CF for the region, e.g., 10 kHz in panel (A), the velocity has slowed to  $\sim 8$  m/s. Waves due to frequencies above the CF appear to reach a velocity plateau between 5–6 m/s. As expected, the waves slow down rapidly even before they reach their respective resonance location. An important traveling wave characteristic is its wavelength. As the wave approaches its resonant location it slows down and therefore its wavelength shortens [panels (B) and (D)]. The wavelength corresponding to a 10-kHz tone is  $\sim 0.6$  mm at its CF location. Therefore, wavelength is smaller than the depth of the cochlear scalae in this region, implying that this is a short-wave condition for the traveling wave consistent with prevailing theory (de Boer, 1984). This has implications for models of cochlear mechanics that often employ simplifications based on a long-wave condition.

#### D. Radial relations

Figure 9 illustrates transfer functions for locations that were nearly radial [see the inset of panel (B)]. Point 4 was located  $130 \mu\text{m}$  longitudinally from point 5, which was located  $\sim 80 \mu\text{m}$  from point 9. All the amplitude transfer functions indicate that the locations had similar CFs even though they could have substantially different amplitudes. The vibration at locations that were close to the edges of the cochlear partition was up to 20 dB less in amplitude than those near the center of the partition [Fig. 9(A)]. The phase portions of the transfer functions are all similar, with the most basal location (bead 4) having the smallest phase accumulation at any frequency. Since it is difficult to see the relative phase at each frequency in these curves, a phase-difference curve is computed by subtracting the phase for the fifth location from all others [panel (C)]. The phase differences between location 5 and locations 2, 6, and 8 are relatively small,  $<30^\circ$ . There are somewhat larger phase differences between locations 5 and 9, where up to a  $90^\circ$  phase difference is noted. However, the phase differences can be reconciled based on longitudinal position of the beads alone (e.g.,  $80 \mu\text{m}$  between the longitudinal positions of the 5th and 9th locations). More important, the maximum phase differences for any combination of locations give no indication of  $180^\circ$  variation in phase in the radial direction anywhere along the basilar membrane. For locations near the edges of the partition, the amplitudes are often 20 dB less than those at the center of the partition, and occasionally result in noisier sig-

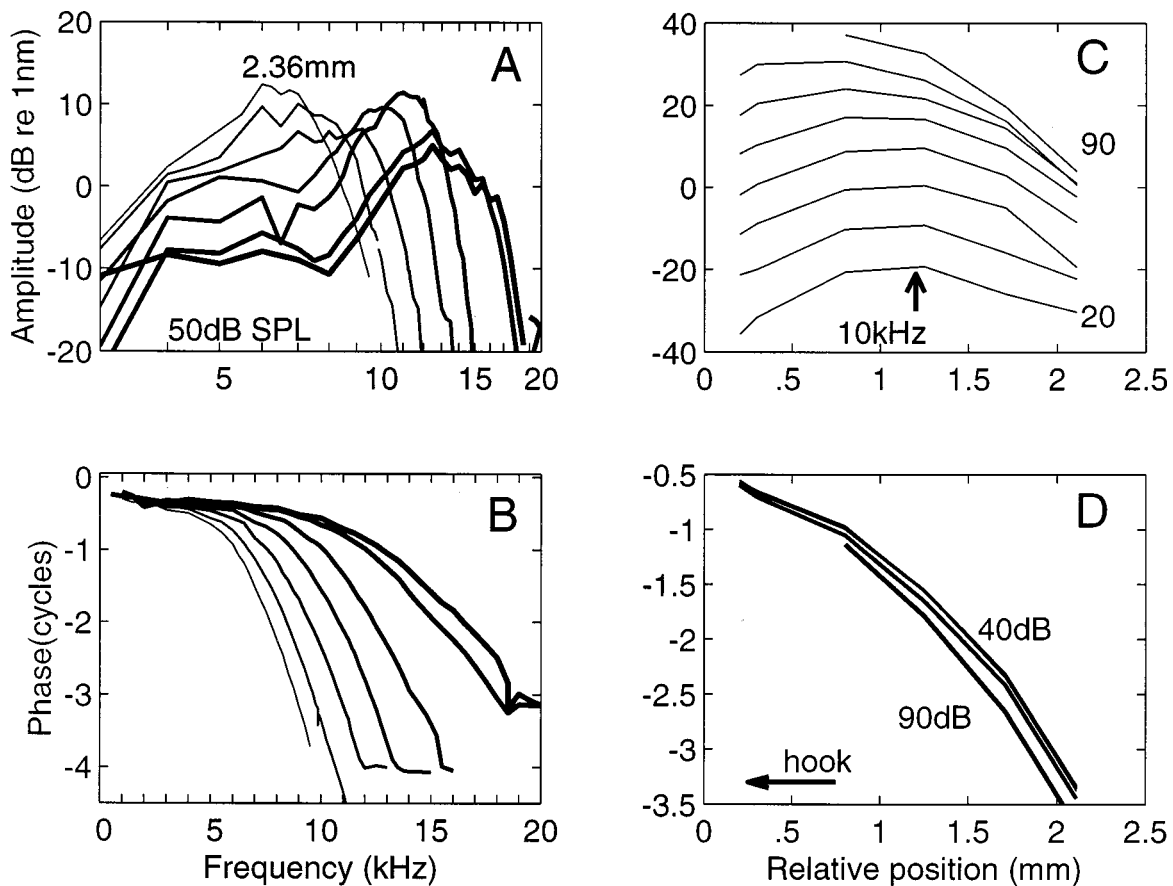


FIG. 5. Data collected in a single experiment at seven locations along the basal cochlea (cb31). CFs and their locations relative to the most basal location of the group were 14 700 Hz/200  $\mu\text{m}$ , 14 300 Hz/300  $\mu\text{m}$ , 12 100 Hz/800  $\mu\text{m}$ , 10 700 Hz/1250  $\mu\text{m}$ , 7900 Hz/1750  $\mu\text{m}$ , 7000 Hz/2110  $\mu\text{m}$ , and 6000 Hz/2360  $\mu\text{m}$  derived from the BM/umbo transfer ratio at 40 dB SPL (cb31). (A) Seven isolevel-transfer functions at 50 dB SPL. The 6000-Hz amplitude function is labeled with its relative location: 2.36 mm. The line thickness is increased with increasing CF and is used the same way in panel (B). (B) Phase portion of the corresponding transfer functions at 70 dB SPL. (C) Spatial distribution of amplitude for 10 kHz derived from the 7 sets of I/O functions at 20 to 90 dB SPL. The arrow indicates the location of CF. (D) Spatial phase-transfer function corresponding to the amplitude data in part (C). Data derived from the phase I/O relations at the levels indicated. The 40-, 50-, 60-, and 70-dB SPL phase functions all superimpose. Thicker lines indicate higher levels.

nals and the possibility that the phase measures are also noisy. The phase curves were computed at 70 dB SPL in order to avoid low-amplitude, possibly noisy data as much as possible. Comparisons with the phase curves at lower stimulus levels showed no essential difference.

Sets of mechanical I/O functions were measured at eight locations spread longitudinally over  $\sim 250 \mu\text{m}$  in another study where the nominal CF was 8.75 kHz [Fig. 10(A)]. The amplitude characteristics reflect the radial position of the beads [inset, Fig. 10(B)]: the beads closest to the strial ligament exhibit the smallest amplitudes.

The phase relations around CF are highlighted in Fig. 10(B) by limiting the range of frequencies. All the phase curves are monotonically decreasing functions of frequency. The phase curves for locations 2 and 7 overlap nearly exactly, which is reassuring since they are for the same bead taken at a 2-h interval. The beads that were nearly radial have the smallest phase differences (beads 6 and 7 and 8 and 9). The longitudinal positions of beads 3 through 8 were 80, 140, 220, 50, 0, 220, and 280  $\mu\text{m}$ , respectively.

For purposes of comparison, the phase characteristic for location 5 was subtracted from all others. The resulting phase differences are relatively small [Fig. 10(C)]. For example, between beads 8 and 9 the phase difference is only a

few degrees at CF, even though they are at different edges of the basilar membrane. Similarly, the phase difference between beads 6 and 7 is less than  $70^\circ$ , an amount that is attributed to a 40- $\mu\text{m}$  difference in their longitudinal position.

The important point is that there is no indication of  $180^\circ$  phase differences between the vibration of the beads as a function of radial position. Based on these two studies and seven others, the basilar membrane appears to vibrate in a first-order mode.

### E. Notches

A dip or notch has been observed in the rate-level function of auditory-nerve fibers at levels  $>90 \text{ dB SPL}$ . A phase change of  $180^\circ$  usually accompanies the notch but is often seen at high levels even if there is no notch. A two-component theory of operation for auditory nerve has been proposed to explain this phenomenon (Liberman and Kiang, 1984; Kiang *et al.*, 1986). Few observations of a notch in mechanical studies have been made including those in the basal region of the cochlea (e.g., Patuzzi *et al.*, 1984; Recio *et al.*, 1998). In contrast, in nearly every transfer curve measured in this study, for which there were low CAF

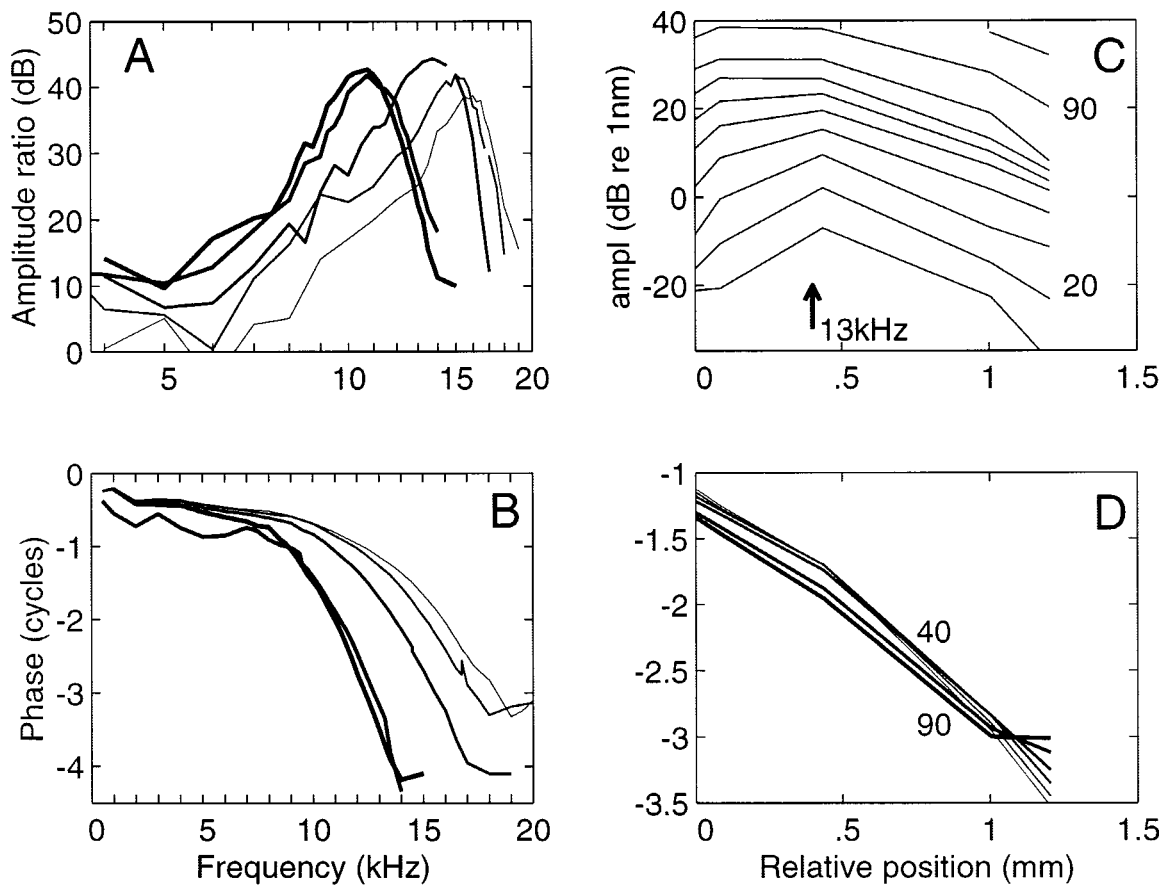


FIG. 6. Data collected in a single experiment at five locations in the basal cochlea. The CFs and their locations relative to the most basal point were 15 700 Hz/0  $\mu\text{m}$ , 14 000 Hz/85  $\mu\text{m}$ , 12 700 Hz/435  $\mu\text{m}$ , 10 800 Hz/1005  $\mu\text{m}$ , and 10 500 Hz/1205  $\mu\text{m}$  (cb030). (A) The five amplitude-transfer functions. (B) The phase portion of the corresponding transfer functions. (C) The amplitude distribution for 13 kHz derived from the five sets of I/O functions at each of the levels indicated. (D) The spatial phase-transfer function corresponding to the amplitude data in part (C). Data derived from the phase I/O relations at the levels indicated.

thresholds and accompanying good compression (implying the cochlear amplifier was working nearly normally), a notch was present.

Figure 11(A) shows a set of isolevel curves that corre-

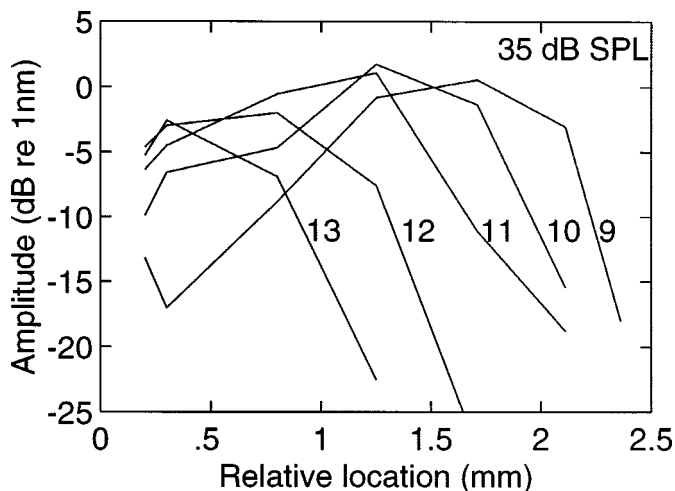


FIG. 7. A set of spatial displacement functions derived from a set of the frequency transfer curves for six locations in a single experiment (cb30). The representation gets poorer as the location moves basally due to fewer locations remaining on which to base the spatial pattern. Stimulus level was 30 dB SPL.

sponds to the 10 200-Hz region of the cochlea. Typical compression at and above the CF with linear growth of the vibration below CF is observed. Above CF, at the frequencies indicated by the arrows (10 700, 12 250, and 14 500 Hz) there are dips in the curves where the displacement-level curve was nonmonotonic. This pattern was always seen whenever there was strong compression. One to three dips were typically observed. A displacement-level I/O function at 11.5 kHz for which a particularly deep notch at 90 dB SPL was seen is shown in panel (B). Note the rate of growth for levels >86 dB SPL is faster than linear; that is, it is expansive. The corresponding phase-level relation shows a phase lead that increases with level and is nearly 90° by 80 dB SPL. As the notch is traversed, a >180° phase change occurs. This behavior is very similar to that reported for frequencies corresponding to notches in ANF rate-level functions. A difference is that only a single notch frequency has been observed in ANFs, never as many as three as seen here. It is possible they exist in ANF responses but have not been seen since the dips are small.

At the minimum of the displacement-level curve, the waveform of the displacement shows that the vibration is “nearly” canceled throughout the stimulus presentation [panel (C)]. Also, a small increase in amplitude is seen between 5 and 30 ms. Exceptions to cancellation are at the



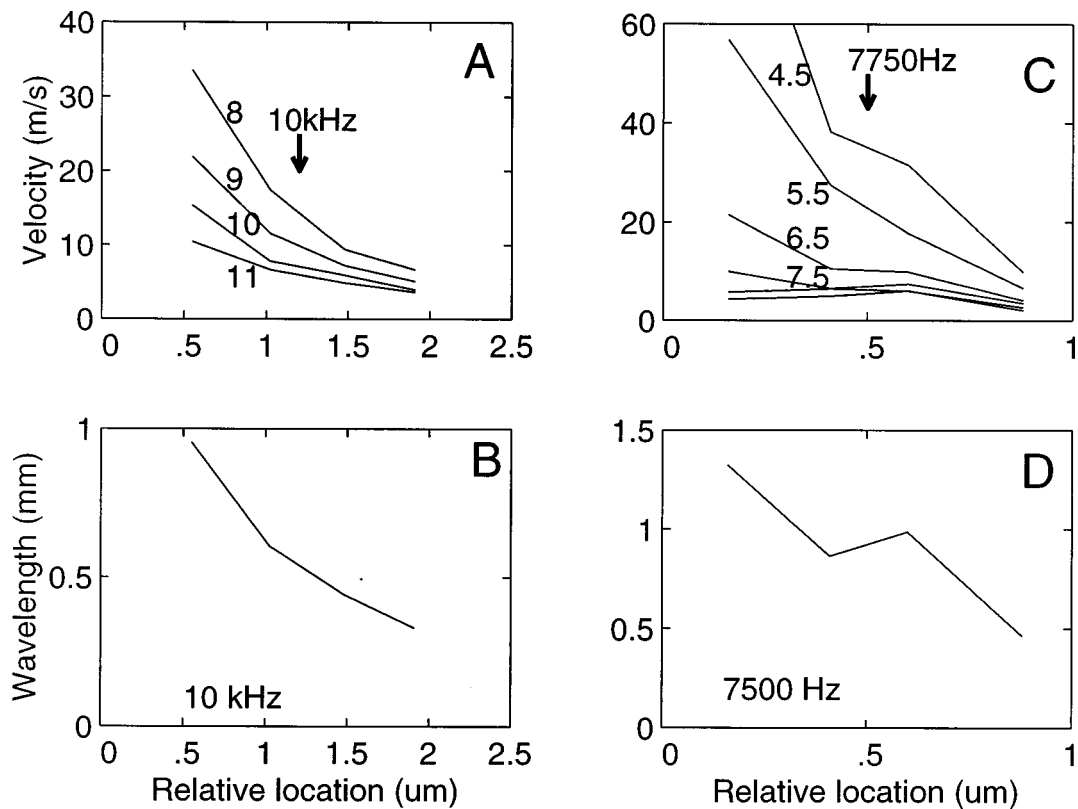


FIG. 8. Two sets of velocity versus location relations derived from data in two chinchillas [(A) cb31, (C) cb30]. The phase data at 70 dB SPL was used to determine the velocity using the relation:  $velocity = (x_2 - x_1) * frequency / (\theta_2 - \theta_1)$  where  $x_1$  and  $x_2$  are distances in meters and  $\theta_2$  and  $\theta_1$  are phases in cycles. The resulting velocities are plotted at the midpoint of the adjacent locations. The label near the individual curves is the frequency in kHz. (B) Illustration of how the wavelength for a 10-kHz tone and a tone varies longitudinally (cb31). (D) Illustration of how the wavelength for a 7.5-kHz tone varies longitudinally (cb30).

onset and offset of the stimulus, where large transients are seen. Expanding the initial portion of the waveform, it is clear there is over a 1-ms delay before the cancellation begins [panel (D)]. The initial portions of the waveforms for two levels at the notch frequency are shown, one below the notch at 80 dB SPL and one above at 100 dB SPL in panel (E). During the initial response ( $t < 1$  ms) the two responses are in phase. By  $t > 1.9$  ms they are  $180^\circ$  out of phase and remain there throughout the stimulus presentation. It appears that there is a delay before the phase shift occurs of over 1 ms from the onset of the response. It is also apparent that there is a latency difference in the response with the high-level response occurring  $\sim 100 \mu\text{s}$  earlier than the lower-level response. It is only after the stimulus is turned off that the two waveforms again have the same phase [panel (F)].

The frequency at which the notch occurred varied appropriately as the CF was varied. In experiments in which transfer functions were measured at several locations, the notch remained on the high-frequency side of the transfer function in approximately the same relative position with respect to the CF. Interestingly, the phase difference between the frequencies of occurrence of the notches in an experiment was usually about  $360^\circ$ . Depth of the notch varied, though this could be attributed to the probability that the best frequency–SPL combination for the notch was not presented. There were no notches seen upon death or in experiments where the cochlear amplifier was not operational, i.e., when the displacement-level functions were linear. In other words, a

viable cochlea was necessary for the occurrence of these notches.

### III. DISCUSSION

Issues that this study addressed were: (1) normal characteristics of vibration in the hook region of the chinchilla cochlea; (2) spatial vibration patterns of the cochlear partition in the longitudinal direction; and (3) radial patterns of vibration of the basilar membrane. Outcomes of some longitudinal and radial studies do not conform in all respects to the existing views and models of cochlear mechanics (Cooper, 1999a; Nilsen and Russell, 1999; Russell and Nilsen, 1997). These differences present a challenge to our understanding of cochlear mechanics that needs to be addressed.

#### A. Basal mechanics

It has proven difficult to measure the vibration of the cochlear partition in the hook region of the cochlea for reasons that are not entirely understood. There is some indication that the cochlear temperature may be a factor in determining sharp, sensitive, and compressive vibration (Ohlemiller and Siegel, 1994). However, in exploring this factor, it did not seem to enter in for the temperatures ( $\sim 35.5^\circ\text{C}$ ) that the cochlea was normally at in these experiments. Others have heated the earpiece that connected the phone to the ear to maintain cochlear temperature and found that there was still relatively poor sensitivity and gain in the

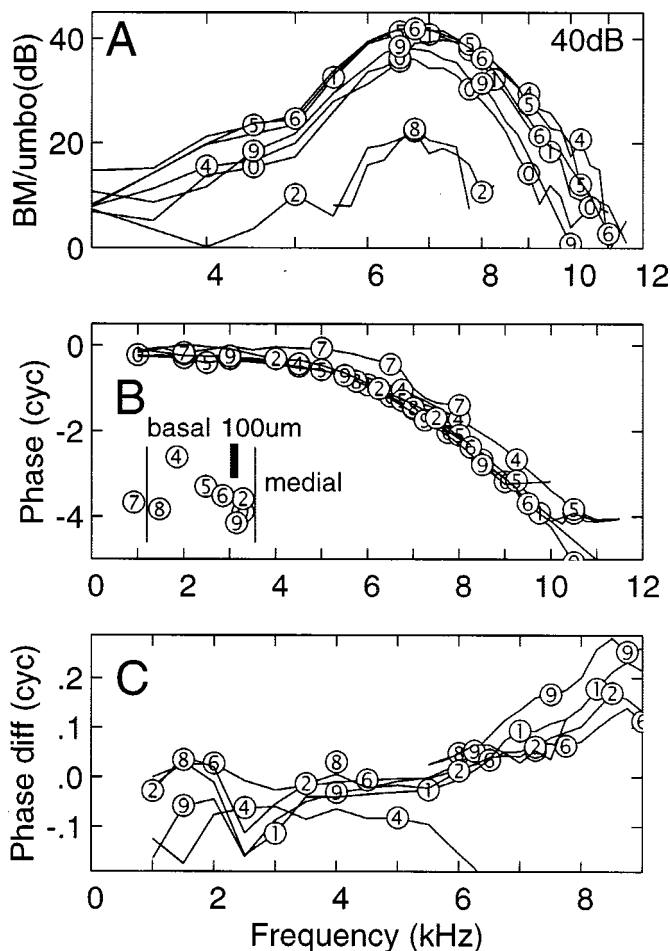


FIG. 9. Study of the vibration of the basilar membrane for several points that were nearly radially located along the cochlear partition. (A) Individual amplitude portions of the mechanical tuning functions obtained in chinchilla, cb61. The amplitude for location 7 was too low for the curve to appear on the graph. (B) Phase portions of the mechanical tuning functions. Included is an inset that shows the relative location of the gold beads used in making the measurements. The scale bar corresponds to 100  $\mu\text{m}$  in the longitudinal direction. The width of the basilar membrane at this location is  $\sim 200 \mu\text{m}$ . The size of the symbols does not correspond to the size of the 25- $\mu\text{m}$  beads. (C) Phase difference relations determined by subtracting the phase relation for the 5th point from the other phase characteristics. Medial: the basilar-membrane attachment medially is the spiral osseus lamina. The vertical lines indicate the lateral and medial extents of the basilar membrane.

hook region (Narayan and Ruggero, 1999). Of course, the reduced sensitivity in the hook region could be the correct result, especially in animals with high initial CAPs corresponding to frequencies for this region. Further, it has been suggested that the cochlear amplifier is composed of a series of OHCs that occupies  $\sim 1.25 \text{ mm}$  around the CF location (Russell and Nilsen, 1997). As the basal extreme of the cochlea is reached, there are no more OHCs basal to the CF site to contribute to cochlear amplification. This reduction of OHCs contributing to cochlear amplification could underlie the rapid increase in the thresholds of high-frequency auditory-nerve fibers observed experimentally. We did not reach the high-frequency limit of the chinchilla cochlea, so this hypothesis was not tested directly.

At least in the 12–14-kHz region of the cochlea, the sensitivity and nonlinear compression appear to be similar to that of the 9-kHz region of the cochlea (see Fig. 1). Mechani-

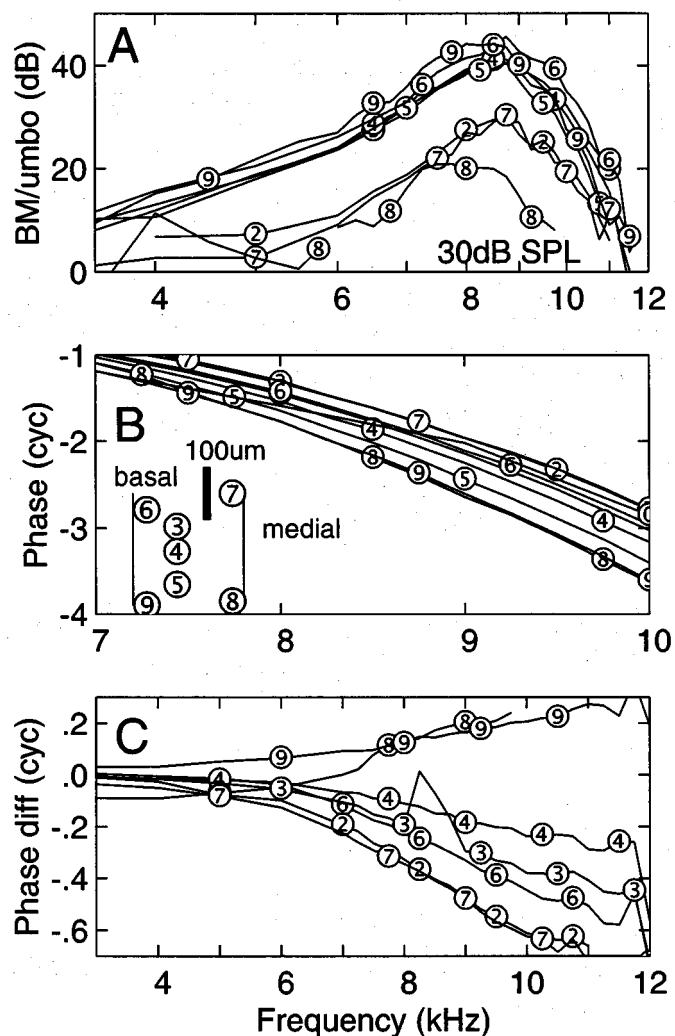


FIG. 10. Vibration of the basilar membrane for several points that were located almost radial along the cochlear partition is illustrated. (A) Amplitude portion of the basilar membrane to umbo transfer function at 40 dB SPL. Location of the beads is illustrated in panel (B). Measurements of bead 7 were a repeat of those on bead 2. The scale bar corresponds to 100  $\mu\text{m}$ . The width of the basilar membrane at this location was 200  $\mu\text{m}$ . (B) Portions of the phase transfer function (BM–umbo) corresponding to the amplitude curves in panel (A). The frequency axis is restricted to the CF region. Inset shows the relative locations of the gold beads used in making the measurements. Measurements of bead 7 were repeat measurements on bead 2. (C) The phase-difference relations determined by subtracting the phase relation for the 5th point from the others. More basal beads will show an accumulated phase lead with increasing frequency and more apical beads will show phase lag (cb17).

cal transfer functions in the two regions are similar, indicating a gradual variation in the mechanical response properties from base to apex. When different locations exhibited good sensitivity, the amplitude characteristic, phase characteristic, and gain were similar across the 5.5- to 15-kHz region of the cochlea.

Amplitude compression was as strong as 0.2 dB/dB at CF and could reach 0 dB/dB above CF similar to previous results (Ruggero *et al.*, 1997). As the stimulus frequency is increased above CF, there is a plateau region of the transfer function where the I/O functions are linear. There is often a notch in the amplitude I/O function for frequencies  $> \text{CF}$  and a corresponding phase change that is a complex function of frequency and level and can be as large as 180°. This behav-

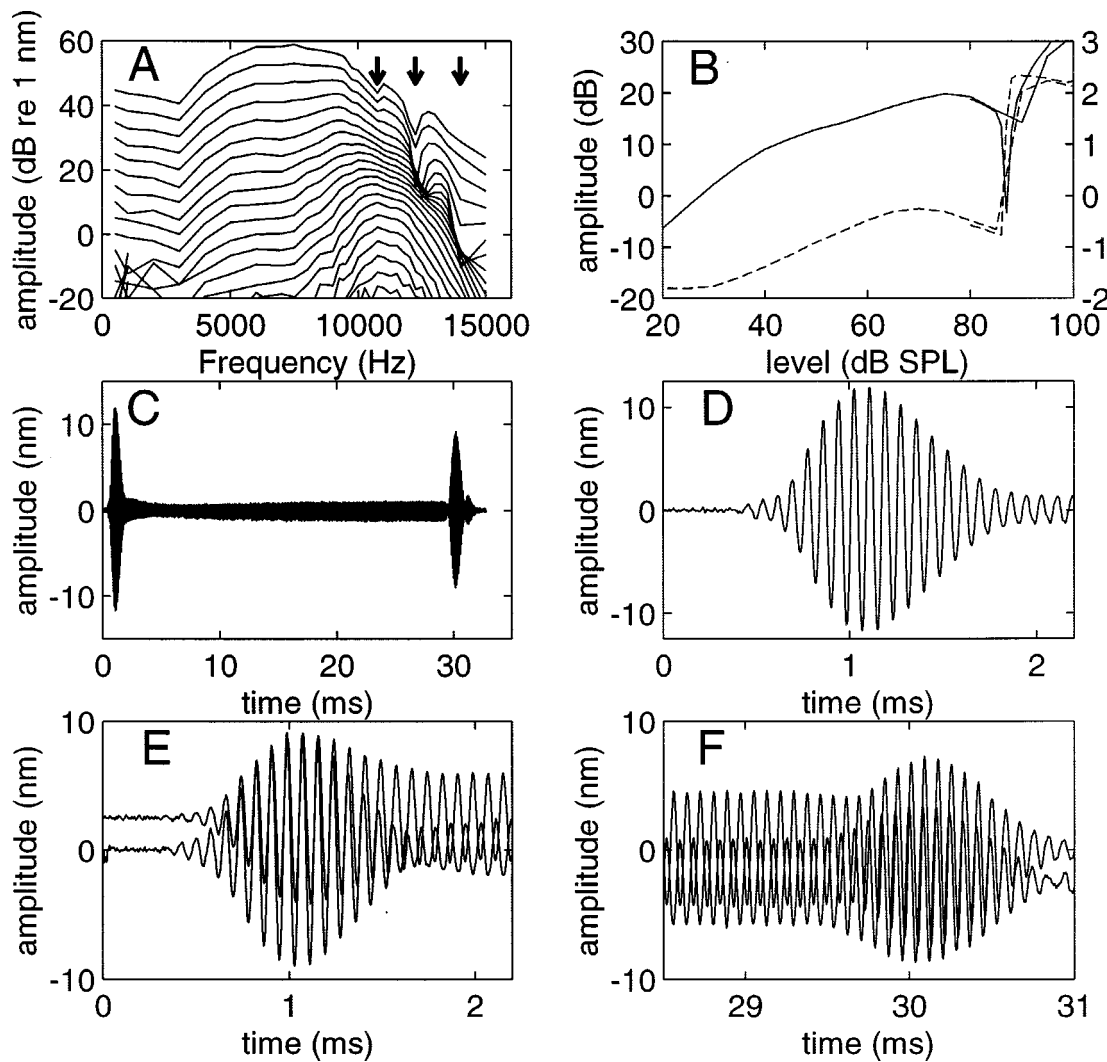


FIG. 11. (A) A set of isolevel amplitude functions for chinchilla cb30. The highest curve is for 100 dB SPL. The curves correspond to stimulus levels of 0 to 100 dB SPL in 10-dB steps. Arrows point to the location of the notches in the curves. (B) The two pair of curves illustrate the notch in a displacement-level curve (solid line) and the phase curve (dashed line) at 12 250 Hz. The function covering 20 to 100 dB SPL in 5-dB steps was repeated in 1-dB steps from 80 to 100 dB SPL 20 min later (the curve with the sharp notch). (C) The waveform corresponding to the lowest point in the notch. (D) The initial portion of the waveforms corresponding to the 80-dB SPL point (dashed line) and the 100-dB SPL point (solid line) for the curve shown in panel (B), (E), (F). The initial [panel (E)] and the final [panel (F)] portions of the displacement waveform for stimulus levels below and above the notch level are shown. The steady-state portions of the two waveforms are 180° out of phase.

ior has been attributed to the fast wave interacting with the traveling wave and results in portions of the I/O functions being expansive. The fast wave is due to a compression wave in the fluid that travels at the speed of sound in water.

The sensitivity of the basilar membrane in this region was shown to be around 0.5 nm at 0 dB SPL and 1.6 nm at 20 dB SPL ( $n=13$ ). The gain basilar-membrane vibration over that of the stapes was  $\sim 79$  dB ( $n=13$ ). These values are more sensitive by around 20 dB than those previously obtained in the hook region (Narayan and Ruggero, 1999).

### B. Phase characteristic

Anderson *et al.* (1971) have shown ANFs exhibit increasing phase lags below CF and increasing phase leads above CF with increasing level. Basilar-membrane mechanics exhibit similar behavior (Rhode and Robles, 1974; Geisler and Rhode, 1982; Cooper and Rhode, 1992; Sellick *et al.*, 1982; Nuttall and Dolan, 1996). Ruggero *et al.* (1997)

demonstrated phase behavior in basilar-membrane mechanics of the 9-kHz region of the chinchilla cochlea compatible with auditory-nerve phase behavior. They showed phase changes as large as 180° over an 80-dB range (20–100 dB SPL) both above and below CF. Similarly, in our most sensitive cochleas, the phase changes are limited to the nonlinear region and can approach 180° at some frequencies [e.g., Fig. 4(B)]. In general, the phase changes were smaller, often  $\pm 45^\circ$  over most of the frequency range and are consistent with the lag/lead behavior below/above CF [Fig. 2(F)]. However, at levels  $>80$  dB SPL, it was often observed that relative phase lags occurred across all frequencies. Phase versus level changes were usually absent in cochleas with low sensitivity, as was also reported by Ruggero *et al.* (1997).

### C. Phase after death. Is there a problem?

Upon death and the subsequent loss of the cochlear amplifier, phase lags the values found in the live cochlea

(Rhode, 1973). In contrast, Nuttall and Dolan (1996) found phase leads of up to  $270^\circ$  for frequencies at and above CF in one guinea pig after death. Similarly, Ruggero *et al.* (1997) found smaller phase leads at CF upon death.

In two comparisons of live/dead phase-transfer characteristics, a phase lag at levels  $>80$  dB SPL across all frequencies was found. There was one condition where the lag/lead change was found, that is, when the phase characteristic was measured immediately upon death (within a few minutes). Presumably the cochlear amplifier does not cease to function instantaneously (cf. Xue *et al.*, 1996). If the time after death was greater than 1 h, the phase always lagged that in the live cochlea. Even in the live cochlea for stimulus levels  $>80$  dB SPL, phase lagged that at lower levels. This is consistent with the cochlear amplifier being saturated at high levels and thereby the mechanism becoming linear. It is likely that phase effects after death are complex and “likely affect several sites and processes in a time-varying manner” as suggested by Ruggero *et al.* (1997).

#### D. Spatial vibration patterns

One of the very difficult aspects of measuring the basilar-membrane vibration along the longitudinal direction is the need to access several millimeters of the cochlea to make the observations interesting. This increases the likelihood that the large opening in the cochlea will result in cochlear impairment. In fact, the basal cochlea has always seemed to be very susceptible to insult for as yet poorly understood reasons. A spatial study of the basal region was performed in guinea pig with a laser-diode interferometer, that works by feeding light reflected from the cochlear partition back into the laser-diode cavity, without the use of a retroreflector, to amplitude modulate the diode output (Russell and Nilsen, 1997). This feedback results in modulating the amount of power out of the diode depending on the distance from the laser and is relatively linear over a 40-nm range. The system is calibrated *in situ*. They usually measured the vibrations for  $\sim 5$  locations in a single study but measured them at many locations at 15 kHz using the footplate of the outer pillar cell as the reference point in one animal. Their results in some respects appear similar to the present results. However, one principal difference is that they conclude that the location of the maximum displacement for a given frequency does not change with level. This is an unexpected result because, if the presumed  $\log(\text{frequency})$  versus cochlear-position relation holds, one would conclude from the shift in the peak frequency observed when the frequency transfer function is measured at one location, that a basal shift in the location of the maximum should be obtained when a single frequency vibration is observed over the spatial domain.

Our results differ from Russell and Nilsen’s on this matter. We performed experiments where several complete transfer functions were measured at up to eight longitudinally distributed locations. Each reconstruction of a spatial frequency distribution based on the I/O functions for individual locations showed that with increasing level, the location of the maximum shifted basally, as was expected theoretically. This shift also agrees with the noise-damage data

that show the location of maximum damage is basal to the CF region of the exposure frequency or center frequency of a noise band (McFadden, 1986). In fact, it is difficult to discern how any other outcome is possible.

The discrepancy between the two outcomes is difficult to explain and we can but list the differences in experimental protocols in search of the causes. Perhaps there was partial damage of the cochlea in the Russell and Nilsen experiments (1997). Their I/O curves appear to be linear up to 40–60 dB SPL, which is substantially higher than the 10–20-dB SPL levels of compression onset in Nuttall and Dolan’s data (1996). In comparing Russell and Nilsen’s experiments to the present experiments, the cochlea was opened in a similar manner in both cases though a different species was used. The use of beads in the present experiments dictates that only a small number of locations can be measured in a single preparation since the location of the beads is not controlled. They fall where they may and they may not reflect well enough to insure a good signal-to-noise ratio, in which case they cannot be used. However, the use of a retroreflector insures that the motion that is reported is that of the bead which is presumed to reflect the motion of the underlying basilar membrane. Without a retroreflector it would be possible for the reflecting surface(s) to switch between various elements of the organ of Corti without knowing that the switch occurred or to reflect from several surfaces with different vibration patterns within the depth of field of the microscope objective. There are multimicron noise movements during the experiment due to breathing and heartbeat-associated movements that could result in switching the reflecting membrane. Further, the signal-to-noise ratio is considerably less than with a retroreflector.

It has been reported that in comparison of measurements of the vibration of the cochlear partition using a bead and a nearby site without a bead that the two transfer functions differed substantially (Khanna *et al.*, 1998). New modes of vibration were created or the bead just did not follow the motion of the basilar membrane. One difference between their use of a bead and ours is that the specific gravity of their glass bead was high ( $\sim 2$ ). The polystyrene beads used here had a specific gravity of 1.05; that is, about the same as water. This was done so that the bead would have a minimum load on the basilar membrane. In modeling the effect of placing a bead on Claudius cells, they state that if the bead had the same density as the fluid, the bead would move with the fluid to a good approximation. In the present study, there was never any indication that the bead was not well behaved in its vibration; that is, it did not appear to introduce any new modes of vibration that would have been manifest by either rapid phase or amplitude fluctuations. Cooper has recently demonstrated that these polystyrene beads follow the motion of the basilar membrane in the guinea pig basal region of the cochlea using an interferometer that does not require the presence of a bead (Cooper, 1999b). Thus, we conclude that the use of retroreflectors does not introduce appreciable errors.

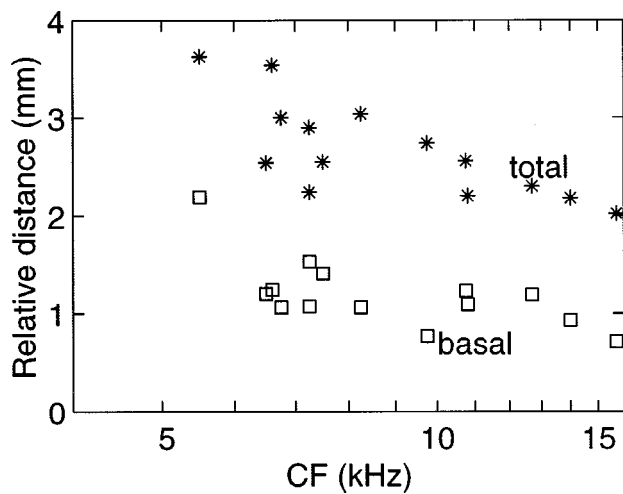


FIG. 12. The longitudinal extent of the compressive region in mm. The distances are determined by determining the intercept of the isolevel growth-rate curves [e.g., Fig. 3(C), (F), and (I)] with the growth rate=1 dB/dB horizontal line. \*'s mark the total and □'s mark the basal compressive region extents, respectively.

### E. Spatial extent of the cochlear amplifier

The spatial extent of the cochlear amplifier is of interest to cochlear modelers. One method of determining the spatial extent of the cochlear amplifier is based on the hypothesis that there is log-frequency versus distance relation in the cochlea. The implication is that the mechanical transfer function for a single location in the cochlea can be transformed to a spatial pattern for a single frequency. The region of compression based on the growth-rate curves is used to determine the spatial extent of the compressive region for a set of 15 experiments with the use of the Greenwood map for chinchilla (Fig. 12). Two sets of data are shown, in the region basal to the CF location (□, basal) and the entire compressive extent (\*, total). The total extent varies fairly systematically with location (CF), decreasing from ~3.5 to 2 mm as the base is approached. The basal portion of the cochlear amplifier varies less along the extent for which measurements are available and is of the order of 1.2 mm. There is an alternative definition of the cochlear amplifier based on power dissipation and modeling the cochlea that derives a value of ~1.75 mm for the extent of the cochlear amplifier and locates it “basal” to the CF location (de Boer and Nuttall, 2000).

The study of Russell and Nilsen (1997) at the 15-kHz location of the guinea pig cochlea suggested the spatial extent was around 1.25 mm. This was taken to imply that around 375 OHCs contributed to cochlear amplification. Cody (1992) recorded auditory-nerve frequency threshold curves (FTCs) after making mechanical lesions in the cochlea. He found normal FTCs at distances from 0.2 to 1.25 mm from the lesion site that implies 60 to 360 OHCs are involved in cochlear amplification. While our estimates are higher than either of these, the extent of the cochlear amplifier is level dependent and this may account for part of the difference between the various estimates. In any case, these data indicate that the cochlear amplifier has a significant longitudinal extent to be considered in any cochlear model.

### F. Are there radial modes of vibration in the basilar membrane?

An important consideration is whether the motion of the cochlear partition along the radial direction is a simple function or whether the known varying properties of the cochlear partition between the *pars arcuata* and *pars pectinata* result in more complex vibration modes (Xue *et al.*, 1996). Nilsen and Russell (1999) found considerable variation in the phase of vibration of the basilar membrane in the radial direction. They argued that OHCs provide the drive near the arcuate-pectinate junction via their active properties. According to this view, OHCs are depolarized and contract as the basilar membrane moves toward the *scala vestibuli* causing the basilar membrane to move further in the same direction, and they are hyperpolarized and expand when the basilar-membrane motion is reversed, thereby driving the basilar membrane further toward the *scala tympani*. However, the pillar cells, Dieter cells, and their connections to the basilar membrane constrain the motion in complex ways. This argument, plus the fact that the bimodal pattern was shown to be absent after the death of the animal, implied that the cochlear amplifier was somehow involved.

By contrast, Cooper's measurements (1999a) did not reveal a complex radial pattern and only a first-order mode of vibration in the 30-kHz region of the guinea pig cochlea. A heterodyne laser interferometer was used and the laser spot was focused to a spot size comparable to that of the Nilsen and Russell interferometer. No beads were used in either case and the motion was sampled at very small spatial intervals in both cases.

The laser-diode technique does require that the system be calibrated each time it is repositioned. In addition, the optical path lengths of the signal and reference arms of the interferometer must be maintained at a multiple of  $\lambda/2$  ( $\lambda$  is the wavelength of the laser) in order to function properly. This is a difficult task due to rapid motion introduced by breathing and heartbeats. In addition the reflectivity of the basilar membrane varies along the radial direction. Whether these factors could produce predictable, and possibly repeatable, variations is not known to us. If there were similar mode changes at lower frequencies one would expect them to be observable in auditory-nerve studies that utilize phase locking of the neural discharges.

Nuttall and Dolan (1996) measured the phase at two nearly radial locations, at the outer pillar cell and the OHCs. No phase differences  $>45^\circ$  were found, a result that is comparable to those reported here. The vibration of the basilar membrane along its radial extent in response to electrical stimulation has been shown to be of uniform phase (Nuttall *et al.*, 1999, see Fig. 4). The phase differences found here increased with frequency at or above CF, yet remained  $<90^\circ$ . Most of the differences are explainable by a lack of perfectly radial alignment of the beads. The 25- $\mu\text{m}$  diameter beads we used are relatively large compared to the 200- $\mu\text{m}$  width of the basilar membrane and therefore might have obscured rapid phase changes by averaging over a radial segment. This could explain missing the out-of-phase motion that was found for a few microns segment of the basilar membrane under the inner hair cells (IHCs) by Nilsen and

Russell (1999). However, this explanation is unlikely in the vicinity of the OHCs where their out-of-phase measurements occurred over a much larger expanse.

The crucial motions are those of the feet of the outer pillar cells, since they effectively support the reticular lamina, and that of the basilar membrane under the outer hair cells. These movements are critical to the drive of the OHC stereocilia. The motion of the basilar membrane (BM) and reticular lamina in an isolated cochlea were shown to be in opposite directions while applying electrical stimulation (Mammano and Ashmore, 1993; Xue *et al.*, 1996). OHCs contracted during depolarization with the reticular lamina moving 5–10 times further than the BM. This implies that the reticular lamina has an additional hinge point and is substantially less stiff than the basilar membrane.

The results shown here indicate that with the use of retroreflectors there is no indication of higher-order modes of vibration in the radial direction. Some phase change is observed that ranges from 30–60°, but in some measurements the phase-transfer function was always very orderly and could be described as well-behaved. The changes that were seen could easily be explained as a result of the beads not being aligned exactly radially.

However, it should be noted that several attempts at modeling the organ of Corti motion, using finite-element approaches to represent the three-dimensional structure, have resulted in higher modes in the vibrations of the simulated basilar membrane. This is a natural result of the attachment of the basilar membrane and the attachment of the pillar cells to the basilar membrane (Zhang *et al.*, 1997; Kolston, 1999). In this respect the Nilsen and Russell observations are consistent with these model results.

## G. Notches

There is a precedent for notches in the rate-level functions of ANFs that were most often clearest for lower CF fibers. The general features were a sharp dip in the rate-level function usually in the vicinity of 90 dB SPL and an accompanying 180° phase change. However, these features do not occur in all ANFs. This behavior is hypothesized to result from the interaction of two components in the cochlear response to sound. The first component is dominant at low levels and is associated with the nonlinear compressive response. The second component grows linearly, starts small, and is 180° out of phase with the first.

Aspects of the mechanical data are in agreement with the auditory-nerve data. The clearest examples of notches were for levels between 85 and 95 dB SPL where the 180° phase shift also occurs. However, the mechanical data have additional features that confound such a straightforward interpretation: (1) there can be more than a single notch, though secondary dips were never very deep; (2) the level at which the notch (dip) occurred varied from 30 to 90 dB SPL; (3) the phase change was not always a jump of 180°, and (4) when there were multiple notches at several frequencies the phase differences between adjacent dips were ~360°. These additional features may result from the use of sensitive preparations and smaller frequency sampling intervals than used in the past. A consideration of why there may be some

areas of disagreement is that often stimulus levels of up to 110 dB SPL were employed to study the second component in ANF responses. Levels above 100 dB SPL were not used since the interferometer would saturate; furthermore, high levels could cause temporary threshold shifts.

There have been descriptions of notches in the mechanical data. Cooper and Rhode (1996) found in the guinea pig apex of the cochlea that a notch could occur that was prominent on the high-frequency side of the mechanical tuning curve. The notch was hypothesized to be the result of the interaction between a fast fluid-conducted wave and the slower basilar-membrane traveling wave. Later, they showed that much of the notch was likely due to the opening that is made in the *scala vestibuli* that provides a shunt for the acoustic pressure wave, thereby resulting in a place where the pressure can leave the cochlea. Upon sealing the cochlea (chinchilla) most, if not all, of the notch disappeared. For some unexplained reason, the notch was more prominent in guinea pigs.

The physical basis of the notch remains an open question, noting that the notch in auditory-nerve fibers rate-response occurs in a closed cochlea. Further, why haven't clear indications of it been seen in high CF auditory-nerve response maps? Perhaps most response maps, either mechanical or neural, have not sampled the stimulus plane in sufficient detail or at high enough levels as done here. One indication of a second component in the absence of a notch is the rapid phase change (~180°) at high stimulus levels that cannot be observed for frequencies above 4–5 kHz due to an absence of phase locking. It is also possible that the notch is labile; that is, the notch is not always seen in every unit and every subject and may shift or disappear with high exposures. In our experience, on occasion, a repeat of the rate-level curve in both nerve and mechanics at the notch frequency showed a change of depth and position of the notch. Nevertheless, it is clear that the notch is dependent on cochlear nonlinearity and compression since it is absent when the I/O curves are linear or the animal is dead.

## IV. SUMMARY

Basilar-membrane mechanics in the base of the chinchilla cochlea covering the 5.5- to 15-kHz frequency range appear to be consistent in a number of characteristics. The basilar membrane has a maximum displacement of 1.6 nm at a sound-pressure level of 20 dB and a maximum gain of ~80 *re* the stapes vibration. It has a compressive nonlinearity in the region of the characteristic frequency as measured by the growth-rate functions, similar in magnitude and extent over the entire frequency range and comparable to previously reported data. A consistent notch in displacement-level functions for several frequencies above the characteristic frequency was found in those cochleas that had strong compression. Notches were never present in damaged or dead cochleas. These notches have a resemblance to those reported in rate-level functions in auditory nerve where a 180° phase shift is often observed near sound-pressure levels of 90 dB. In addition, for frequencies near and immediately higher than the notch frequency, there is an expansive nonlinearity above the levels for which the notch occurs. The

notch and expansive nonlinearity likely result from two processes that interact in both a constructive and destructive manner.

The phase versus level dependency is consistent with auditory-nerve data, i.e., a phase lag/lead with increasing level for frequencies below/above CF. For sound-pressure levels above 80 dB there is only a phase lag that may be due to the saturation of the cochlear amplifier. The spatial pattern of vibration for a tone is consistent with expectations based on the log-frequency relation. That is, as level increases the location of the maximum displacement moves toward the base. The velocity of the traveling wave near the resonance location was shown to slow up to  $\sim 5$  m/s. In the radial direction, it appears that the motion is analogous to that of a simple beam model for a transverse section of the cochlear partition at stimulus levels less than 100 dB SPL.

## ACKNOWLEDGMENTS

This work was supported by the National Institute of Deafness and Communications Disorders, Grant No. R01 DC 01910. Special thanks to C. Daniel Geisler for his careful reading of this manuscript.

Anderson, D. J., Rose, J. E., Hind, J. E., and Brugge, J. F. (1970). "Temporal position of discharges in single auditory nerve fibers within the cycle of a sine-wave stimulus: Frequency and intensity effects," *J. Acoust. Soc. Am.* **49**, 1131–1139.

Brown, M. C., Smith, D. I., and Nuttall, A. L. (1983). "The temperature dependency of neural and hair cell responses evoked by high frequencies," *J. Acoust. Soc. Am.* **73**, 1662–1670.

Brownell, W. E., Bader, C. R., Bertrand, D., and de Ribaupierre, Y. (1985). "Evoked mechanical responses of isolated cochlear outer hair cells," *Science* **227**, 194–196.

Cody, A. R. (1992). "Acoustic lesions in the mammalian cochlea: Implications for the spatial distribution of the active process," **62**, 166–172.

Cody, A. R., and Johnstone, B. M. (1981). "Acoustic trauma: Single neuron basis for the half-octave shift," *J. Acoust. Soc. Am.* **70**, 707–711.

Cooper, N. P. (1999a). "Radial variations of the cochlear partition," in *Recent Developments in Auditory Mechanics*, edited by H. Wada and T. Takasaka (World Scientific, Singapore).

Cooper, N. P. (1999b). "Vibration of beads placed on the basilar membrane in the basal turn of the cochlea," *J. Acoust. Soc. Am.* **106**, L59–L64.

Cooper, N. P., and Rhode, W. S. (1992). "Basilar membrane mechanics in the hook region of cat and guinea-pig cochlea," *Hear. Res.* **63**, 163–190.

Cooper, N. P., and Rhode, W. S. (1996). "Fast traveling waves, slow traveling waves and their interactions in experimental studies of apical cochlear mechanics," *Aud. Neurosci.* **2**, 289–299.

de Boer, E. (1984). "Auditory physics: Physical principals in hearing theory. II," *Phys. Rep.* **105**, 141–226.

de Boer, E. (1990). "Wave-propagation modes and boundary conditions or the Ulfendahl–Flock–Khanna preparation," in *Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele (Springer, New York), pp. 333–339.

de Boer, E., and Nuttall, A. F. (2000). "The mechanical waveform of the basilar membrane. III. Intensity effects," *J. Acoust. Soc. Am.* **107**, 1497–1507.

Geisler, C. D., and Rhode, W. S. (1982). "The phases of basilar membrane vibrations," *J. Acoust. Soc. Am.* **71**, 1201–1203.

Greenwood, D. D. (1990). "A cochlear frequency-position function for several species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.

Hubbard, A. E. (1993). "A traveling-wave amplifier model of the cochlea," *Science* **259**, 68–71.

Khanna, S. M., and Leonard, D. G. B. (1982). "Basilar membrane tuning in the cat cochlea," *Science* **215**, 305–306.

Khanna, S. M., Ulfendahl, M., and Steele, C. R. (1998). "Vibration of reflective beads placed on the basilar membrane," *Hear. Res.* **116**, 71–85.

Kiang, N. Y. S., Liberman, M. C., Sewell, W. F., and Guinan, J. J. (1986). "Single unit clues to cochlear mechanisms," *Hear. Res.* **22**, 171–182.

Kolston, P. J. (1999). "Comparing *in vitro*, *in situ*, and *in vivo* experimental data in a three-dimensional model of mammalian cochlear mechanics," *Proc. Natl. Acad. Sci. USA* **96**, 3676–3681.

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," *Hear. Res.* **16**, 75–90.

Mammano, F., and Ashmore, J. F. (1993). "Reverse transduction measured in the isolated cochlea by laser Michelson interferometry," *Nature (London)* **365**, 838–841.

McFadden, D. (1986). "The curious half-octave shift: Evidence for a basalward migration of the traveling-wave envelope with increasing intensity," in *Basic and Applied Aspects of Noise-induced Hearing Loss*, edited by R. J. Salvi, D. Henderson, R. P. Hamernik, and V. Colletti (Plenum, New York), pp. 295–312.

Mountain, D. C., and Cody, A. R. (1999). "Multiple modes of inner hair cell stimulation," *Hear. Res.* **132**, 1–14.

Narayan, S. S., and Ruggero, M. A. (1999). "Basilar-membrane mechanics at the hook region of the chinchilla cochlea," in *Recent Developments in Auditory Mechanics*, edited by H. Wada and T. Takasaka (World Scientific, Singapore).

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., and Dolan, D. F. (1996). "Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig," *J. Acoust. Soc. Am.* **99**, 1556–1565.

Nuttall, A. L., Guo, M., and Ren, T. (1999). "The radial pattern of basilar membrane motion evoked by electrical stimulation of the cochlea," *Hear. Res.* **131**, 39–46.

Ohlemiller, K. K., and Siegel, J. H. (1994). "Cochlear basal and apical differences reflected in the effects of cooling on responses of single auditory nerve fibers," *Hear. Res.* **80**, 174–190.

Patuzzi, R., Johnstone, B. M., and Sellick, P. M. (1984). "The alteration of the vibration of the basilar membrane produced by a loud sound," *Hear. Res.* **13**, 99–100.

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.

Rhode, W. S. (1971). "Observations of the vibration of the basilar membrane in squirrel monkey using the Mössbauer technique," *J. Acoust. Soc. Am.* **49**, 1218–1231.

Rhode, W. S. (1973). "An investigation of post-mortem cochlear mechanics," in *Basic Mechanisms of Hearing*, edited by A. R. Møller (Academic, New York), pp. 49–63.

Rhode, W. S., and Robles, L. (1974). "Evidence from Mössbauer experiments for nonlinear vibration in the cochlea," *J. Acoust. Soc. Am.* **55**, 588–596.

Robles, L., Ruggero, M. A., and Rich, N. C. (1986). "Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input–output functions, tuning curves, and response phases," *J. Acoust. Soc. Am.* **80**, 1364–1374.

Ruggero, M. A., Rich, N. C., and Recio, A. (1996). "The effect of intense stimulation on basilar-membrane vibration," *Aud. Neurosci.* **2**, 329–345.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (1997). "Basilar-membrane responses to tones at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* **101**, 2151–2163.

Russell, I. J., and Nilsen, K. E. (1997). "The location of the cochlear amplifier: Spatial representation of a single tone on the guinea pig basilar membrane," *Proc. Natl. Acad. Sci. USA* **94**, 2660–2664.

Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (1982). "Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique," *J. Acoust. Soc. Am.* **72**, 131–141.

Xue, S., Mountain, D. C., and Hubbard, A. E. (1995). "Electrically evoked basilar membrane motion," *J. Acoust. Soc. Am.* **97**, 3030–3041.

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.

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.

# Efficient across-frequency integration: Evidence from psychometric functions<sup>a)</sup>

Michelle L. Hicks<sup>b)</sup>

Communication Research Laboratory, Department of Speech-Language Pathology and Audiology (133 FR), Northeastern University, Boston, Massachusetts 02115

Søren Buus

Institute of Hearing, Speech, and Language, and Communications and Digital Signal Processing Center, Department of Electrical and Computer Engineering (442 DA), Northeastern University, Boston, Massachusetts 02115

(Received 10 May 1999; accepted for publication 17 March 2000)

Across-frequency integration of complex signals was investigated by measuring psychometric functions [ $\log(d')$  versus signal level in dB SPL] for detection of brief and long signals presented in broadband noise. The signals were tones at 630, 1600, and 4000 Hz, and a nine-tone complex with components spaced at one-third-octave frequencies between 630 and 4000 Hz. The phase relationship of the components in the complex was varied such that adjacent components were in phase (at 0 degrees), 90, or 180 degrees out of phase. Signal durations (defined in terms of the number of cycles between the half-amplitude points of the Gaussian envelopes) of 4.7 and 150 cycles were tested. Results for six normal-hearing listeners showed that the slopes of the psychometric functions were steeper for the brief than for the long signals, and steeper for the tone complexes than for the tones, particularly for the brief signals. This suggests that the transformation from signal intensity to decision variable may be different for brief complex signals than for tonal signals and long complex signals. Thresholds obtained from the psychometric functions were in excellent agreement with those obtained with an adaptive procedure that employed three interleaved tracks. For the long signals, the threshold improvement for the tone complexes relative to a single tone was well described by a  $5 * \log(n)$  integration rule. However, the threshold improvement for brief signals obeyed a more efficient integration rule of 7 to  $8 * \log(n)$ . A portion of this effect could be accounted for by the phase relationship of the tone complexes; thresholds for brief signals were lowest when the components were in phase at the envelope peak of the signal. This finding indicates that temporal synchrony across auditory channels may enhance detection of brief multi-tone complexes. © 2000 Acoustical Society of America. [S0001-4966(00)05806-9]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Mk, 43.66.Nm [RVS]

## INTRODUCTION

It is generally well known that information in relatively independent frequency channels can be integrated across frequency under certain listening situations. The present study examines some factors that affect the efficiency of this integration by measuring psychometric functions for simple and complex signals. For complex signals composed of equally detectable components that excite independent frequency-selective auditory channels, the optimal decision rule predicts that  $d'$  ( $d$  prime) should increase as the square root of the number of components (Green and Swets, 1966). Because  $d'$  in detection experiments appears to be nearly proportional to signal intensity, the threshold for an  $n$ -component signal should be  $10 * \log(\sqrt{n})$  or  $5 * \log(n)$  dB lower than that for a single component. Whereas this rule is well established for long signals (e.g., Green, 1958; Green

*et al.*, 1959; Buus *et al.*, 1986; van den Brink and Houtgast, 1988, 1990a, b; Higgins and Turner, 1990; Langhans and Kohlraush, 1992), it often fails for short signals. Scholl (1962) reported that duration was a critical parameter for the effect of bandwidth on masked thresholds of noise signals. The interpretation of Scholl's data is difficult owing to the statistical properties of noises, which vary with bandwidth and duration, but more recent data with deterministic signals point in the same direction. For brief multi-tone complexes, thresholds can be  $8 * \log(n)$  dB lower than the thresholds for the individual components (van den Brink and Houtgast, 1988, 1990a, b), resulting in more efficient integration than that predicted by the optimal decision rule and observed for long signals. The effect of duration is difficult to reconcile with some of the models that have been proposed to account for across-frequency integration, such as the independent-thresholds model (Schafer and Gales, 1949; Scholl, 1962) or the multi-band decision model (Green, 1958; de Boer, 1966; Green and Swets, 1966), in which the auditory filter bandwidths and channel weights are invariant over time.

One possible explanation for the more efficient across-frequency integration (i.e., the lower-than-expected thresholds) for very brief complex signals is that psychometric

<sup>a)</sup>A portion of this research was presented at the Association for Research in Otolaryngology Midwinter Meeting, St. Petersburg, FL, February 1999.

<sup>b)</sup>Author to whom correspondence should be addressed; Psychoacoustics Laboratory, Department of Speech and Hearing Science, Arizona State University, P.O. Box 871908, Tempe, AZ 85287-1908. Electronic mail: michelle.hicks@asu.edu



functions for detection of tones under masking may be shallower for brief signals than for long signals. It is well established that psychometric functions for detection of masked long-duration signals closely follow the predictions of energy-detector models and yield a sensitivity,  $d'$ , which is approximately proportional to signal energy (e.g., Green and Swets, 1966; Buus *et al.*, 1986). To our knowledge, only one study of psychometric functions for detection of very brief (<10 ms) masked tones by normal-hearing listeners has been published (Green *et al.*, 1957). However, their results may be confounded by spectral splatter caused by turning the signal on and off abruptly, and the psychometric functions may be artificially steep owing to their use of a fixed signal level throughout a block of 100 trials (cf. Taylor *et al.*, 1983).

The description of the psychometric functions in terms of  $d'$  being proportional to the signal intensity (or signal energy) is only one of several possible alternatives. Because the signal level at threshold for long signals typically is lower than the masker level within the auditory filter, the increase in overall level within an auditory filter caused by the presence of the signal,  $\Delta L$ , is approximately proportional to the signal intensity. Therefore, the psychometric functions for detection may equally well be described as  $d'$  being proportional to  $\Delta L$ . This idea is consistent with findings by Buus and Florentine (1991) and Buus *et al.* (1995) that  $d'$  for detecting an increase in the level of tones or noises is proportional to the level increment,  $\Delta L$  [=  $10 \log(I + \Delta I/I)$ , where  $\Delta I$  is the increase in intensity and  $I$  is the intensity of the standard].

The difference between these two possible descriptions of the psychometric functions for detection is of little consequence when the signal is long, because signal intensity and  $\Delta L$  are nearly proportional to one another at the low signal-to-noise ratios required for detection of a long signal. When the signals are brief, on the other hand, the difference between the two possible forms of the psychometric functions for detection is considerable, because the threshold signal-to-noise ratios are relatively large. When the signal-to-noise ratio is large,  $\Delta L$  grows much more slowly than the signal intensity. Thus,  $d'$  will grow more slowly than the signal intensity, if  $d'$  is proportional to  $\Delta L$  and the signal-to-noise ratio is large—as it must be when the signal is brief. For example, assume that  $d'$  is 1.0 for a brief tone with a signal-to-noise intensity ratio of 4 (=6 dB). This means that  $d'$  is 1.0 for a  $\Delta L$  of 7.0 dB [=  $10 \log(1+4)$ ]. Now assume that the signal-to-noise ratio is doubled to become 8 (=9 dB), which yields a  $\Delta L$  of 9.5 dB [=  $10 \log(1+8)$ ]. In this case,  $d'$  should increase to 2.0 if  $d'$  is proportional to the signal intensity and to 1.36 if  $d'$  is proportional to  $\Delta L$ . Thus, a 3-dB increase in the intensity of a brief signal (with a high signal-to-noise ratio) increases  $d'$  much less if  $d'$  is proportional to  $\Delta L$  than if  $d'$  is proportional to the signal intensity. Conversely, the change in signal intensity needed to compensate for a threefold increase in  $d'$  which may be obtained by presenting nine independent brief components with large signal-to-noise ratios is much larger if  $d'$  is proportional to  $\Delta L$  than if  $d'$  is proportional to the signal intensity. If  $d'$  is proportional to  $\Delta L$ , the brief signal must be lowered to yield a  $\Delta L$  of 2.33 dB (=7 dB/3)—corresponding to a signal-to-

noise ratio of 0.71 (=−1.5 dB)—to produce a  $d'$  of 1.0 for the complex. Thus, the threshold for the brief nine-component complex would be 7.5 dB lower than the threshold for a single component. On the other hand, if  $d'$  is proportional to the signal intensity, the signal-to-noise ratio for each component must be lowered to  $\frac{4}{3}$  (=1.23 dB), such that the threshold difference between the complex and a single component would be 4.77 dB. This shows that the effect of across-frequency integration depends on the form of the psychometric function. If  $d'$  is proportional to  $\Delta L$  and the listeners' decision is based on optimum combination of independent observations whether the signal is brief or long, the threshold improvement (in terms of level per component) obtained for tone complexes will be larger when the signal is brief than when it is long. Thus, it is possible that the efficient across-frequency integration reported for brief signals simply reflects that  $d'$  grows more slowly with signal level when the signal is brief than when it is long, which could result if  $d'$  were proportional to  $\Delta L$  rather than to signal intensity.

A second possible explanation for the lower-than-expected thresholds for brief complex signals is that temporal synchrony across auditory channels may enhance detection of the brief multi-tone complexes with Gaussian envelopes. The complex signals used by van den Brink and Houtgast (1988, 1990a, b) had the components in sine phase around the peak of the envelope, which produces a very impulsive waveform and is likely to promote near-simultaneous neural activity across all the auditory channels excited by the brief signal. Such near-simultaneous neural activity across a broad range of channels is unlikely to occur in the absence of a signal and, thus, could provide a cue for detection of that signal.

To investigate whether the form of the psychometric functions or the temporal synchrony across auditory channels could account for the efficient across-frequency integration of brief complex signals, psychometric functions for masked brief and long tones and tone complexes with different phase relationships were measured in normal-hearing listeners.

## I. METHODS

### A. Subjects

Six subjects, ranging in age between 24 and 71 years, participated.<sup>1</sup> All had absolute thresholds at or below 15 dB HL (ANSI, 1989) at the octave frequencies from 250 to 4000 Hz. Four young subjects (S1–S4) also had normal thresholds at 8000 Hz, whereas two elderly subjects had thresholds of 45 (S5; age 70) and 65 (S6; age 71) dB HL at 8000 Hz. Subjects S1 (the first author), S2, S3, and S5 had previous experience in psychoacoustic tasks. Subjects S4 and S6 received about one hour of practice to become familiar with the listening task. Subjects S1 and S2 were employees of the laboratory; all others were paid for their participation.

### B. Stimuli and apparatus

The signals were either single tones with Gaussian envelopes and frequencies of 630, 1600, and 4000 Hz, or a

nine-tone complex, consisting of nine such components spaced at one-third-octave frequencies between 630 and 4000 Hz. The phase relationship of the components in the nine-tone complex was varied, such that adjacent components were in phase (at 0 degrees), or 90 or 180 degrees out of phase at the envelope peak. Two signal durations (defined in terms of the number of cycles between the half-amplitude points of the Gaussian envelopes of each component) were tested; brief signals were 4.7 cycles in duration, and long signals were 150 cycles in duration. Defining the duration in terms of a fixed number of cycles means that the absolute duration (in ms) of each component decreased with frequency. Such a definition may seem unusual, but it is identical to that used in the original studies by van den Brink and Houtgast (1990a, b) and it ensures that the relative bandwidth of the spectral splatter is independent of frequency.<sup>2</sup> Note that when the components were in phase at the envelope peak, the overall energy of the brief complex signal was 0.9 dB higher than the sum of the individual component energies, and when they were 180 degrees out of phase, the overall energy was 1.1 dB below the sum of the component energies. The signals were digitally generated via a 16-bit D/A converter (TDT DD1) with a 50-kHz sampling rate. The output of the D/A converter was attenuated (TDT PA4), low-pass filtered (TDT FT5,  $f_c = 20$  kHz, 135 dB/oct), and attenuated again (TDT PA4) before it was added to the masking noise.

The masking noise was a uniformly masking noise<sup>3</sup> (Zwicker and Feldtkeller, 1967) with an overall level of 45 or 85 dB SPL. It was presented continuously within a block of trials. The masking noise was produced by a waveform generator (TDT WG2), filtered (TDT PF1), attenuated (TDT PA4) and then added (TDT SM3) to the signal. The signal and masker were led to a headphone buffer (TDT HB6), which fed one earphone of a Sony MDR-V6 headset.

### C. Procedure

Subjects were tested individually in a double-walled, sound-attenuating booth. A two-interval, two-alternative forced-choice paradigm was used; each trial consisted of two 500-ms observation intervals marked by lights on a response panel. The observation intervals were separated by 500 ms. The signal was presented in either the first or second interval with equal *a priori* probability for each interval. Subjects responded by pressing a button corresponding to the interval that contained the signal. Immediately following the response, lights provided correct-response feedback.

In the first stage of the study, masked thresholds were measured using an adaptive procedure with a decision rule that estimates 79.4% correct on the psychometric function (Levitt, 1971). This was done to familiarize the listeners with the stimuli and to aid in setting the appropriate signal levels for measurement of the psychometric functions. As in previous studies from this laboratory (e.g., Florentine *et al.*, 1996), a single threshold measurement was obtained from three interleaved adaptive tracks. The signal level for each track began approximately 10 dB above the estimated threshold; it was decreased following three correct responses and increased following one incorrect response. The step size

was initially 5 dB, but was reduced to 2 dB following the second reversal. Each track terminated after five reversals and threshold for one track was calculated as the average of the signal levels at the last two reversals. A single threshold estimate was taken as the average of the thresholds for the three tracks. Threshold estimates were accepted for inclusion in the final result only if the standard error of the thresholds for the three tracks was less than 3 dB. Each threshold reported is the mean from three such threshold estimates (for a total of nine tracks) obtained over at least two days.

In the second stage of the study, psychometric functions were obtained using a method of constant stimuli (MCS). For each stimulus, ten signal levels, 1.5 dB apart (ranging from  $-7.5$  to  $+6$  dB relative to the adaptive threshold), were chosen to encompass the range from approximately 55% to 95% correct. Each block of 103 trials consisted of three practice trials set to the highest or second highest signal level, and 10 trials of each of the ten signal levels presented in random order. Initially, a total of ten such blocks were obtained, but to ensure homogeneous performance across blocks, only those blocks in which the overall percentage correct was within the 80% confidence interval were retained for further analysis. Additional blocks (usually two to three) were obtained until 100 usable trials for each signal level were obtained. (Note: a comparison of the thresholds obtained using the adaptive and MCS procedures is discussed in the Appendix.)

### D. Data analysis

For each subject, stimulus, signal duration, masker level, and signal level, the 90% confidence interval was calculated across the 100 trials. Only successive points whose confidence intervals were contained within the interval from 50% to 100% were used for further analysis. The  $d'$  values for each of the points fulfilling this criterion (typically 4 to 5) were calculated from the hit and false-alarm rates, and plotted on a logarithmic scale as a function of signal level (in dB SPL). The data were fitted with straight lines using the method of least squares. These lines characterized the psychometric functions by their slope and by a threshold, which was defined as the SPL at the intercept with  $d' = 1.16$  (corresponding to 79.4% correct responses).

Four-way (stimulus type  $\times$  signal duration  $\times$  masker level  $\times$  subject) analyses of variance (ANOVAs) for repeated measures were performed on the slopes and the threshold signal-to-noise ratios in dB (Data Desk 6.0, Data Description, Inc., Ithaca, NY, 1997). In these analyses, stimulus type, signal duration, and masker level were considered fixed effects, whereas subject was considered a random effect. Scheffé *post hoc* tests for contrasts were performed as appropriate to evaluate the sources of significant effects and interactions. Differences are reported as significant when  $p < 0.05$ .

## II. RESULTS

### A. Slopes of the psychometric functions

Figure 1 shows psychometric functions obtained for one subject (S1) at a masker level of 45 dB SPL. Percent correct (top panels) and  $d'$  calculated according to the method de-

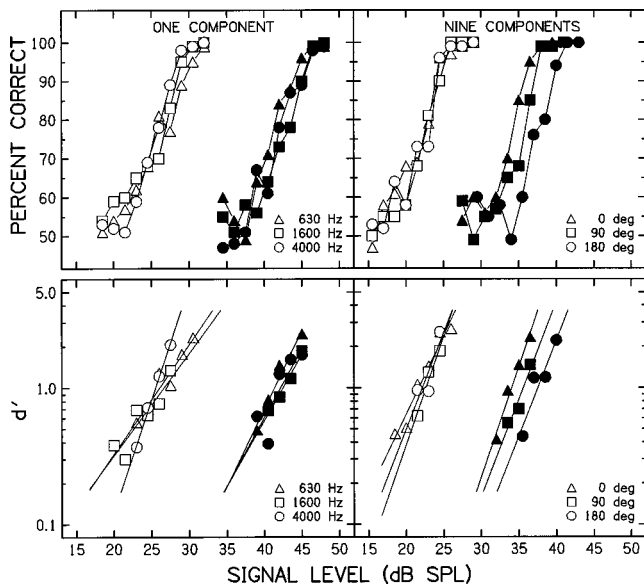


FIG. 1. Representative psychometric functions obtained from subject S1 at a masker level of 45 dB SPL. Percentage correct (top panels) and  $d'$  (bottom panels) are plotted as functions of signal level. Results for the tones and nine-tone complexes are shown across panels; filled and unfilled symbols represent the brief and long signals, respectively. Different symbols represent either the frequency (single tone) or the component phase relationship (nine-tone complex) of the signal.

scribed earlier (bottom panels) are plotted as a function of signal level. These functions are examples of the psychometric functions obtained in the present study and reflect the trends observed in most of the data. In general, the functions obtained with the tones (left panels) show considerable overlap at the three signal frequencies and are largely parallel to one another, whether the tones were brief (filled symbols) or long (unfilled symbols). The exception is the 4000-Hz signal, which tends to show steeper slopes than either the 630- or 1600-Hz signal, although this was not the case for all subjects. Overall, the functions for the tones also tend to be shallower than those obtained with the nine-tone complexes (right panels). For the nine-tone complexes, the psychometric functions obtained with the long signals overlap one another, whereas those obtained with the brief signals are

shifted to the right by about 5 dB as the phase relationship changes from 0 to 180 degrees. This shift indicates poorer performance as the components move out of phase with one another.

These trends can be examined more closely by evaluating the slopes of the lines fitted to the  $d'$  values. Tables I and II show the slopes of the lines fitted to the functions for each subject and stimulus for masker levels of 45 and 85 dB SPL, respectively. The standard errors of the slope estimates are shown in parentheses. A four-way ANOVA for repeated measures performed on the slope values showed no significant effects of subject or masker level, and thus Fig. 2 shows the slope values averaged across those parameters. (As mentioned later, however, an analysis for the tones alone revealed a small but significant effect of masker level.)

In general, the slope estimates obtained for the tones (filled symbols) are close to unity for both signal durations, consistent with the idea that  $d'$  ought to be proportional to the signal energy of single tones presented in noise (Egan *et al.*, 1959). However, there is a slight tendency for the long tones to show somewhat shallower slopes than the brief tones, although this is primarily due to the very shallow slopes obtained for the 630-Hz, long-duration signal. There is also a tendency for the 4000-Hz signal to exhibit slightly steeper slopes than either the 630- or 1600-Hz signal. The slope estimates for the complex signals (unfilled symbols) are generally steeper than the approximately unity slopes obtained for the tones, although the difference between the slopes for the tones and tone complexes is clearly larger for the brief than for the long signals. The average slopes for the brief complexes range from 1.40 to 1.82, whereas those for the long complex signals range from 1.14 to 1.37. There does not appear to be any effect of phase relationship.

Three four-way ANOVAs for repeated measures performed on the slope estimates largely confirmed the trends noted above. The first ANOVA retained all stimulus conditions; the stimulus dimension comprised three tones at different frequencies and three complexes with different phase relationships. The second and third ANOVAs analyzed the data for the tones and tone complexes separately to examine

TABLE I. Slopes of the psychometric functions [ $10 \cdot \log(d')$  versus signal level] obtained for each subject and condition at a masker level of 45 dB SPL. The values in parentheses are the standard errors of the slope estimates. The means (and standard deviations) across subjects are shown in bold.

Subject	Duration (cycles)	Tones			Tone complexes		
		630 Hz	1600 Hz	4000 Hz	0 degrees	90 degrees	180 degrees
S1	4.7	1.14 (0.13)	0.96 (0.11)	1.00 (0.37)	1.61 (0.18)	1.42 (0.42)	1.41 (0.37)
	150	0.80 (0.12)	0.74 (0.17)	1.66 (0.07)	1.16 (0.12)	1.59 (0.32)	1.42 (0.86)
S2	4.7	0.85 (0.15)	0.81 (0.17)	0.94 (0.13)	1.12 (0.32)	1.43 (0.11)	1.33 (0.21)
	150	0.64 (0.13)	1.36 (0.24)	1.20 (0.07)	0.97 (0.00)	1.16 (0.51)	1.02 (0.28)
S3	4.7	1.30 (0.12)	1.02 (0.04)	1.42 (0.30)	1.62 (0.20)	1.36 (0.23)	1.65 (0.14)
	150	0.81 (0.17)	0.88 (0.15)	1.09 (0.13)	1.45 (0.06)	1.30 (0.39)	0.98 (0.21)
S4	4.7	0.93 (0.15)	1.28 (0.22)	1.34 (0.14)	2.23 (0.00)	1.61 (0.35)	1.37 (0.14)
	150	0.71 (0.14)	0.69 (0.14)	1.34 (0.06)	1.30 (0.19)	0.85 (0.17)	1.17 (0.18)
S5	4.7	1.18 (0.23)	0.97 (0.27)	0.89 (0.17)	1.05 (0.39)	1.92 (0.76)	1.38 (0.23)
	150	1.32 (0.21)	1.00 (0.11)	0.97 (0.24)	1.59 (0.14)	1.76 (0.91)	1.39 (0.32)
S6	4.7	0.96 (0.19)	1.66 (0.03)	1.12 (0.31)	1.81 (0.18)	1.44 (0.17)	1.43 (0.21)
	150	0.79 (0.16)	0.84 (0.10)	0.82 (0.38)	1.76 (0.30)	1.14 (0.28)	1.58 (0.27)
<b>Mean</b>	<b>4.7</b>	<b>1.06 (0.17)</b>	<b>1.12 (0.31)</b>	<b>1.12 (0.22)</b>	<b>1.57 (0.44)</b>	<b>1.53 (0.21)</b>	<b>1.43 (0.11)</b>
	<b>150</b>	<b>0.84 (0.24)</b>	<b>0.92 (0.24)</b>	<b>1.18 (0.30)</b>	<b>1.37 (0.29)</b>	<b>1.30 (0.33)</b>	<b>1.26 (0.24)</b>

TABLE II. Slopes of the psychometric functions [ $10 \cdot \log(d')$  versus signal level] obtained for each subject and condition at a masker level of 85 dB SPL. The values in parentheses are the standard errors of the slope estimates. The means (and standard deviations) across subjects are shown in bold.

Subject	Duration (cycles)	Tones			Tone complexes		
		630 Hz	1600 Hz	4000 Hz	0 degrees	90 degrees	180 degrees
S1	4.7	1.18 (0.31)	1.15 (0.10)	1.59 (0.27)	2.29 (0.53)	1.56 (0.30)	0.79 (1.01)
	150	0.91 (0.13)	1.43 (0.23)	1.28 (0.60)	1.94 (0.00)	1.37 (0.21)	1.41 (0.16)
S2	4.7	1.19 (0.21)	0.89 (0.14)	1.69 (0.09)	0.92 (0.07)	1.52 (0.46)	1.41 (0.27)
	150	0.76 (0.10)	0.84 (0.22)	1.06 (0.11)	0.65 (0.46)	0.90 (0.16)	1.60 (0.27)
S3	4.7	0.97 (0.07)	0.88 (0.09)	1.64 (0.11)	1.35 (0.37)	1.86 (0.51)	1.87 (0.34)
	150	0.86 (0.17)	1.30 (0.19)	1.20 (0.21)	0.85 (0.16)	1.25 (0.10)	1.22 (0.21)
S4	4.7	1.25 (0.18)	1.09 (0.21)	1.32 (0.20)	2.38 (0.00)	1.47 (0.09)	1.35 (0.13)
	150	0.86 (0.05)	0.85 (0.16)	0.88 (0.08)	1.31 (0.56)	1.48 (0.16)	1.24 (0.11)
S5	4.7	1.90 (0.07)	0.78 (0.24)	1.17 (0.15)	2.30 (0.16)	2.77 (0.23)	1.76 (0.18)
	150	0.72 (0.11)	1.34 (0.19)	1.57 (0.19)	0.96 (0.12)	0.82 (0.27)	1.00 (0.11)
S6	4.7	1.14 (0.14)	0.87 (0.10)	1.24 (0.08)	1.58 (0.58)	1.72 (0.35)	1.19 (0.59)
	150	0.66 (0.14)	0.80 (0.23)	1.75 (0.13)	1.79 (0.27)	1.00 (0.28)	1.39 (0.28)
<b>Mean</b>	<b>4.7</b>	<b>1.27 (0.32)</b>	<b>0.94 (0.14)</b>	<b>1.44 (0.22)</b>	<b>1.80 (0.61)</b>	<b>1.82 (0.49)</b>	<b>1.40 (0.39)</b>
	<b>150</b>	<b>0.80 (0.10)</b>	<b>1.09 (0.29)</b>	<b>1.29 (0.32)</b>	<b>1.25 (0.52)</b>	<b>1.14 (0.27)</b>	<b>1.31 (0.21)</b>

more fully the effects of frequency (tones) and phase relationship (tone complexes). The first ANOVA revealed a significant effect of stimulus [ $F(5,25) = 9.47, p < 0.0001$ ], and Scheffé tests for contrasts confirmed that the slopes were shallower for the tones than for the complexes, although this depended on the frequency of the tone. The slope for 630-Hz tones was significantly shallower than those obtained for any of the complexes and the slope for 1600-Hz tones was significantly shallower than those obtained for the 0- and 90-degree complexes. The slope for 4000-Hz tones was not significantly different from those obtained for any of the complexes. The ANOVA also revealed that the functions for the long signals were shallower than those for the brief signals [ $F(1,25) = 16.54, p < 0.01$ ]. As stated previously, there were no significant effects of subject or masker level, and there were no interactions.

The ANOVA for the tones revealed a significant effect of frequency on the slopes of the psychometric functions

[ $F(2,10) = 11.89, p < 0.005$ ]. Scheffé tests for contrasts showed that the functions obtained at 4000 Hz were significantly steeper than those obtained at 630 and 1600 Hz. There was also a small but significant effect of duration [ $F(1,10) = 7.86, p < 0.05$ ]; the slopes for the long tones were on average 0.16 shallower than those for the brief tones. Finally, there was a small effect of masker level that was not evident in the first ANOVA when both the tones and tone complexes were included in the analysis [ $F(1,10) = 8.37, p < 0.05$ ]. The slopes obtained at a masker level of 85 dB SPL were on average 0.10 steeper than those obtained at 45 dB SPL. The ANOVA for the tone complexes showed no significant effects of the phase relationship of the components or of masker level on the slope estimates. However, there was a significant effect of duration [ $F(1,10) = 10.84, p < 0.05$ ]; the slopes obtained for the long complexes were on average 0.32 shallower than those obtained for the brief complexes.

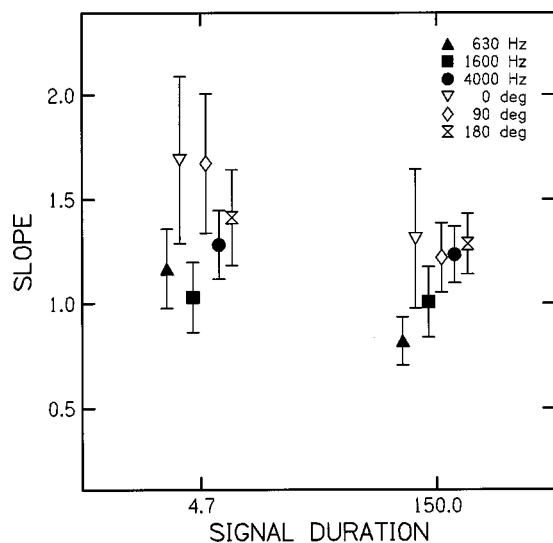


FIG. 2. Slopes of the psychometric functions averaged across subject and masker level for both signal durations. Slopes obtained for the tones and nine-tone complexes are shown by the filled and unfilled symbols, respectively. Error bars show standard deviations.

## B. Thresholds from the psychometric functions

The masked thresholds obtained from the psychometric functions are shown in Fig. 3; results for the 45- and 85-dB SPL masker levels are shown in the left and right panels, respectively. Thresholds for the brief (triangles) and long (squares) signals are shown as a function of signal frequency (tones; filled symbols; top axis) or phase relationship (tone complexes; unfilled symbols; bottom axis). A repeated-measures ANOVA on the threshold estimates showed that the effects of the stimulus variables agreed across subjects, and therefore only the mean data are shown. However, it should be noted that the thresholds for all conditions tended to be slightly higher for the two elderly subjects than for the young subjects. The analysis also revealed a significant effect of masker level (after transforming the threshold values to signal-to-noise ratios to account for the 40-dB level difference between the maskers) [ $F(1,25) = 151.35, p < 0.001$ ]. That is, subjects required the signal-to-noise ratio to be approximately 1.0 dB greater for the 45-dB SPL masker than for the 85-dB SPL masker for almost all conditions. Thus,

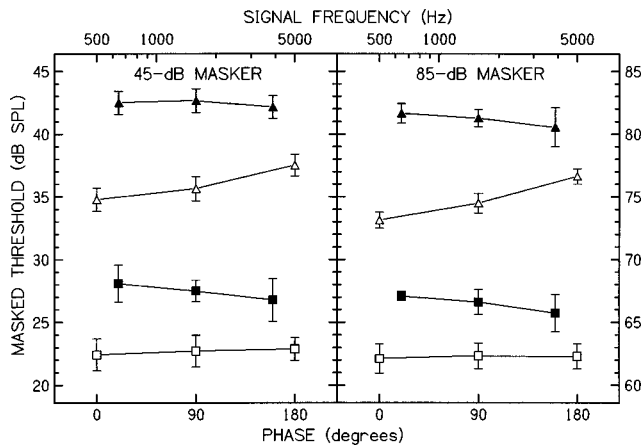


FIG. 3. Masked thresholds (corresponding to 79.4% correct) obtained from the psychometric functions as a function of signal frequency (tones; filled symbols; referenced to the top axis) or component-phase relationship (tone complexes; unfilled symbols; referenced to the bottom axis). The thresholds, averaged across subjects, are shown for the brief (triangles) and long (squares) signals. Results for the 45 and 85 dB SPL masker levels are shown in the left and right panels, respectively.

the trends observed in the results are essentially the same at the two masker levels (compare the right and left panels) and will be considered together.

For both short and long tones, masked thresholds tended to decrease as the frequency increased from 630 to 4000 Hz. This occurred despite our masker having a higher spectrum level at high frequencies than the uniformly masking noise described by Zwicker and Feldtkeller (1967) (see footnote 3). On average, the effect of frequency was only about 1 dB. Thus, the components that made up the complex signal, although not exactly equally detectable, were very nearly so. Accordingly, it is reasonable to use an average of the masked thresholds for the three individual tones for comparison with the thresholds for the tone complexes.

For the long complex signals, the masked thresholds were nearly independent of the phase relationship of the components, and were about 4 to 5 dB lower than the average thresholds for the tones. Thus, increasing the number of components from one to nine increased sensitivity for the long stimuli threefold (i.e., proportionally to  $\sqrt{n}$ ), representing a threshold improvement of about  $5 * \log(n)$ . For the brief complex signals, however, masked thresholds were not independent of the phase relationship; thresholds tended to increase 3 to 4 dB as the phase relationship changed from 0 to 180 degrees. This increase is larger than the 2-dB change that would be expected based only on the change in overall energy of the signals. Thus, the amount of across-frequency integration for brief signals depended on the phase relationship of the components that made up the complex. When the components were 180 degrees out-of-phase, increasing the number of components from one to nine decreased threshold by 4 to 5 dB, consistent with the  $5 * \log(n)$  across-frequency integration rule. However, when the components were in phase, increasing the number of components decreased threshold by 7 to 8 dB, consistent with a more efficient  $8 * \log(n)$  integration rule. It should be noted, however, that

these values refer to nominal levels of the individual components and have not been corrected for the 0.9-dB gain in total energy for the 0-degree complex and the 1.1-dB loss in total energy for the 180-degree complex. Making the proper corrections reduces the phase effect to about 1 dB and the amount of across-frequency integration for the brief, in-phase complexes to 6 to 7 dB, which is still more efficient than the 4 to 5 dB of integration observed for the long complexes, or brief, out-of-phase complexes.

Three four-way ANOVAs for repeated measures performed on the threshold signal-to-noise ratios confirmed the trends noted above. As with the slope estimates, the first ANOVA retained all stimulus conditions, whereas the second and third ANOVAs analyzed the data for the tones and tone complexes separately, in order to more fully examine the effects of frequency (tones) and phase relationship (tone complexes). The first ANOVA revealed a significant effect of duration [ $F(1,25) = 3262.1, p < 0.0001$ ], confirming that thresholds for the long signals were significantly lower than those for the brief signals, and a significant effect of stimulus [ $F(5,25) = 389.47, p < 0.0001$ ]. Scheffé tests for contrasts confirmed that thresholds for the tone complexes were significantly lower than those for the tones. There was also a significant interaction between stimulus and duration [ $F(5,25) = 37.95, p < 0.0001$ ]. The Scheffé tests revealed that the brief signals were more likely to show differences across stimuli. Also, as already described above, the effect of masker level was also significant. Recall that the signal-to-noise ratio at masked threshold was 1.0 dB higher in the 45-dB masker than in the 85-dB masker.

The ANOVA on the threshold values for the tones revealed a small but significant effect of frequency [ $F(2,10) = 5.55, p < 0.05$ ]. Scheffé tests for contrasts revealed no differences between thresholds at 630 and 1600 Hz, nor between 1600 and 4000 Hz. However, thresholds at 4000 Hz were significantly lower than those at 630 Hz. As in the first ANOVA, there were significant effects of duration [ $F(1,10) = 2841.2, p < 0.0001$ ] and masker level [ $F(1,10) = 83.25, p < 0.001$ ]. There were no significant interactions.

The ANOVA for the tone complexes revealed a significant effect of phase relationship [ $F(2,10) = 169.32, p < 0.0001$ ]; Scheffé tests for contrasts revealed that threshold in the 180-degree condition was significantly higher than the 90-degree condition, which was significantly higher than the 0-degree condition. The ANOVA also revealed that thresholds for the long complexes were significantly lower than those for the brief complexes [ $F(1,10) = 2771.70, p < 0.0001$ ], and that there was a significant interaction between the phase relationship and duration [ $F(2,10) = 55.89, p < 0.0001$ ]. Tests for contrasts of this interaction confirmed that the effect of phase relationship was present only for the brief complexes. Also, as in the previous ANOVAs, there was also a significant effect of masker level [ $F(1,10) = 48.10, p < 0.001$ ].

### III. DISCUSSION

#### A. Comparison with previous results

As in previous studies (Green *et al.*, 1959; Buus *et al.*, 1986), the slopes of the psychometric functions for the single

tones were generally close to unity, consistent with what would be expected from an energy-detector model, which predicts that  $d'$  ought to be nearly proportional to the signal energy of a tone presented in noise (Egan *et al.*, 1959). It is unclear why the 4000-Hz signal in the present study tended to produce somewhat steeper slopes than either the 630- or 1600-Hz signal, although a similar result was reported by Buus *et al.* (1986) who found that psychometric functions for a 3850-Hz tone were steeper than those for 220- or 1100-Hz tones. However, unlike these previous studies that also reported that the slopes were independent of the number of components in the signal, we found that the slopes for the tone complexes were steeper than for the single tones, particularly for the brief signals. The difference between the slopes for the tones and tone complexes averaged 0.44 for the brief stimuli, and 0.25 for the long stimuli. It is likely that the previous studies did not observe this difference due to using longer signals ( $\geq 100$  ms). Certainly, the present study shows that the difference between the slopes for the tones and tone complexes is clearly smaller for long signals than for brief signals. The steeper slopes for the tone complexes indicate that the decision variable used for detection of complex stimuli is unlikely to be based on energy, particularly when the signal is very brief.

The thresholds obtained from the psychometric functions were similar to those expected based on previous reports of across-frequency and temporal integration. As expected, duration had a pronounced effect on masked thresholds. For the tonal signals, thresholds for the long signals were approximately 15 dB lower than those for the brief signals, corresponding almost exactly with the classical temporal integration rule of  $10 \cdot \log(T_l/T_s)$ , where  $T_l$  and  $T_s$  are the durations of the long and short signals, respectively. Van den Brink and Houtgast (1990a, b) found the same correspondence with similar narrow-band stimuli. For the complex signals, however, the amount of temporal integration was dependent on the phase relationship of the components. When the components were 180 degrees out of phase, the thresholds for the long signals were approximately 14 dB lower than those for the brief signals, whereas the difference was only 12 dB when they were in phase. When our values are corrected for the gain or loss of total energy incurred at the different component phases of these brief signals, however, this 2-dB difference disappears. After correction, the amount of temporal integration for complexes with either phase relationship is 12.9 dB, which is less than the 15-dB improvement that would be expected based on the classical temporal integration rule of  $10 \cdot \log(T_l/T_s)$ .

The amount of across-frequency integration also depended on the phase relationship of the components in the complex signals, but only for the brief stimuli. For the long signals, thresholds for the complexes were 4–5 dB lower than the thresholds for the tones regardless of the phase relationship of the components in the complex. Thus, increasing the number of components from one to nine increased sensitivity proportionally to  $\sqrt{n}$ , consistent with the optimal decision rule that predicts a threshold improvement of  $5 \cdot \log(n)$ . For the brief stimuli, however, the  $5 \cdot \log(n)$  across-frequency integration rule only held when the compo-

nents in the complex were 180 degrees out-of-phase with one another. When the components were in-phase, increasing the number of components from one to nine increased sensitivity nearly proportionally to  $n$ , consistent with a more efficient  $8 \cdot \log(n)$  integration rule (prior to correction for the 0.9-dB gain in total energy obtained for the brief in-phase complex). This effect of duration on across-frequency integration is difficult to reconcile with either the independent-thresholds model or the multi-band decision model, in which the auditory filter bandwidths and channel weights are invariant over time, suggesting that other listening strategies may play a role in detecting very brief complex stimuli.

## B. Cues for detection of multi-tone complexes

The results for the long stimuli show that  $d'$  increases proportionally with the square root of the number of components, and that threshold for the nine-component signal is about  $5 \cdot \log(9)$  dB lower than for a single component, independent of the phase relationship of the components in the complex. These results are consistent with a multi-band energy-detector model where information is combined across frequency channels in an optimal manner. The results for the brief stimuli, however, are not consistent with this model. For the brief stimuli,  $d'$  increases proportionally with the number of components, and threshold for the nine-component signal is about  $8 \cdot \log(9)$  dB lower than for a single component, but only when the components in the complex were in phase with one another. (Note that this value is not corrected for the 0.9-dB gain in total energy obtained for the brief in-phase complex.)

As suggested in the Introduction, one possible explanation for the lower-than-expected thresholds for the brief stimuli was based on the form of the psychometric functions. That is, if detection for a tone in noise is similar to level discrimination (i.e., where detection is based on the level increment caused by the signal), then the psychometric functions, described in terms of  $\log(d')$  as a function of signal level, should be shallower for brief than long signals, as they are for level discrimination (Buus and Florentine, 1991; Buus *et al.*, 1995). However, in the present study the reverse was observed. The long signals showed significantly shallower functions than the brief signals. This finding indicates that the more efficient across-frequency integration observed for brief, than for long, complex signals cannot be explained by the slopes of the psychometric functions.

A second explanation is that the decision rule for detection is nonoptimal for very brief stimuli. Wright and Dai (1994a, b) investigated detection of short and long tones that were presented at either an expected frequency or at an unexpected frequency. (The expectation was formed by presenting the expected frequency in 75% of the trials and each of eight unexpected frequencies in only 3% of the trials.) Wright and Dai showed that unexpected frequencies were detected much more poorly than the expected frequency when the signal was long, whereas the expected frequency and a broad range of unexpected frequencies were detected roughly equally well when the signal was brief. They proposed that subjects monitored only the auditory filter centered at the expected frequency when the signal was long,

allowing for optimal detection of expected signals, but very poor detection of unexpected signals. However, when the signal was brief, subjects monitored several adjacent auditory filters, allowing for good performance for signals encompassing a broad range of frequencies. Wright and Dai (1994a, b) argued that subjects might be unable to locate the optimal filter for detecting brief signals because their pitches were unclear; thus, the subjects were likely to adopt a broadband listening strategy. Such a broadband listening strategy for brief signals could account for the efficient across-frequency integration for the brief stimuli observed in the present study. However, the fact that the psychometric functions for the brief tones and tone complexes had unequal slopes makes this unlikely. If the same broadband listening strategy were used for all brief stimuli, the slopes of the psychometric functions should have been equal, regardless of the type of stimuli.

A third explanation for the efficient across-frequency integration of brief complex signals is that temporal synchrony across several adjacent auditory channels may enhance detection of the multi-tone complexes. The brief in-phase multi-tone complexes used here and by van den Brink and Houtgast (1988, 1990a, b) had the components of the complexes in sine phase around the peak of the envelope. This produces a very impulsive waveform that is likely to promote near-simultaneous neural activity across all the auditory channels excited by the brief signal. Such near-simultaneous neural activity across a broad range of channels is unlikely to occur in the absence of a signal and thus, could aid in the detection of it. That thresholds became poorer when the components were out of phase in the present study, or when they were shifted by some time delay (van den Brink and Houtgast, 1990b), indicates that temporal synchrony is likely to be an important cue underlying the efficient across-frequency integration.

The detection of such temporal synchrony may be occurring at the level of the cochlear nucleus. Octopus cells of the posteroventral cochlear nucleus, which are known to carry precise information about the timing of action potentials, have large dendrites that receive converging input from nerve fibers originating from broad regions of the cochlea (Osen, 1969). The change in potential contributed by any individual input generally is small and is insufficient to elicit an action potential. However, when several inputs are stimulated synchronously, they can elicit a response from the octopus cell. Thus, octopus cells are likely to respond only when several auditory-nerve fibers tuned to different frequencies fire nearly simultaneously (Golding *et al.*, 1995, 1999). This behavior is consistent with the across-frequency synchrony that is necessary to obtain efficient across-frequency integration for detection of brief tone complexes. Therefore, it seems likely that the octopus cells may play an important role in the detection of complex sounds with a high degree of across-frequency synchrony, such as the brief tone complexes used in the present study.

#### IV. SUMMARY

Detection cues for efficient across-frequency integration of brief, complex signals were investigated by measuring

psychometric functions ( $\log d'$  versus signal level in dB SPL) for detection of brief and long tones and tone complexes presented in broadband noise. The results can be summarized as follows:

- (1) The slopes of the psychometric functions were steeper for the brief than for the long signals, and steeper for the tone complexes than for the tones, particularly for the brief stimuli. This finding indicates that the decision variable for signal detection is dependent on duration and is different for simple and complex stimuli. Moreover, it indicates that the slopes of the psychometric functions cannot explain the more efficient across-frequency integration observed for brief than for long signals.
- (2) For the long signals, increasing the number of components from one to nine increased sensitivity proportionally to  $\sqrt{n}$ . That is, the amount of across-frequency integration was well described by an optimal decision rule that predicts a threshold improvement of  $5 \cdot \log(n)$  for an  $n$ -component signal, independent of the phase relationship of the components.
- (3) For the brief signals, increasing the number of components from one to nine increased sensitivity proportionally to  $\sqrt{n}$  when the components in the complex were 180 degrees out-of-phase with one another, but nearly proportional to  $n$  when the components were in-phase with one another. Given that the psychometric functions had slopes around 1.3, the thresholds for brief, in-phase complexes show a relatively efficient across-frequency integration rule of  $8 \cdot \log(n)$ , or about  $7 \cdot \log(n)$  when corrected for the gain in total energy. The larger improvement in threshold when increasing the number of components from one to nine for brief in-phase stimuli than for out-of-phase stimuli suggests that temporal synchrony across multiple frequency-selective channels is likely to be an important cue for the efficient across-frequency integration of brief multi-tone complexes.

#### ACKNOWLEDGMENTS

We wish to thank Marjorie Leek and an anonymous reviewer for their helpful comments on an earlier version of this manuscript. This work was supported by NIH-NIDCD Grant No. RO1DC00187 and manuscript preparation was funded in part by NIH-NIDCD Grant No. RO1DC01376.

#### APPENDIX: COMPARISON BETWEEN THRESHOLDS OBTAINED WITH THE ADAPTIVE AND MCS PROCEDURES

The adaptive procedure used in the first stage of this study (described in detail in Sec. IC) employed three interleaved adaptive tracks to obtain a single threshold measurement. Threshold was calculated as the average signal level at the last two (of five) reversals for each of the three tracks. The reliability of this method, used extensively in our laboratory, can be examined by comparing the threshold values obtained in this manner with those obtained using the more classical method of constant stimuli. The MCS procedure is generally considered to be highly accurate due to the large

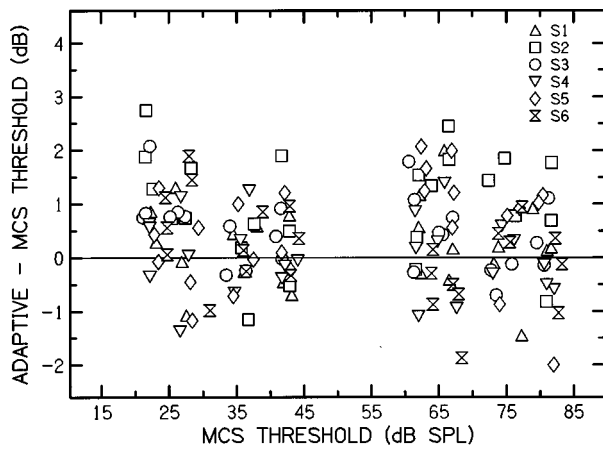


FIG. A1. Differences between the thresholds obtained using the adaptive procedure and the method of constant stimuli (MCS) plotted as a function of the MCS thresholds. Results are shown for all six subjects.

number of observations (in this case, 100) that make up each point on the psychometric function. The adaptive procedure, on the other hand, may have inherent biases and variability due to (1) inefficient placement of the observations relative to the point of interest (i.e., signal level at threshold) on the psychometric function, and (2) only a small number of observations being used to calculate threshold.

Figure A1 shows the differences (in dB) between the thresholds obtained using our adaptive procedure and the MCS procedure as a function of the MCS thresholds. The data points are normally distributed around a mean difference of about 0.4 dB, with approximately 73% of the data points falling within  $\pm 1$  dB of the mean. More than 97% of the data fall within  $\pm 2$  dB of the mean, and all are within 2.6 dB of the mean. That the adaptive thresholds tended to be somewhat higher than the MCS thresholds may reflect a slight practice effect as the adaptive procedure was completed prior to beginning the MCS procedure. However, the differences in threshold between the two procedures are generally very small, indicating that our adaptive procedure is a consistent, valid method for obtaining accurate threshold measurements. The relatively narrow distribution of the threshold differences also indicates that our procedure very rarely yields a threshold estimate that is far from the true value. Simulations show that this desirable property results, in part, from the procedure's reliance on many independent short tracks, as opposed to a few long tracks, to obtain a threshold. "Lucky guesses," especially in the beginning of a track, may cause the signal level to migrate far below the threshold and make convergence somewhat slow. Therefore, any single track—even quite long ones—can occasionally produce a divergent threshold estimate. If only a few tracks are used, the final threshold may be affected considerably by a single divergent track. The present procedure ensures that many independent tracks go into the final threshold estimate, which reduces the impact of an occasional divergent track. Furthermore, the three-track procedure typically provides an indication of divergent tracks by a large variability across the three concurrent threshold estimates. The simulations show that discarding threshold estimates with a standard error larger than 2.5 or 3 dB across the three tracks eliminates

most of the divergent measurements without eliminating an excessive number of good measurements. These considerations, together with the well-behaved distribution of thresholds in Fig. A1, indicate that the use of many short tracks is an advantageous feature of the present adaptive procedure.

<sup>1</sup>A fifth young normal-hearing subject was also recruited to participate in these experiments. However, after analysis of her data, it was determined that she exhibited a strong response bias towards the second observation interval while completing the constant-stimuli procedure. This resulted in unrealistic and uninterpretable values of  $d'$  for this subject. The bias persisted after reinstruction and retraining, and thus her data are not included in the analysis presented here. Her data obtained using the adaptive procedure, although somewhat variable, were generally consistent with the results of the other subjects.

<sup>2</sup>The duration of the signals is defined in terms of number of cycles to make the in-phase stimuli identical to those used by van den Brink and Houtgast (1990a, b). Keeping the number of cycles constant across frequency ensures that the spectral splatter encompasses the same relative bandwidth (e.g.,  $\frac{1}{3}$  octave for the brief stimuli) at all frequencies. Of course, defining the duration as a number of cycles causes the absolute duration of each component to decrease with frequency. For example, the absolute durations of the individual components ranged from 1.18 (4000 Hz) to 7.5 (630 Hz) ms for the brief signal, and from 37.5 (4000 Hz) to 238 (630 Hz) ms for the long signal. Because the envelopes of the individual components all peaked at the same time (by delaying the onsets of the higher frequency components relative to the lower components), the brief in-phase complexes were highly likely to produce nearly synchronous activity across auditory channels tuned to different frequencies. Whether similar synchrony can be elicited by brief complex signals consisting of equally audible components with equal absolute durations and rise-fall times may depend on the details of the envelope shape. Thus, it is unclear whether the efficient cross-frequency integration obtained with the present stimuli can be generalized to complex stimuli in which the absolute duration is constant across component frequencies.

<sup>3</sup>The uniformly masking noise from Zwicker and Feldtkeller (1967) is a noise designed to produce masked thresholds for long tones that are approximately equal across frequency. It has a flat spectrum below about 500 Hz, and decreases at a rate of approximately  $-3$  dB/oct from 500 to 20 000 Hz. Pilot data collected from S1 and S2, however, revealed that this noise produced thresholds, at least for the stimuli used here, that were approximately 2 dB lower at 4000 Hz than at 630 and 1600 Hz. Thus, the noise was altered such that its spectrum decreased at a rate of about 1.4 dB/oct above 1600 Hz, producing more similar (although still not exactly equal) thresholds across frequency (see, for example, Fig. 3).

ANSI (1989). ANSI S3.6-1989, "Specifications for audiometers" (American National Standards Institute, New York).

Buus, S., and Florentine, M. (1991). "Psychometric functions for level discrimination," *J. Acoust. Soc. Am.* **90**, 1371-1380.

Buus, S., Florentine, M., and Zwicker, T. (1995). "Psychometric functions for level discrimination in cochlearly impaired and normal listeners with equivalent-threshold masking," *J. Acoust. Soc. Am.* **98**, 853-861.

Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (1986). "Decision rules in detection of simple and complex tones," *J. Acoust. Soc. Am.* **80**, 1646-1657.

de Boer, E. (1966). "Intensity discrimination of fluctuating signals," *J. Acoust. Soc. Am.* **40**, 552-560.

Egan, J. P., Schulman, A. I., and Greenburg, G. Z. (1959). "Operating characteristics determined by binary decisions and by ratings," *J. Acoust. Soc. Am.* **31**, 768-773.

Florentine, M., Buus, S., and Poulsen, T. (1996). "Temporal integration of loudness as a function of level," *J. Acoust. Soc. Am.* **99**, 1633-1644.

Golding, N. L., Ferragamo, M. J., and Oertel, D. (1999). "Role of intrinsic conductances underlying responses to transients in octopus cells of the cochlear nucleus," *J. Neurosci.* **19**, 2897-2905.

Golding, N. L., Robertson, D., and Oertel, D. (1995). "Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision," *J. Neurosci.* **15**, 3138-3153.

Green, D. M. (1958). "Detection of multiple component signals in noise," *J. Acoust. Soc. Am.* **30**, 904-911.



- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).
- Green, D. M., Birdsall, T. G., and Tanner, W. P. (1957). "Signal detection as a function of signal intensity and duration," *J. Acoust. Soc. Am.* **29**, 523–531.
- Green, D. M., McKey, M. J., and Licklider, J. C. R. (1959). "Detection of a pulsed sinusoid in noise as a function of frequency," *J. Acoust. Soc. Am.* **31**, 1446–1452.
- Higgins, M. B., and Turner, C. W. (1990). "Summation bandwidths at threshold in normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **88**, 2625–2630.
- Langhans, A., and Kohlraush, A. (1992). "Differences in auditory performance between monaural and diotic conditions. I: Masked thresholds in frozen noise," *J. Acoust. Soc. Am.* **91**, 3456–3470.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Osen, K. K. (1969). "Cytoarchitecture of the cochlear nuclei in the cat," *J. Comp. Neurol.* **136**, 453–472.
- Schafer, T. H., and Gales, R. S. (1949). "Auditory masking of multiple tones by random noise," *J. Acoust. Soc. Am.* **21**, 392–398.
- Scholl, H. (1962). "Über die Bildung der Hörschwellen und Mithörschwellen von Dauerschall," *Frequenz* **15**, 58–64.
- Taylor, M. M., Forbes, S. M., and Creelman, C. D. (1983). "PEST reduces bias in forced choice psychophysics," *J. Acoust. Soc. Am.* **74**, 1367–1374.
- van den Brink, W. A. C., and Houtgast, T. (1988). "Across-frequency integration in signal detection," in *Basic Issues in Hearing—Proceedings of the 8th International Symposium on Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London).
- van den Brink, W. A. C., and Houtgast, T. (1990a). "Efficient across-frequency integration in short-signal detection," *J. Acoust. Soc. Am.* **87**, 284–291.
- van den Brink, W. A. C., and Houtgast, T. (1990b). "Spectro-temporal integration in signal detection," *J. Acoust. Soc. Am.* **88**, 1703–1711.
- Wright, B. A., and Dai, H. (1994a). "Detection of unexpected tones with short and long durations," *J. Acoust. Soc. Am.* **95**, 931–938.
- Wright, B. A., and Dai, H. (1994b). "Detection of unexpected tones in gated and continuous maskers," *J. Acoust. Soc. Am.* **95**, 939–948.
- Zwicker, E., and Feldtkeller, R. (1967). *Das Ohr als Nachrichtenempfänger* (Hirzel, Stuttgart, Germany). Also available in English translation by H. Müsch, S. Buus, and M. Florentine as *The Ear as a Communication Receiver* (Acoustical Society of America, Woodbury, NY, 1999).

# Temporal dynamics of pitch strength in regular-interval noises: Effect of listening region and an auditory model

L. Wiegrebe<sup>a)</sup>

Zoologisches Institut der Universität München, Luisenstrasse 14, 80333 München, Germany

H. S. Hirsch

Centre for the Neural Basis of Hearing, Physiology Department, Cambridge University, Cambridge, United Kingdom and Lehrstuhl für Mensch-Maschine Kommunikation, Technische Universität München, Germany

R. D. Patterson

Centre for the Neural Basis of Hearing, Physiology Department, Cambridge University, Cambridge, United Kingdom

H. Fastl

Lehrstuhl für Mensch-Maschine Kommunikation, Technische Universität München, Germany

(Received 22 October 1999; revised 3 January 2000; accepted 8 March 2000)

Recently, it was demonstrated that the pitch strength of a stimulus denoted “AABB” differed from rippled noise (RN) despite the fact that their long-term spectra and autocorrelation functions are identical (Wiegrebe *et al.*, 1998). Rippled noise is generated by adding a delayed copy of Gaussian noise back to itself; AABB is generated by concatenating equal-duration, Gaussian-noise segments where every segment is repeated once. It was shown that a simple model based on a two-stage integration process separated by a nonlinear transformation explains the pitch-strength differences quantitatively. Here, we investigate how the spectral listening region influences pitch-strength differences between RN and AABB. Bandpass filtering the two stimuli with a constant bandwidth of 1 kHz revealed a systematic effect of center frequency. For relatively high pitches (corresponding to delays,  $d$ , of 4 or 5.6 ms, pitch strength differences between AABB and RN were absent when the pass band was between 0 and 1 kHz. When the pass band was between 3.5 and 4.5 kHz, pitch-strength differences were substantial. For lower pitches ( $d$  equal to or longer than 8 ms), AABB had a substantially greater pitch strength independent of the filter center frequency. The model presented in Wiegrebe *et al.* (1998) cannot capture these effects of center frequency. Here, it is demonstrated that it is possible to simulate the RN-AABB pitch-strength differences, and the effect of listening region, with a computer model of the auditory periphery. It is shown that, in an auditory model, pitch-strength differences are introduced by the nonlinear transformation possibly associated with half-wave rectification or mechano-electrical transduction. In this experimental context, however, the nonlinearity has perceptual relevance only when the differences in short-term fluctuations of AABB and RN are preserved in auditory-filter outputs. The current experiments relate the purely functional model introduced in the preceding paper to basic properties of the peripheral auditory system. The implication for neural time constants of pitch processing is discussed. © 2000 Acoustical Society of America. [S0001-4966(00)03406-8]

PACS numbers: 43.66.Ba, 43.66.Hg, 43.66.Jh, 43.66.Mk [SPB]

## INTRODUCTION

Both rippled noise (RN) and “AABB” noise are tonal stimuli derived from random noise by introducing temporal regularity into the waveform. Thus they both belong to the category of “regular-interval noises” (Yost, 1998). RN is produced by delaying a sample of random noise by “ $d$ ” ms and adding it back to itself; iterated rippled noise (IRN) is generated by iterating the delay-and-add process. Both have been used extensively to study mechanisms of pitch perception (Yost, 1996a, 1996b; Yost *et al.*, 1996; Patterson *et al.*, 1996; Wiegrebe and Patterson, 1999). Recently, Wiegrebe *et al.* (1998) used an AABB stimulus to investigate the time

constants associated with the pitches of regular-interval noises. AABB is generated by concatenating many independent noise segments of the same duration,  $d$ , with the constraint that every noise segment is repeated once before the next independent segment is concatenated. If we denote independent noise segments with letters A,B,C,..., the general form for the sound is AABBCC.... For convenience, this stimulus is called AABB. Wiegrebe *et al.* (1998) showed that AABB and RN have identical long-term spectra and autocorrelation functions when generated with the same  $d$ . Nevertheless, the pitch strengths of the two stimuli differ for pitches below about 250 Hz: AABB produces a stronger pitch than RN. To account for the pitch-strength differences, an autocorrelation model composed of a two-stage temporal integration mechanism separated by a nonlinear transforma-

<sup>a)</sup>Electronic mail: wiegrebe@zi.biologie.uni-muenchen.de

tion was introduced. The first stage was a running autocorrelation performed on the stimulus waveform. In this stage, temporal integration is imposed by an exponentially decaying integration window shifted over the waveform. The second stage relates the height of the peak in the running autocorrelation function with pitch strength in an exponentially nonlinear fashion. The third stage (second integration stage) involved averaging over the entire stimulus duration to produce a single pitch-strength value. The time constant of the first integration stage, the running autocorrelation of the stimulus waveform, was used to fit pitch-strength matching data between AABB and an IRN matching stimulus. The best-fitting time constant was 1.5 ms.

Although the model provided a good fit to the experimental data, it is not clear how the different model stages relate to properties of the auditory system. Before any kind of regularity analysis like running autocorrelation can be implemented in the auditory system, the stimulus is transformed by the outer and middle ear, cochlear nonlinear narrow-band filtering and the mechano-electrical transduction of the inner hair cells. The current study presents experiments whose results are incompatible with the purely functional modeling approach presented in Wiegrefe *et al.* (1998). However, it is demonstrated that the current experimental results as well as those in Wiegrefe *et al.* (1998) can be simulated in a computer model of the auditory periphery similar to that presented by Meddis and Hewitt (1991a, 1991b), Yost *et al.* (1996) and Meddis and O'Mard (1997). It is shown how the principal processing stages of the functional model in Wiegrefe *et al.* (1998) relate to specific processing stages in the peripheral human auditory system.

## I. EXPERIMENT I: THE EFFECT OF SPECTRAL LISTENING REGION ON RN-AABB DISCRIMINATION

In Wiegrefe *et al.* (1998) it was shown that despite the identity of the long-term spectra and autocorrelation functions, listeners could reliably discriminate between AABB and RN when  $d$  was longer than about 4 ms. As broadband stimuli were used, it was not possible to determine whether discrimination was supported by information available over the whole spectral range or if discrimination may depend on the spectral listening region. Here, we asked listeners to discriminate AABB from RN with bandpass filtered stimuli having constant bandwidth but different center frequencies. When using bandpass filtered stimuli, audible distortions may be generated below the pass band which can potentially interfere with the required experimental task. Because distortion products differ between RN and AABB, a small control experiment was performed where a random noise with a spectrum level only 10 dB lower than the stimulus spectrum level was added below the stimulus pass band. Introducing this low-pass masker, however, did not significantly alter performance in experiment I, and thus we are confident that the experimental results are not influenced by audible distortion products.

## A. Method

### 1. Stimuli

RN and AABB stimuli were produced exactly as described in the Introduction. The pitch, which corresponds to  $1/d$ , ranged from 31.25 to 250 Hz in half-octave steps. Within each run,  $d$  was the same for RN and AABB. Stimuli were generated in the time domain and then, using the fast Fourier transform, translated into the frequency domain. By setting the magnitude spectrum outside the pass band to zero and then performing an inverse Fourier transform, the bandpass filtered stimuli were obtained. The spectrum level was 30 dB; the pass-band width was fixed at 1 kHz; the center frequency was 500 Hz, 1 kHz, 2 kHz or 4 kHz. The stimulus duration was 327.68 ms, including 20-ms, cosine-squared ramps. Stimuli were generated and played back digitally using the Tucker Davis System II and software delay lines. The sampling rate was 25 kHz. All sounds were presented diotically via AKG K240 DF headphones. Listeners were tested individually in a double-walled sound-attenuating booth.

### 2. Procedure

As in Wiegrefe *et al.* (1998), psychometric functions for the discrimination of AABB from RN were obtained using a two-alternative, three-interval, forced-choice task. The first interval was always a RN. Either the second or the third interval contained an AABB stimulus to be discriminated from RN, whereas the remaining interval contained another RN. The listeners were instructed to choose the interval that differed from the other two. The listeners were provided with visual feedback. Each point on the psychometric function represents the mean of 99 trials obtained in three runs. Within a run, the value of  $d$  for each trial was chosen randomly from the full range with the constraint that each value of  $1/d$  was presented 33 times. Within each run, the bandpass center frequency was fixed at one of the four values given above. The different center-frequency conditions were presented in a nonsystematic order.

### 3. Listeners

Three listeners aged between 26 and 33 years took part in all experiments; L1 was the first author, L2 was the second author. Listeners L2 and L3 were given training until stable performance was achieved. Listener L1 had taken part in the experiments in Wiegrefe *et al.* (1998) and needed no further training. All listeners had absolute thresholds at all audiometric frequencies within 10 dB of the 1969 ANSI standard.

## B. Results

The results were very similar across listeners, and thus the results were pooled. The psychometric functions for the AABB-RN discrimination are plotted as a function of  $1/d$  in Fig. 1. Error bars represent cross-listener standard errors. When  $1/d$  is below about 100 Hz, discrimination of AABB from RN is above 80%, irrespective of filter center frequency. When  $1/d$  is 125 Hz and above, discrimination performance deteriorates with increasing  $1/d$  and there is a systematic effect of center frequency. The lower the center

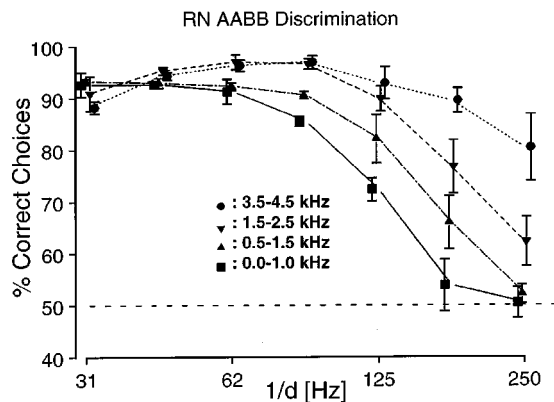


FIG. 1. Discrimination of AABB stimuli from RN stimuli as a function of the fundamental frequency ( $1/d$ ). Different line styles and symbols represent discrimination performance for bandpass filtered stimuli with a constant bandwidth of 1 kHz and a center frequency of 500 Hz (squares and solid line), 1 kHz (upward triangles and dash-dotted line), 2 kHz (downward triangles and dashed line) and 4 kHz (circles and dotted line). Error bars represent cross-listener standard errors.

frequency of the pass band, the worse the discrimination performance. When  $1/d$  is 250 Hz, discrimination is at chance levels for a center frequency of 500 Hz, but it is on average above 80% for a center frequency of 4 kHz.

## II. EXPERIMENT II: PITCH-STRENGTH MATCHING FILTERED AABB WITH EQUALLY FILTERED IRN STIMULI

Experiments II and III of Wiegrebe *et al.* (1998) showed that the basis of the RN-AABB discrimination with broadband stimuli is a difference in pitch strength. To establish if pitch-strength differences may also serve as a possible discrimination cue for the bandpass filtered stimuli, the pitch strength of the filtered AABB stimuli was matched with that of equally filtered IRN stimuli having the same  $d$ . Over trials, the actual value of  $d$  was randomized within  $\pm 3\%$  of the denoted  $d$  value to help to focus the listeners' attention to pitch strength, not pitch. IRN was produced in the add-same configuration (Yost, 1996a, 1996b) with 16 iterations. The pitch strength of RN and IRN stimuli can be continuously varied by changing the gain in the delay-and-add loop (Yost, 1996b; Wiegrebe *et al.*, 1998).

### A. Method

#### 1. Stimuli

Stimulus duration, presentation level and filter conditions were the same as in the previous experiment. AABB was generated as described above. IRN was generated with the same  $d$  and 16 iterations. Pitch-strength matches were obtained for  $1/d$ 's between 31.25 and 250 Hz in half-octave steps.

#### 2. Procedure

As in Wiegrebe *et al.* (1998), a two-alternative, forced-choice procedure was used to determine the pitch-strength matches. One randomly chosen interval contained a sample of filtered AABB; the other interval contained an equally filtered IRN. At the beginning of each run, the IRN gain was

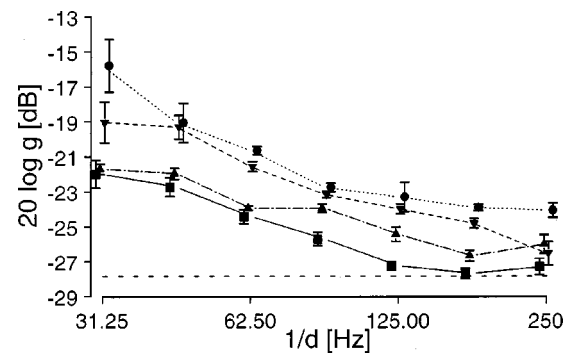


FIG. 2. Pitch-strength matches of AABB test stimuli with IRN matching stimuli. IRN was produced with 16 iterations and its gain ( $g$ ) was manipulated to match the IRN pitch strength to the AABB pitch strength. An increase in  $g$  corresponds to an increase in pitch strength. The different filter conditions are represented in the same format as in Fig. 1. The horizontal dashed line represents the pitch-strength matches as they would be obtained for RN test stimuli according to Yost (1996b). Error bars represent cross-listener standard errors. Generally, pitch-strength differences between AABB and RN decrease with increasing  $1/d$ . However, for a pass band from 3.5 to 4.5 kHz (dotted line and filled circles), there is still a pitch-strength difference when  $1/d$  equals 250 Hz where there is no pitch-strength difference when the stimuli are lowpass filtered at 1 kHz (solid line and filled squares).

set to unity so that the IRN pitch strength was initially higher than that of the AABB. Listeners were instructed to choose the stimulus with the greater pitch strength. Feedback was not provided. When they chose the IRN stimulus, the IRN gain was reduced for the next trial, when they chose the AABB stimulus, the IRN gain was increased for the next trial. When listeners switched from choosing the IRN to the AABB or vice versa, a reversal was registered. The step size for the IRN gain initially was 3 dB; it was reduced to 2 dB after the third reversal and to 1 dB after the fifth reversal. The average gain from reversals six to eleven was taken as the matching gain for that run. For each condition, four runs were obtained and their matching points were averaged. The listeners were the same as in the previous experiment. They were trained until stable performance was reached, which typically occurred after only a few runs for each value of  $d$ .

### B. Results and discussion

The IRN gain required to match the IRN pitch strength to that of equally filtered AABB with the same  $d$  is shown as a function of  $1/d$  in Fig. 2. Again performance across listeners was very consistent and so the data were pooled. Error bars represent cross-listener standard errors. Results for the different filter conditions are shown with the same symbols and line styles as in Fig. 1. The horizontal dashed line at an IRN gain of  $-28.8$  dB represents the pitch-strength match estimated for a RN test stimulus instead of an AABB test stimulus. The estimate is based on Yost's (1996b) finding that RN and IRN stimuli have the same pitch strength when the height of the first peak in the long-term autocorrelation function ( $h_1$ ) is the same. The pitch-strength difference between AABB and RN is represented as the difference between the IRN matching gain for AABB test stimuli and the horizontal dashed line. For all filtering conditions, there is a consistent trend toward lower matching gains with  $1/d$ . In

the lowest filter condition (0 to 1 kHz, solid line and filled squares), the pitch-strength difference was 6 dB when  $1/d$  was 31.25 Hz and reduces to zero for  $1/d$  of 125 Hz and above. In the highest filter condition (3.5 to 4.5 kHz, dotted line and filled circles), the pitch-strength difference was 12 dB at  $1/d$  equals 31.25 Hz, and it was still about 5 dB even for  $1/d$  equals 250 Hz. The results indicate that pitch-strength differences between RN and AABB may have served as a cue for the RN-AABB discrimination measured in experiment I. For broadband stimuli, this has been shown in Wiegrebe *et al.* (1998). For values of  $d$  and filter center frequency where pitch-strength differences are absent in experiment II, discrimination performance in experiment I is at chance levels. For two reasons, however, it must not be assumed that the pitch-strength matching curves of experiment II resemble the psychometric functions of experiment I: First, the psychometric functions for the discrimination are likely to be limited by ceiling effects, and second, the relation between pitch strength and IRN gain is monotonic but it is not linear (cf. Yost, 1996b). It is important to note that the pitch-strength matching paradigm only allows for comparisons of pitch strengths within one value of  $d$  and within one frequency band. As stated above, if a RN had been used as a test stimulus, matches would be fixed at about  $-28.8$  dB. This does not imply, however, that the pitch strength of RN is independent of  $1/d$ . It is well documented that pitch strength of RN stimuli increases with increasing  $1/d$  at least for the range of  $1/d$  tested here (Yost and Hill, 1978). Additionally, high-pass filtering complex harmonic stimuli decreases their pitch strength as shown, for example, by Houtsma and Smurzynski (1990). This is also supported by experiments on the dominance region of pitch (Plomp, 1967). Both these sets of experiments underline the perceptual dominance of the low harmonics or the low spectral peaks in regular-interval noises. The current experiments and the simulations in Sec. III only try to capture pitch-strength differences between RN and AABB within each of the experimental conditions of  $d$  and center frequency, not pitch strength across the different conditions.

### III. MODELING THE EFFECT OF LISTENING REGION ON THE PITCH-STRENGTH DIFFERENCE OF REGULAR-INTERVAL NOISES

#### A. The functional model and the effect of spectral listening region

Wiegrebe *et al.* (1998) used a functional model to account for RN-AABB pitch-strength differences; a block diagram of the model is shown in the left panel of Fig. 3. In this functional model, the waveform is subjected to a running normalized autocorrelation with a time constant of 1.5 ms. The correlation is calculated as a function of time, but only for one correlation lag which equals  $d$ , because only at this lag, a correlation peak ( $h1$ ) occurs in the long-term autocorrelation of RN and AABB. The result of this processing stage is called  $Rh1(t)$ ; it is subjected to an expansive non-linear transformation (exponentiation). The result is then averaged over the stimulus duration to obtain a scalar value for the pitch-strength estimate. In this section, this model is

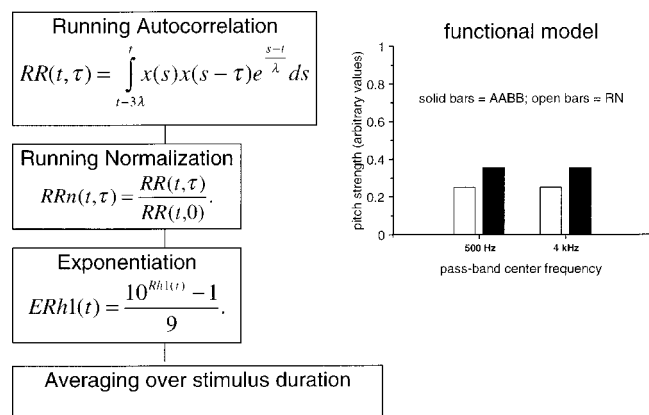


FIG. 3. Left panel: Block diagram of the functional model used to explain the AABB-RN pitch-strength differences in Wiegrebe *et al.* (1998). See text for a more detailed description of the model. The right panel shows simulated pitch-strength values for RN (open bars) and AABB (filled bars) for bandpass center frequencies of 500 Hz and 4 kHz. The simulation results are independent of center frequency and thus not suitable to simulate the results of experiments I and II.

tested to determine whether it can accommodate the effect of the listening region. To test the functional model, average pitch strength was calculated for 100 samples of RN and AABB;  $d$  was fixed at 4 ms and the stimuli were filtered with a center frequency of either 500 Hz or 4 kHz and a constant bandwidth of 1 kHz. The results of this simulation are shown in the right panel of Fig. 3. For both filter conditions, AABB with  $d$  equals 4 ms has a higher pitch strength than RN with the same  $d$ . Thus the model does not predict the effect of filter center frequency.

#### B. Toward an auditory model of pitch strength

In recent years, multichannel computer models of the human auditory periphery have been successfully used to simulate psychoacoustical experiments on pitch perception (Slaney and Lyon, 1990; Meddis and Hewitt, 1991a, 1991b, 1992; Patterson *et al.*, 1995; Yost *et al.*, 1996; Meddis and O'Mard, 1997; de Cheveigne, 1998). Common to these models is the desire to simulate the spectrotemporal distribution of auditory-nerve firing. The spiking activity or spike probability pattern is then subjected to an analysis of temporal regularity, typically implemented by channel-wise autocorrelation. The autocorrelation is typically implemented as a running autocorrelation with a time constant between 2.5 and 10 ms. We investigated the representation of RN and AABB stimuli in an auditory model whose peripheral preprocessing is identical to that of Meddis and O'Mard (1997).

Meddis and Hewitt (1991a, 1991b) and Meddis and O'Mard (1997) tested their model with a large variety of pitch experiments; some of which are related to pitch strength: the model successfully predicts the dominance region of harmonic complexes as determined by Plomp's (1967) paradigm, the sensitivity to mistuning of individual harmonics (Moore *et al.*, 1985), and the increase in threshold  $f0$ -differences with increasing listening region for band-pass filtered harmonic complexes (Houtsma and Smurzynski, 1990). The procedure of all these experiments involves the comparison of complex tones *within* the same frequency re-

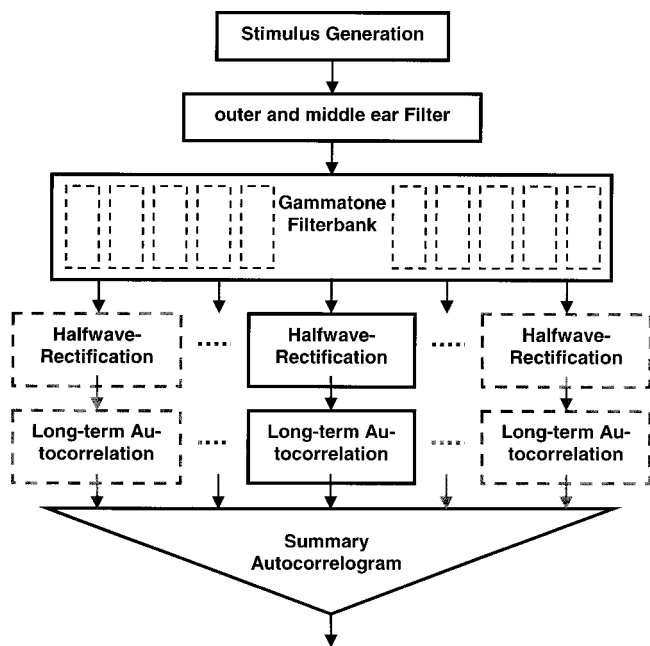


FIG. 4. Block diagram of the auditory model used to simulate pitch-strength differences between AABB and RN stimuli. The model is very similar to that of Meddis and O'Mard (1997) apart from the use of simple half-wave rectification instead of the Meddis hair cells and a long-term autocorrelation instead of a running autocorrelation with a 10-ms time constant.

gion and with similar  $f_0$ s. The same is true for the current experiments where pitch strength is compared for AABB and RN with the same  $d$  (i.e., with similar pitch) and in the same frequency region. In the next section, it will be shown the model can predict pitch-strength differences correctly, although the model is not able to predict absolute pitch strength and its changes across different  $d$ 's and frequency regions.

Outer and middle ear filtering is implemented by a second-order bandpass filter with cutoffs of 450 and 8500 Hz: spectral analysis is implemented by a bank of 60 gammatone filters equally spaced on a log frequency scale between 100 and 8000 Hz. Mechano-electrical transduction was very coarsely simulated by simple half-wave rectification. A bank of Meddis hair cells (Meddis, 1988) was shown to yield very similar results to half-wave rectification. Unlike Meddis and O'Mard (1997), autocorrelation of the rectifier output was implemented as a long-term autocorrelation, i.e., the width of the correlation window equaled the stimulus duration. The autocorrelograms were summed over frequency

channels to produce a summary autocorrelogram (SACG). A block diagram of the model is shown in Fig. 4.

### C. Simulation results

Yost (1996b) showed that for RN and IRN stimuli, the position of the first peak in the autocorrelation of the stimulus indicates the pitch of the stimulus and its height is monotonically related to pitch strength. The same has been suggested for the first peak in the SACG (Yost *et al.*, 1996; Patterson *et al.*, 1996); however, this hypothesis has not been quantitatively tested, so far. For the current simulations, it is assumed that pitch strength is monotonically related to the height of the first peak in the SACG. In the following, we compare the SACG peak heights for RN and AABB stimuli in the same stimulus conditions as used in experiment II. For each filter condition and each delay, 100 model responses were calculated and their peak heights averaged. In the left panel of Fig. 5, the SACG peak heights are plotted as a function of  $1/d$  for the four filter conditions, in the same format as for the experimental data. Simulation results for RN are plotted with solid lines; results for AABB stimuli are plotted with dashed lines. The right panel of Fig. 5 shows the SACG peak-height difference functions between RN and AABB.

The left and right panel of Fig. 5 represent the shortcomings and success of the model, respectively. The absolute SACG peak heights in the left panel are rather independent of  $d$ , suggesting a pitch strength independent of  $d$ . This is in contrast to the experimental results of Yost and Hill (1978), who found that when  $d$  increases from 5 ms, pitch strength decreases and is almost entirely absent for  $d > 20$  ms. Experiments on the dominance region and mistuning of individual harmonics showed that the strongest pitch is always mediated by the low harmonics (typically between two and five). The SACG peak heights in the left panel of Fig. 5, however, suggest that pitch strength is always lowest in the lowest frequency band and increases when the listening band is shifted to between 1.5 and 2.5 kHz. Thus the model in its current form is clearly not able to predict absolute pitch strength and its changes across different frequency regions and different  $d$ 's.

The general shapes of the peak-height difference functions in the right panel of Fig. 5, however, are consistent with the pitch-strength differences obtained in experiment II and shown in Fig. 2. For low values of  $1/d$ , pitch-strength

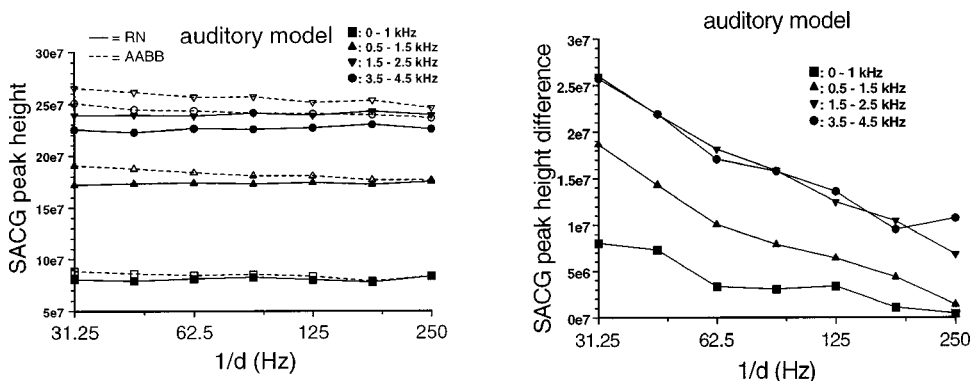


FIG. 5. Left panel: Height of the first peak in the summary autocorrelogram for AABB (dashed lines and open symbols) and RN stimuli (solid lines and filled symbols) as a function of  $1/d$ . The different symbols represent the different filter conditions. Peak-height differences between AABB and RN stimuli are plotted in the right panel. They are in good qualitative agreement with the pitch-strength differences obtained in experiment II (Fig. 2).

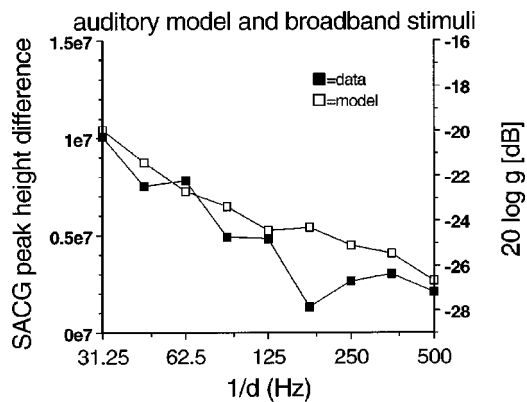


FIG. 6. SACG peak-height differences between broadband AABB and RN. The shape of the peak-height differences (open squares referring to the left Y-axis) are in reasonable *qualitative* agreement with the experimental data of experiment II in Wiegrebe *et al.* (1998) (filled squares referring to the right Y-axis). Note that the left and right Y-axes are not linearly related.

differences and SACG peak-height differences are large. With increasing  $1/d$ , pitch-strength differences and SACG peak-height differences decrease. When the pass band is from 0 to 1 kHz and  $1/d$  is 250 Hz, pitch-strength differences and peak-height differences become negligible. Pitch-strength differences are still substantial, however, when the pass band is from 3.5 to 4.5 kHz. To simulate the results of experiment II in Wiegrebe *et al.* (1998), the model was presented with broadband stimuli with  $1/d$  ranging from 31.25 to 500 Hz. The simulation results are shown in Fig. 6. Again, SACG peak-height differences are in good qualitative agreement with the experimental data of experiment II in Wiegrebe *et al.* (1998).

#### D. Crucial processing stages of the auditory model and their differential effect on RN and AABB stimuli

In this section, the manipulations imposed on the waveforms at the different processing stages of the auditory model are reviewed and the processing stages crucial to the differential representation of AABB and RN are highlighted. It will be shown that in the model, the cochlear filtering, together with the nonlinear transformation implemented by half-wave rectification, are responsible for the differences in the coding of AABB and RN periodicity.

##### 1. The effect of cochlear filtering

Cochlear filtering is approximated by a gammatone filterbank in the auditory model. The equivalent rectangular bandwidth (ERB) of a gammatone filter is

$$\text{ERB} = 24.7 + 0.108 * F_c.$$

The filter bandwidth grows with increasing filter center frequency  $F_c$  (in Hz). As it does, the filter impulse response becomes shorter, reducing the effective integration time of the filter. The effect of this frequency-dependent integration on the temporal structure in RN and AABB stimuli can be observed by comparing the running autocorrelation of the stimuli with a very short time constant, 1.5 ms, before filtering and after filtering with a gamma tone filter. As shown in Wiegrebe *et al.* (1998), the peak in the normalized running

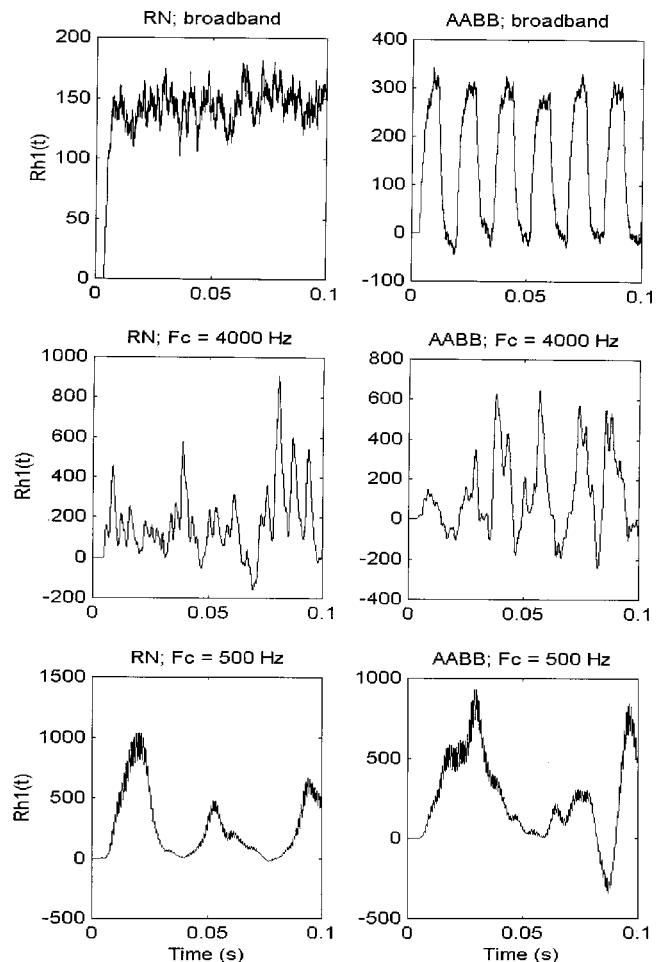


FIG. 7.  $Rh1(t)$  as a function of time for RN (left column) and AABB stimuli (right column). In the upper row,  $Rh1(t)$  is calculated on the raw waveform; in the middle row, stimuli were first filtered with a gammatone filter with  $F_c$  equals 4 kHz; in the lowest row,  $F_c$  equals 500 Hz. After the 4-kHz filtering, the  $Rh1(t)$  oscillations of AABB are still preserved whereas after the 500-Hz filtering where the filter bandwidth is only about one-sixth of the 4-kHz filter, the  $Rh1(t)$  oscillations are strong but random like those of RN. Note that, unlike in Wiegrebe *et al.* (1998),  $Rh1(t)$  is not normalized.

autocorrelation,  $Rh1(t)$ , fluctuates randomly around its average for RN stimuli whereas, for AABB stimuli, it oscillates with a period equal to two times  $d$ .  $Rh1(t)$  for RN stimuli with  $d$  equals 8 ms is shown in the left column of Fig. 7;  $Rh1(t)$  for AABB with  $d$  equals 8 ms is shown in the right column.  $Rh1(t)$  is shown for unfiltered stimuli in the upper row, for stimuli filtered with a 4-kHz gammatone filter in the middle row and for stimuli filtered with a 500-Hz gammatone filter in the lower row. The  $Rh1(t)$  oscillations for AABB are clearly visible for the unfiltered stimuli (upper right panel) and are still reasonably well preserved after the 4-kHz gammatone filter (middle right panel). However, for the AABB stimulus filtered with the 500-Hz gammatone filter (lower right panel), the period of the oscillations is smeared by the long integration time of the 500-Hz filter. It is important to point out that the degree of smearing does not depend on the absolute filter bandwidth but on the relationship between filter bandwidth and  $d$ . Thus although the oscillations in the autocorrelation of AABB are smeared out in the 500-Hz gammatone filter when  $d$  is 8 ms, the oscillations

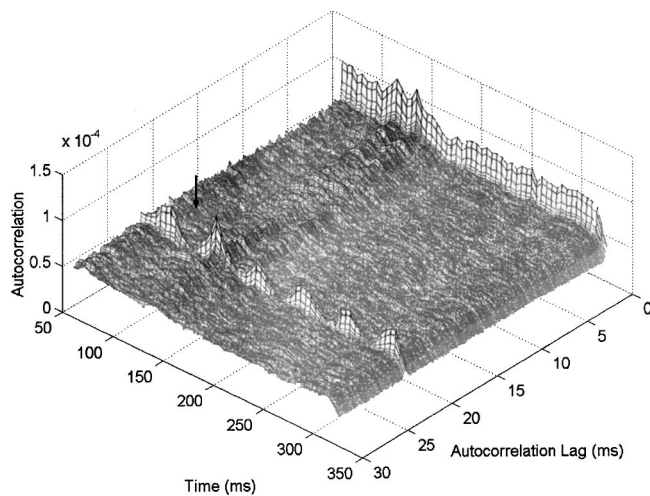


FIG. 8. Summary autocorrelograms (SACG) as a function of time for an AABB stimulus with  $d$  equal to 22.4 ms. The first peak in the SACG (arrow) fluctuates strongly and periodically over time. However, these fluctuations are not perceived as shown in Wiegrebe *et al.* (1998).

would still be preserved in the same filter when  $d$  is say 16 or 32 ms. In the next section, we show that the preservation of  $Rh1(t)$  fluctuations interacts with the nonlinear transformation (the half-wave rectification) following the gammatone filters.

## 2. The effect of half-wave rectification

As noted in the Introduction, the long-term autocorrelation functions of RN and AABB are identical. However, when the stimuli are half-wave rectified, this is no longer the case. Following half-wave rectification, the height of the first peak in the normalized stimulus autocorrelation function ( $h1$ ) is higher for AABB than for RN. The mathematical evidence for this interaction of half-wave rectification and temporal regularity is not trivial but the effect is found very consistently. When averaged over 10 stimuli,  $h1$  for AABB was 0.67, whereas  $h1$  for RN was only 0.60; the standard deviations were 0.0053 and 0.0119, respectively. The differential effect of half-wave rectification disappears, however, when the temporal oscillations of AABB are smeared out which, as shown above, happens when the oscillation period of the stimulus is in the same range as or shorter than the auditory-filter impulse response.

Thus the cumulative effect of frequency-dependent filter width and nonlinearity like half-wave rectification can account for the frequency-dependent pitch-strength differences found experimentally.

## 3. Running autocorrelation versus long-term autocorrelation

The third stage of the auditory model involves long-term autocorrelation instead of running autocorrelation with either a 10-ms time constant (Meddis and O'Mard, 1997) or a 2.5-ms time constant (Meddis and Hewitt, 1991a, 1991b). This was done for the following reason: Even with a 10-ms time constant, the running autocorrelation for AABB with a long  $d$  fluctuates strongly and periodically over time. This is illustrated in Fig. 8, where the summary autocorrelogram for

an AABB stimulus with  $d$  equals 22.4 ms is plotted as a function of time. Note, however, that these fluctuations are not perceived. This was demonstrated in experiment III of Wiegrebe *et al.* (1998). When  $d$  was 22.4 ms and the IRN pitch-strength was matched to that of AABB, the two stimuli were no longer discriminable although the autocorrelation fluctuations were periodic and large for AABB versus random and small for RN. Meddis and O'Mard (1997) and Meddis and Hewitt (1991a, 1991b) did not actually calculate *running* autocorrelation; rather they calculated the short-term autocorrelation of a stimulus snapshot at the end of the stimulus. Figure 8 shows that this approach is not applicable to AABB stimuli because, depending on the phase at the end of the stimulus, a pitch-strength estimate could be well above or well below the pitch strength that listeners hear. Thus a considerably longer integration time constant is required if running autocorrelation is used. The value of this time constant is not known and so, for the time being, we used long-term autocorrelation where the integration time is only limited by the stimulus duration.

## IV. SUMMARY AND CONCLUSIONS

The current experiments like those in Wiegrebe *et al.* (1998) reveal pitch-strength differences between AABB and RN stimuli, which represent a challenge for models of pitch perception because the long-term characteristics of AABB and RN are identical. The current experiments show that pitch-strength differences of AABB and RN stimuli depend systematically on the center frequency of the bandpass filter. When  $d$  is 4 ms, AABB-RN discrimination is poor and pitch-strength differences are small when the stimuli are lowpass filtered at 1 kHz, whereas discrimination is good and pitch-strength differences are large when the stimuli are filtered between 3.5 and 4.5 kHz. This effect cannot be explained with a purely waveform-based model as in Wiegrebe *et al.* (1998). Simulations with a model of the auditory periphery similar to Meddis and O'Mard (1997) show that auditory filtering and half-wave rectification affect the representation of RN and AABB differentially. The long-term autocorrelation of the AABB and RN stimuli processed in this way and summed up over frequency channels yields estimates of pitch-strength differences in good agreement with the current data. However, it must be pointed out that the model in its current form is not able to predict changes of pitch strength across different  $d$ 's and frequency regions. The current experiments, together with those of Wiegrebe *et al.* (1998), question the value of running autocorrelation with a fixed time constant in the range of 10 ms as suggested by Meddis and O'Mard (1997). With such a short time constant, the model would predict that AABB pitch-strength for higher  $d$ 's (lower pitches) would fluctuate periodically, which it does not (cf. Wiegrebe *et al.*, 1998, experiment III).

Based on the current experiments and simulations, it is possible to relate the processing stages of the functional model in Wiegrebe *et al.* (1998) to properties of the peripheral auditory system. The first integration stage implemented by the integration window of the running autocorrelation corresponds to the frequency-dependent temporal integration of the cochlear filters. The nonlinear transformation in the



functional model corresponds to the half-wave rectification in auditory mechano-electrical transduction. The second integration stage is implemented by averaging over the stimulus duration in both the functional and the auditory model. Despite the fact that the temporal properties of the second integration stage still remain obscure, the current research shows that pitch strength is substantially influenced not only by the filtering properties but also by the nonlinear transformations of the peripheral auditory system.

## ACKNOWLEDGMENTS

This study was supported by the Medical Research Council (G9703469), U.K. by a research grant from the "Deutsche Forschungsgemeinschaft" to the first author, and by a grant from the "Prof. Dr.-Ing. Erich Müller Stiftung" to the second author.

- de Cheveigné, A. (1998). "Cancellation model of pitch perception," *J. Acoust. Soc. Am.* **103**, 1261–1271.
- Houtsma, A. J. M., and Smurzynski, J. (1990). "Pitch identification and discrimination for complex tones with many harmonics," *J. Acoust. Soc. Am.* **87**(1), 304–310.
- Meddis, R. (1988). "Simulation of auditory-neural transduction: Further studies," *J. Acoust. Soc. Am.* **83**, 1056–1063.
- Meddis, R., and Hewitt, J. (1991a). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification," *J. Acoust. Soc. Am.* **89**(6), 2866–2882.
- Meddis, R., and Hewitt, J. (1991b). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity," *J. Acoust. Soc. Am.* **89**(6), 2883–2894.
- Meddis, R., and Hewitt, J. (1992). "Modeling the identification of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **91**(1), 233–245.
- Meddis, R., and O'Mard, L. (1997). "A unitary model of pitch perception," *J. Acoust. Soc. Am.* **102**(3), 1811–1820.
- Moore, B. J. C., Glasberg, B. R., and Peters, R. W. (1985). "Relative dominance of individual partials in determining the pitch of complex tones," *J. Acoust. Soc. Am.* **77**, 1853–1860.
- Patterson, R. D., Allerhand, M., and Giguere, C. (1995). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.* **98**, 1890–1894.
- Patterson, R. D., Handel, S., Yost, W. A., and Datta, A. J. (1996). "The relative strength of the tone and noise components in iterated rippled noise," *J. Acoust. Soc. Am.* **100**, 5.
- Plomp, R. (1967). "Pitch of complex tones," *J. Acoust. Soc. Am.* **41**, 1526–1533.
- Slaney, M., and Lyon, R. F. (1990). "A perceptual pitch detector," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Albuquerque, New Mexico, pp. 357–360.
- Wiegrefe, L., Patterson, R. D., Demany, L., and Carlyon, R. P. (1998). "Temporal dynamics of pitch strength in regular interval noises," *J. Acoust. Soc. Am.* **104**(4), 2307–2313.
- Wiegrefe, L., and Patterson, R. D. (1999). "The role of envelope modulation in spectrally unresolved iterated rippled noise," *Hear. Res.* **132**, 94–108.
- Yost, W. A. (1996a). "Pitch of iterated rippled noise," *J. Acoust. Soc. Am.* **100**(1), 511–518.
- Yost, W. A. (1996b). "Pitch strength of iterated rippled noise," *J. Acoust. Soc. Am.* **100**(5), 3329–3335.
- Yost, W. A., and Hill, R. (1978). "Strength of the pitches associated with ripple noise," *J. Acoust. Soc. Am.* **64**, 485–492.
- Yost, W. A., Patterson, R. D., and Sheft, S. (1996). "A time domain description for the pitch strength of iterated rippled noise," *J. Acoust. Soc. Am.* **99**, 1066–1078.
- Yost, W. A. (1998). "Auditory processing of sounds with temporal regularity: Auditory processing of regular interval stimuli," in *Psychophysical and Physiological Advances in Hearing, Proceedings 11th International Symposium on Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

# Relative loudness perception of low and high frequency sounds in the open and occluded ear

Gitte Keidser, Rickard Katsch, Harvey Dillon, and Frances Grant

*The National Acoustic Laboratories, 126 Greville Street, Chatswood, NSW 2067, Australia*

(Received 20 July 1999; revised 2 March 2000; accepted 17 March 2000)

A comparison of published equal loudness contours indicates that different shapes are obtained at a comfortable level when the measurements are done in an occluded ear than when they are done in an open ear, even though all measurements are expressed as dB SPL at the eardrum. This paper presents the result from a loudness balancing test which confirms this observation. Eleven normal-hearing listeners balanced the level of a 500- and a 3000-Hz octave band babble-noise to the level of a 1500-Hz octave band babble-noise. The balancing test was completed in open and occluded ears using a loudspeaker and a hearing aid receiver, respectively. A probe tube microphone was used to measure the actual levels presented near the individual's eardrum. The results show that an average of 10 dB higher level was selected for the 500-Hz octave band when listening with the occluded ear than when listening with the open ear. A large range of factors is discussed, but no physical explanation for the discrepancy was found. The findings could have consequences for psychoacoustic experiments and for the use of loudness measurements for hearing aid prescription.

© 2000 Acoustical Society of America. [S0001-4966(00)05706-4]

PACS numbers: 43.66.Cb, 43.66.Lj, 43.66.Yw [RVS]

## INTRODUCTION

Loudness is the psychological perception of a sound presented with certain intensity. For many years loudness has been used in one way or another to estimate the appropriate gain of a hearing aid. For example, Watson and Knudsen (1940) suggested that the frequency response should match the individual most comfortable equal loudness curve, and more recently it has been proposed that nonlinear hearing aids should normalize loudness for each frequency and each input level (e.g., Cox, 1995). Loudness measurements are also commonly used in psychoacoustic experiments to study phenomena like loudness summation and masking, and to establish references, such as the minimum audible field (MAF) response. Across the bulk of literature that presents loudness data, measurements are conducted using any of the usual transducers like loudspeakers, supra-aural headphones, or insert earphones. Throughout the years many comparisons of loudness data across studies have been made, and in the case of hearing aid fitting, hearing aid listening is sometimes compared to free field listening by normal-hearing people. In doing so it is assumed that loudness perception is independent of the transducer used to obtain the loudness data. Unfortunately, this is not so.

The authors first became aware of the issue when comparing two equal loudness contours measured at the National Acoustic Laboratory (NAL) (see Fig. 1). One equal loudness contour was obtained in a balancing test conducted in the free field. Using a 1500-Hz octave band as the reference, normal-hearing listeners were asked to adjust the levels of octave bands from 250 to 4000 Hz to match the overall level of the reference (unpublished data). The other equal loudness contour was derived from a set of loudness growth functions for normal-hearing listeners (Keidser *et al.*, 1999). Figure 1 shows a clear difference in the shape of the two loudness contours. When the levels are converted to dB SPL at the

eardrum, using the correction figures from Bentler and Pavlovic (1989), the result from the balancing test suggests a 0.3-dB difference in equal loudness level between 500 and 3000 Hz whereas the scaling test produce a difference of -9.5 dB. There were many procedural differences between the experiments that produced the equal loudness contours displayed in Fig. 1, but a few pilot tests quickly narrowed the contributing factor down to the transducer. A loudspeaker was used in the balancing test and a hearing aid receiver was used in the loudness scaling test. The hearing aid receiver was mounted in a behind-the-ear (BTE) hearing aid case connected to a Libby horn and a foam ear tip. That is, in the balancing test subjects listened with open ears, but in the loudness scaling test the test ear was occluded.

Many experiments have been reported in the literature that measure loudness contours in the comfortable range for normal-hearing listeners (Beattie *et al.*, 1997; Christen and Byrne, 1980; Cox *et al.*, 1997; Holte and Margolis, 1987; Humes *et al.*, 1996; ISO R226, 1961; Keidser *et al.*, 1999; Khalil, 1989; Morgan and Dirks, 1974; Pollack, 1952; Richards, 1975; Ricketts and Bentler, 1996). These experiments were conducted using various transducers (loudspeaker, headphone, insert earphone, or hearing aid receiver), narrow band test stimuli (from pure tones to 1/1 octave noise or speech), and techniques (loudness scaling, loudness balancing, or an adaptive or tracking procedure). The effect of using different transducers is that listening is either performed via an open ear (loudspeaker), a cushioned ear (headphones), or an occluded ear (insert earphones and hearing aid receivers). From all these studies the mean levels at 500 and 3000 Hz were extracted. Either the mean level at 3000 Hz was presented or it was interpolated from the mean levels measured at 2000 and 4000 Hz. This was considered a valid procedure because all studies that presented data at each of the frequencies 2000, 3000, and 4000 Hz suggested that the

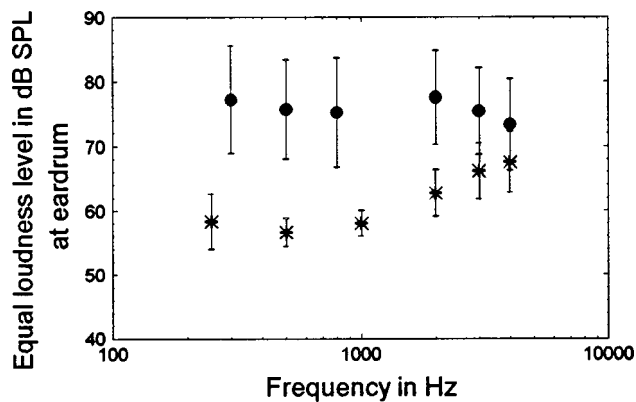


FIG. 1. Normal equal loudness contours measured in the open ear ( $N = 10$ ) using a loudness balancing procedure (\*), and in the occluded ear ( $N = 25$ ) using a loudness scaling procedure (●). The bars show  $\pm$  one standard deviation.

slope of the contour across these frequencies is relatively constant. All mean levels were converted to SPL at the eardrum using average correction figures (Bentler and Pavlovic, 1989). Together with the loudness data obtained at NAL, they were averaged for each listening condition of the ear. Figure 2 shows the results, which demonstrate a significant difference in the relative equal loudness levels between 500 and 3000 Hz, dependent on whether measurements were obtained in the open or occluded ear. The average differences in loudness perception of low- and high-frequency sounds presented in the open and in the occluded ear are  $-8.5$  and  $0.7$  dB, respectively, resulting in a difference between the two test conditions of  $9.2$  dB. As several studies have shown that different types of stimuli with a similar bandwidth produce similar results, and that the same stimulus with various bandwidths mainly produces a shift in overall level (e.g., Cox, 1989; Keidser *et al.*, 1996; Ricketts and Bentler, 1996), the variation in type of stimulus should not greatly influence the loudness contour. When the same test stimulus and transducer were used, Khalil (1989) demonstrated that similar loudness contours were measured with an adaptive test procedure and a scaling test, and a pilot study conducted at NAL revealed no difference in loudness contours produced in a balancing test and a scaling test. Therefore, it appears that

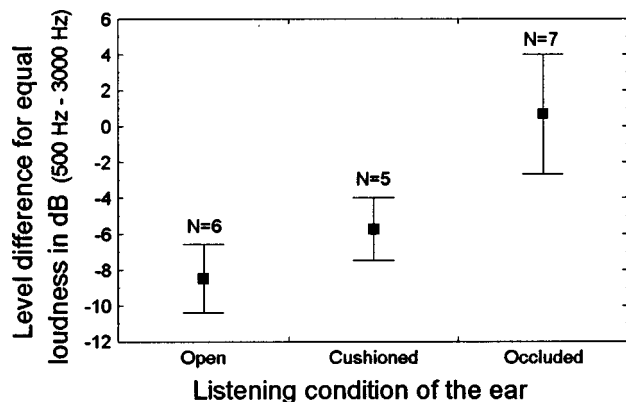


FIG. 2. The average difference in equal loudness levels measured at 500 and 3000 Hz in open, cushioned, or occluded ears. Data were extracted from  $N$  different experiments using a variety of test stimuli and test procedures (see text). The bars show  $\pm$  one standard deviation.

neither the type of stimulus nor the choice of technique has a significant influence on the loudness contour.

However, there are still many procedural differences among the studies referred to earlier and in many ways the observation has some resemblance to the “missing 6 dB” case reported 50 years ago (e.g., Beranek, 1949; Munson and Wiener, 1952). Data presented then suggested that when subjects were asked to balance loudness of low-frequency tones presented alternately from an earphone and from a loudspeaker, approximately 6 dB higher sound pressure levels were selected when the source was the earphone. Rudmose (1982) later proved that nothing was really missing. The discrepancy could be explained by factors such as the formal procedure for performing the balancing test, mechanical vibration coupled to the subject, transducer distortions, and the distance between loudspeaker and listener during free-field testing. A further issue is the exact sound pressure level presented at the individual’s eardrum in each test situation (Kil-lion, 1978).

This paper presents data from a study that aimed at investigating whether the transducer has an influence on the perception of the relative loudness of low- and high-frequency sounds in normal ears. The data from this experiment confirm earlier findings. In the light of the “missing 6 dB” case the possible influence of a large range of factors on the data is discussed.

## I. METHOD

### A. Subjects

Eleven volunteers with normal hearing were accepted for the experiment. All subjects had hearing threshold levels of 10 dB HL, or less, across the frequency range 250–4000 Hz, and an otoscopic examination on both ears revealed no abnormalities.

### B. Instrumentation

Babble-noise was played back from a digital audio recorder (SONY TCD-D10 PRO) and was octave band filtered in real time using the Knowles Electronics Processor for Acoustic Research (KEPAR) system. In KEPAR, bandpass filters of one octave were designed using Butterworth filters of 18th order and the center frequencies 500, 1500, and 3000 Hz. Two filters ran simultaneously in KEPAR, presenting the reference octave band (1500 Hz) through one output channel and the test octave band (500 or 3000 Hz) through another output channel of the system. The two octave band filtered signals were then presented through separate remotely controlled attenuators (NAL design), the outputs of which were taken into a selector box (NAL design). From the selector box the signal was either played back through an amplifier and a loudspeaker in the free field, or directly to a Knowles CI 2955 hearing aid receiver mounted in a BTE hearing aid case coupled to the ear via a Libby horn and a foam eartip.

The babble-noise was a mixture of four male and four female talkers recorded on digital audiotape (DAT). Two versions of the babble-noise were used in the experiment. One version (A) was filtered to match the International long-term average speech spectrum (ILTASS) *re* Byrne *et al.*

(1994), when presented at the subject's position in the free field. The other version (B) was filtered to match ILTASS at the eardrum (using the free field to eardrum correction factor by Bentler and Pavlovic, 1989) when presented in a Zwislocki coupler. All free-field stimuli were derived from version A and all occluded ear stimuli were derived from version B.

### C. Procedure

The experiment was conducted in a large anechoic chamber ( $11.1 \times 9.1 \times 7.1 \text{ m}^3$ ) where the subject was situated by a small table about 2 m in front of the loudspeaker. It was confirmed that the table had no measurable effect on the SPLs at the eardrum. Using the 1500-Hz octave-filtered babble-noise as reference, each subject was asked to balance the loudness of the 500-Hz octave-filtered babble-noise and the 3000-Hz octave-filtered babble-noise to the loudness of the reference stimulus. The subject switched between the reference and the test stimulus using a switch on the selector box, and adjusted the level of the test stimulus using an unmarked knob, which was connected to one of the variable attenuators. Before each loudness balancing the attenuator for the test stimulus was randomly set. The subjects were encouraged to use a bracketing method where they made the test stimulus both louder and softer than the reference stimulus before selecting the level that made the two stimuli equally loud. The procedure was carried out in the free field and when listening with the hearing aid receiver. Ten measurements were obtained for each of the four conditions (two transducers  $\times$  two test frequencies), which were presented in random order to the subjects. The first measurement of each condition was considered as a practice run and was discarded in further analyses.

Before each of the four loudness balancing tests, a probe tube measurement of the one-third octave spectrum of the wideband, ILTASS filtered babble-noise was recorded at the subject's eardrum in the open or in the occluded ear. During free-field testing, the probe tube was removed from the ear, and the subject completed the free-field listening with both ears unblocked. To perform the probe tube measurement in the occluded ear a probe tube was inserted in the foam eartip parallel to the tubing. During the balancing the probe tube remained attached to the probe tube microphone to avoid any low-frequency leakage from the coupling. From the individual one-third octave spectra of the wideband babble-noise measured in the open and in the occluded ear, the octave levels at 500, 1500, and 3000 Hz were determined by adding the powers measured in the appropriate one-third octave bands.

### D. Calibration issues

Prior to each subject's arrival, the setup was checked by observing that stimulus version A was 70 dB SPL at the subject's position in the free field. In addition, stimulus version B was observed to be 73.4 dB SPL in a Zwislocki coupler. This is the level that the free-field stimulus would be expected to have when measured at the eardrum of the average adult.

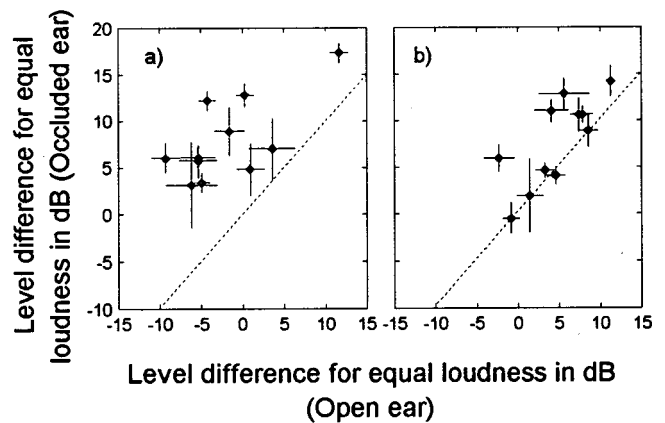


FIG. 3. Individual differences in equal loudness levels measured in the occluded ear versus the open ear for (a) a 500-Hz octave band with reference to a 1500-Hz octave band, and (b) a 3000-Hz octave band with reference to a 1500-Hz octave band. The broken lines show the relationship where the difference in equal loudness levels was the same for the two listening conditions. The bars show  $\pm$  two standard errors from nine measurements.

Probe tube measurements were conducted using a Ras-tronics probe tube microphone. The probe microphone was recalibrated for every individual probe tube used. Prior to testing, the probe tube was positioned to within 6 mm of the subject's eardrum. This was achieved by presenting white noise in the free field while observing the  $\frac{1}{4}$ -wave notch in the frequency response characteristic from the ear canal on a spectrum analyzer as the probe tube was inserted into the open ear. When the notch was positioned above 10 kHz the red marker on the probe tube was positioned to lie on a prominent feature of the tragus.

## II. RESULTS

The mean equal loudness levels and the intersubject standard deviation are given in Table I for each of the three frequencies. Although all 11 subjects were presented with the same physical input level for the 1500-Hz reference stimulus, there is some between-subject variation in the probe tube measurements at this frequency due to physical ear differences. The average equal loudness levels measured for 3000 Hz are very similar in the two test situations. However, a difference of 10 dB is observed between the average equal loudness levels measured for 500 Hz. That is, on average the subjects selected 10 dB higher levels when listening with the ear occluded than when listening with the ear open. The result is illustrated in Fig. 3. Figure 3 shows for each individual the difference in equal loudness level between each test frequency and the reference frequency measured in the occluded ear as a function of the same measurement in the open ear. The bars show plus and minus two times the intrasubject standard error of nine repetitions. It is clear that when equalizing the loudness at 3000 Hz relative to 1500 Hz similar levels were needed for the two transducers. However, when equalizing the loudness of 500 Hz relative to 1500 Hz, the subjects selected 5–15 dB higher levels when listening with the occluded ear than when listening with the open ear.

TABLE I. Average equal loudness levels measured in dB SPL with a probe tube less than 6 mm from the eardrum ( $N=11$ ). The figures in the brackets show the intersubject standard deviation.

Frequency (Hz)	Open ear (dB)	Occluded ear (dB)
500	60.6 (5.1)	70.7 (5.3)
1500	62.5 (3.1)	62.8 (1.8)
3000	67.1 (3.6)	70.4 (4.6)

The average difference in the equal loudness levels measured for 500 and 3000 Hz are  $-6.6$  dB ( $\pm 3.8$  dB) and  $0.3$  dB ( $\pm 4.1$  dB) for the open ear and the occluded ear, respectively (cf. Table I). These mean values correspond well with the average values of  $-8.5$  and  $0.7$  dB displayed in Fig. 2. A  $t$ -test for dependent samples shows that the difference between the open ear and the occluded ear measurements is highly significant ( $p=0.0008$ ). Further, the average difference in the equally loud levels of the low- and high-frequency octave bands presented in the open and in the occluded ear is  $-6.9$  dB ( $-6.6$  dB $-0.3$  dB), which is slightly lower than the  $-9.2$  dB ( $-8.5$  dB $-0.7$  dB) seen in Fig. 2.

### III. DISCUSSION

Data from the present study and from a large range of other published studies suggest that when balancing or rating loudness, normal-hearing listeners tend to select an average of 10 dB higher level for low-frequency sounds at 500 Hz when listening with the ear occluded than when listening with the ear open, cf. Figs. 2 and 3. Largely, the data presented in Fig. 2 were for each type of transducer measured on different populations, in different environments, and using different methods. The data reported in Fig. 3 were, however, collected on the same group of subjects in a test environment where only the transducer was a variable, giving a result similar to that seen in Fig. 2. Looking back at the missing 6 dB case in which researchers found a low-frequency discrepancy in equal loudness levels measured in the open and in the cushioned ear, it is likely that the difference reported in this paper is also due to some physical and/or psychological factors. In the following, a range of potential factors, which we believe we have been able to exclude, is discussed with reference to the instrumentation and procedure used in this study. The factors are, however, still open for discussion, and it would be interesting to see more work presented in this area.

#### A. Implementation

In the case of the occluded ear, the balancing test was performed with a probe tube inserted in the foam eartip parallel to the tubing that carried the sound from the hearing aid receiver to the eardrum. By performing measurements in a Zwislocki coupler with the probe tube clear and then filled with putty, it was confirmed that the measurements were not affected by interactions between the two tubes. Further, because all loudness levels were individually measured at the subjects' eardrum, factors such as coupler measurements, transfer functions, and low-frequency leakage can be disregarded.

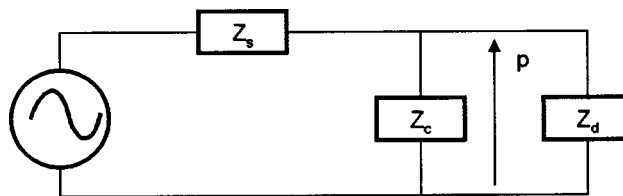


FIG. 4. Lumped element model of the ear, showing the source impedance,  $Z_s$ , the impedance of the unoccluded portion of the ear canal,  $Z_c$ , and the impedance of the eardrum and middle and inner ear,  $Z_d$ . The pressure measured in the ear canal is  $p$ .

#### B. Excitation of the ear

One possibility is that occlusion of the canal changes the relationship between SPL near the eardrum and the power entering the middle ear system. We have repeatedly been told this, but we do not believe that it is the case, for the following reasons.

Let us first model the ear canal as a transmission line, terminated at one end by the impedance of the eardrum (incorporating the reflected middle and inner ear impedances) and at the other by the impedance of the source (incorporating the tubing and miniature receiver). The sound pressure measured at a given point in the canal is the vector addition of the wave progressing inwards towards the eardrum plus the wave reflected outwards from the eardrum. The magnitude and phase of the reflected wave relative to the inwards-traveling wave depend only on the complex impedance of the eardrum relative to the characteristic impedance of the ear canal. That is, the source impedance has no effect on this relationship, and therefore also has no effect on the power entering the eardrum relative to the power in the inwards-traveling wave in the ear canal. The relationship between the pressure measured in the ear canal and the power entering the eardrum is therefore also unaffected by the source impedance, although the magnitude of the pressure at all points in the ear canal is directly affected by the properties of the source.

Let us now model the ear canal and eardrum as lumped impedance elements, as shown in Fig. 4. The pressure,  $p$ , measured in the canal will depend on every one of the three impedances shown. The velocity of the eardrum, and the power entering the eardrum, however, depend only on  $p$  and  $Z_d$ , the impedance of the eardrum, and is therefore unaffected by the properties of the source if  $p$  is known. Consequently, the relationship between SPL and the power entering the eardrum is unaffected by the source, and will therefore be the same for the occluded and open cases.

For both models we reach the conclusion that the relationship between SPL at a given point in the ear canal and the power entering the middle ear system is independent of the source. It seems extremely unlikely that a different conclusion would be reached for a more complex model of the ear comprising combinations of lumped elements and a non-uniform transmission line. The physical difference between the open and occluded ear cannot explain the variation in loudness perception when SPL in the ear canal is controlled.

### C. Bone conduction transmission

Another factor that has been considered is the effect of bone conduction transmission. In the open ear, sound entering via the bone conduction route is estimated to be 40 to 60 dB lower at the cochlea than that entering via air conduction. This factor is thus unlikely to affect loudness perception in the open ear. We are unaware of any evidence that driving the ear canal with an occluding source enhances bone conduction transmission sufficiently for it to play a significant role in perception for people with normal hearing. Furthermore, such an enhancement would have to be frequency dependent to produce the results presented in this article.

### D. Distortion products

Transducer distortion could affect loudness perception. Long-term spectra of the octave band filtered test stimuli were measured in the free field and in a Zwislocki coupler when the overall level of the wideband stimulus corresponded to 70 dB SPL in the free field (FF). The spectra show no trace of distortion products in the background noise in the frequency region from 12.5 to 20 kHz. That is, the distortion products were at least 40 dB down for the 500-Hz octave band in both the free field and in the Zwislocki coupler. For the 1500- and the 3000-Hz octave band the distortion products were down by at least 25 dB in the free field and by 35 dB in the Zwislocki coupler. It is therefore unlikely that transducer distortion can explain the difference reported in this experiment.

### E. Mechanical vibration

In the case of the missing 6 dB, Rudmose (1982) demonstrated that most subjects were affected by vibrational energy transmitted from the loudspeaker to the listener via the floor. By providing a mechanical isolation of the chair from the floor, Rudmose observed up to 2 dB difference in the selected equal loudness levels at 100 Hz. In our experiment, the system made up of the loudspeaker, stand, wire-mesh floor, chair, and subject primarily consists of masses and springs and may be regarded as a low-pass filter with a low cutoff frequency. Without having attempted any calculations on such a system, we believe that mechanical vibrations are unlikely to affect the transmission of sounds at 500 Hz and higher frequencies.

### F. Test procedure

Another problem discussed in the “missing 6 dB” case is the test procedure. The balancing tests were then conducted by either switching from free field listening to listening under headphones (with one ear occluded to ensure monaural listening), or an earphone was on one ear while the other ear was uncovered and used to listen in the free field. In both cases the main problem was to ensure proper attenuation of the low frequencies in the occluded ear. This was in particular a problem in the latter situation as the ear fitted with the earphone would still contribute to the listening in the free field. Rudmose (1982) demonstrated that an improperly occluded nontest ear could change the equal loudness

level by about 2 dB. Another problem arises from remembering the levels when switching between transducers in the first case, and also in balancing the sound from the two transducers which often produced different “quality” sounds. All of these problems were eliminated in the present study by using the same transducer to provide the reference and the test stimulus. Further, in this study a frequency independent of the two test frequencies in question was used as a reference stimulus. This procedure was introduced to eliminate a possible effect from always listening to one frequency at the same level while adjusting the level of the other frequency. Finally, in this study binaural listening was used with one transducer (loudspeaker) and monaural listening was used with the other (hearing aid receiver). If binaural loudness summation is frequency dependent, this difference in the test procedure could affect the result. However, we did repeat the free field testing with six of the original ten subjects using monaural listening (the nontest ear was occluded with a foam earplug). With reference to the binaural listening condition, the six subjects produced a slightly higher equal loudness level for 500 Hz (62.1 dB SPL at eardrum), whereas the equal loudness level for 3000 Hz was about the same (66.6 dB SPL at eardrum). An analysis of variance using the individual equal loudness free field levels as observations and frequency and procedure as repeated measures showed no significant difference between monaural and binaural listening ( $p = 0.72$ ).

### G. Loudspeaker location

The factor that seemed to have most effect on the equal loudness level in Rudmose’s experiments (about 4 dB) was the location of the loudspeaker, which seemed to have a psychological effect on some subjects. That is, some subjects wanted more sound pressure from a loudspeaker placed near to the ear than from a loudspeaker positioned further away. The rationale is that the near source is regarded as having a “smaller acoustic size” and consequently must produce more sound pressure level in the ear canal to equal the loudness of the distant and “larger” loudspeaker. In the experiment reported in this paper the loudspeaker was positioned about 2 m away from the subject. If moving the loudspeaker closer to the subject had produced higher equal loudness levels at 500 Hz, this would have decreased the discrepancy seen between open and occluded ears in this study. But that is provided the loudspeaker position only affects the perception of low-frequency sounds. Rudmose (1982) also noted that training the subjects could eliminate the effect. We did not investigate this factor in a systematic way, but in the papers referred to in Fig. 2 the loudspeaker positions from the test subjects were 0.45, 1, and 2 m, producing average equal loudness level differences between 500 and 3000 Hz of  $-7.2$ ,  $-9.2$ , and  $-8.0$  dB. Further, after the reported experiment was completed, three of the authors and a sound engineer who are all experienced listeners and who all knew the outcome of this study and who were skeptical about a psychological explanation, performed the balancing test. Their average results agreed with those presented in this paper, and

TABLE II. Average level differences for equal loudness (500–3000 Hz) as measured in a range of studies (see text). The figures in the brackets show the standard deviation across studies.

Condition of ear	Threshold level (dB)		Discomfort level (dB)	
Open	-4.2;	$N=1$	-5.2(2.0);	$N=3$
Cushioned	-3.1(1.9);	$N=6$	-4.8(2.8);	$N=3$
Occluded	-1.8(4.0);	$N=6$	-0.6(3.6);	$N=6$

thus we have no other data but Rudmose's to support the theory about the loudspeaker location and training of subjects.

## H. Acoustic reflex

An interesting possibility is that the acoustic reflex may be more easily triggered when the foam eartip is inside the ear canal. The acoustic reflex, when triggered, typically introduces a transmission loss of low-frequency sounds (<1.2 kHz) through the middle ear of 5 to 10 dB. However, the acoustic reflex is usually associated with high level sounds above 85 dB SPL, and the SPLs used in this experiment were well below that. Also, we are not aware of any evidence that the acoustic reflex is triggered by occluding the ear.

## I. Tensor tympani muscle

It has also been suggested to us that our findings may be caused by the activation of the tensor tympani muscle in the middle ear, as tactile stimuli rather than acoustic stimuli activate this muscle. The explanation seems unlikely, as there is no conscious tactile perception of sounds at the levels used in this experiment. If there were, we would not see people with threshold levels worse than approximately 70 dB HL. We have not attempted to explore this issue further.

## J. The effect at other input levels

This paper has solely concentrated on suprathreshold levels around the comfortable loudness category. Equal loudness data are, however, also available for threshold levels and loudness discomfort levels (Arlinger and Kinnefors, 1989; Beattie *et al.*, 1997; Cox *et al.*, 1997; Denley-McFatrige, 1990; Hawkins, 1980; Humes *et al.*, 1996; ISO R226, 1961; Keidser *et al.*, 1999; Morgan and Dirks, 1974; Morgan *et al.*, 1974; Walker *et al.*, 1984; Wilber *et al.*, 1988). The average differences in the levels needed for equal loudness perception of low- and high-frequency sounds presented in the open, the cushioned, and the occluded ear across these studies are shown in Table II. Though the differences in results for open and occluded ears at threshold and discomfort levels are not as large as seen for comfortable loudness levels, the results show a similar tendency, that is that the magnitude of the difference in equal loudness level between 500 and 3000 Hz is smaller when listening with the ear occluded than when listening with the ear open.

## IV. CONCLUSION

We have presented data on normal-hearing people from a large range of studies that suggest that the relative loudness

perception of low- and high-frequency sounds presented at a comfortable level differ in the open and the occluded ear. A specially designed experiment was reported that confirms the findings. We have been unable to find any physical explanation for the discrepancy. A large range of factors that potentially could have influenced the results was discussed, including the factors that were considered to influence the 6-dB difference in loudness levels measured in open and cushioned ears some decades ago. We wish to alert other researchers to the issue, as the effect could distort results from psychoacoustic experiments or cause errors in applications related to how we perceive sounds, such as when fitting hearing aids based on loudness perception.

## ACKNOWLEDGMENTS

This study was supported by the Cooperative Research Center for Cochlear Implant, Speech and Hearing Research. We also wish to thank Knowles Electronics for making their KEPAR system available for this project, and to thank two reviewers for valuable comments to this paper.

- Arlinger, S., and Kinnefors, C. (1989). "Reference equivalent threshold sound pressure levels for insert earphones," *Scand. Audiol.* **18**, 195–198.
- Beattie, R. C., Huynh, R. C., Ngo, V. N., and Jones, R. L. (1997). "IHAFF loudness contour test: reliability and effects of approach mode in normal-hearing subjects," *J. Am. Acad. Audiol.* **8**, 243–256.
- Bentler, R. A., and Pavlovic, C. V. (1989). "Transfer functions and correction factors used in hearing aid evaluation and research," *Ear Hear.* **10**(1), 58–63.
- Beranek, L. L. (1949). *Acoustic Measurements* (Wiley, New York), pp. 731–755.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Kotby, M. N., Nasser, N. H. A., El Kholy, W. A. H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. I., Westerman, S., and Ludvigsen, C. (1994). "An international comparison of long-term average speech spectra," *J. Acoust. Soc. Am.* **96**, 2108–2120.
- Christen, R., and Byrne, D. (1980). "Preferred listening levels for bands of speech in relation to hearing aid selection," *Scand. Audiol.* **9**, 3–10.
- Cox, R. M. (1989). "Comfortable loudness level: stimulus effects, long-term reliability, and predictability," *J. Speech Hear. Res.* **32**, 816–825.
- Cox, R. M. (1995). "Using loudness data for hearing aid selection: The IHAFF Approach," *Hear. J.* **48**(2), 10, 39–44.
- Cox, R. M., Alexander, G. C., Taylor, I. M., and Gray, G. (1997). "The contour test of loudness perception," *Ear Hear.* **18**(5), 388–400.
- Denley-McFatrige, L. K. (1990). "Loudness summation across bandwidth: effect of hearing loss and psychophysical procedure," Masters thesis, Memphis State University.
- Hawkins, D. B. (1980). "The effect of signal type on the loudness discomfort level," *Ear Hear.* **1**(1), 38–41.
- Holte, L., and Margolis, R. H. (1987). "The relative loudness of third-octave bands of speech," *J. Acoust. Soc. Am.* **81**, 186–190.
- Humes, L. E., Pavlovic, C., Bray, V., and Barr, M. (1996). "Real-ear measurement of hearing threshold and loudness," *Trends Amp.* **1**(4), 121–135.
- ISO R 226. (1961). Normal equal-loudness contours for pure tones and normal threshold of hearing under free field listening conditions (International Organization for Standardization, Switzerland).
- Keidser, G., Seymour, J., Dillon, H., Grant, F., and Byrne, D. (1999). "An efficient, adaptive method of measuring loudness growth functions," *Scand. Audiol.* **28**(1), 3–14.
- Keidser, G., Katsch, R., Dillon, H., Byrne, D., Seymour, J., and Banerjee, S. (1996). *Evaluation of fitting rationales for non-linear hearing aids*, NAL Research & Development Annual Report 1995-1996, pp. 15–16.
- Khalil, E. (1989). "Monaural and binaural loudness summation of speech," unpublished independent study, Central Institute for the Deaf, St Louis.
- Killion, M. C. (1978). "Revised estimate of minimum audible pressure: Where is the 'missing 6 dB'?" *J. Acoust. Soc. Am.* **63**, 1501–1508.

- Morgan, D. E., and Dirks, D. D. (1974). "Loudness discomfort level under earphone and in the free field: The effects of calibration methods," *J. Acoust. Soc. Am.* **56**, 172–178.
- Morgan, D. E., Wilson, R. H., and Dirks, D. D. (1974). "Loudness discomfort level: Selected methods and stimuli," *J. Acoust. Soc. Am.* **56**, 577–581.
- Munson, W. A., and Wiener, F. M. (1952). "In search of the missing 6 dB," *J. Acoust. Soc. Am.* **24**, 498–501.
- Pollack, I. (1952). "Comfortable listening levels for pure tones in quiet and noise," *J. Acoust. Soc. Am.* **24**, 158–162.
- Richards, A. M. (1975). "Most comfortable loudness for pure tones and speech in the presence of masking noise," *J. Speech Hear. Res.* **18**, 498–505.
- Ricketts, T. A., and Bentler, R. A. (1996). "The effect of test signal type and bandwidth on the categorical scaling of loudness," *J. Acoust. Soc. Am.* **99**, 2281–2287.
- Rudmose, W. (1982). "The case of the missing 6 dB," *J. Acoust. Soc. Am.* **71**, 650–659.
- Walker, G., Dillon, H., Byrne, D., and Christen, C. (1984). "The use of loudness discomfort levels for selecting the maximum output of hearing aids," *Aust. J. Audiol.* **6**(1), 23–32.
- Watson, N. A., and Knudsen, V. O. (1940). "Selective amplification in hearing aids," *J. Acoust. Soc. Am.* **11**, 406–419.
- Wilber, L. A., Kruger, B., and Killion, M. C. (1988). "Reference thresholds for the ER-3A insert earphone," *J. Acoust. Soc. Am.* **83**, 669–676.



# An effect of temporal asymmetry on loudness

G. Christopher Stecker and Ervin R. Hafter

*Department of Psychology, University of California at Berkeley, 3210 Tolman Hall #1650, Berkeley, California 94720-1650*

(Received 22 May 1999; revised 1 December 1999; accepted 6 March 2000)

A set of experiments was conducted to examine the loudness of sounds with temporally asymmetric amplitude envelopes. Envelopes were generated with fast-attack/slow-decay characteristics to produce F–S (or “fast–slow”) stimuli, while temporally reversed versions of these same envelopes produced corresponding S–F (“slow–fast”) stimuli. For sinusoidal (330–6000 Hz) and broadband noise carriers, S–F stimuli were louder than F–S stimuli of equal energy. The magnitude of this effect was sensitive to stimulus order, with the largest differences between F–S and S–F loudness occurring after exposure to a preceding F–S stimulus. These results are not compatible with automatic gain control, power-spectrum models of loudness, or predictions obtained using the auditory image model [Patterson *et al.*, *J. Acoust. Soc. Am.* **98**, 1890–1894 (1995)]. Rather, they are comparable to phenomena of perceptual constancy, and may be related to the parsing of auditory input into direct and reverberant sound. © 2000 Acoustical Society of America. [S0001-4966(00)02606-0]

PACS numbers: 43.66.Cb, 43.66.Mk [RVS]

## INTRODUCTION

Sounds in the world can often be characterized as the result of two kinds of processes: production (the act of driving an acoustic medium to vibrate) and filtering (the effects of the medium upon the sound). Each of these processes imbues specific characteristics to the sound. Driving forces give a sound its attack; striking a cymbal produces a rapid attack, while bowing the same cymbal produces a very gradual attack. Filtering, on the other hand, primarily shapes the decay; a very stiff cymbal will decay more slowly than one composed of softer metal. The stiffness and mass of the cymbal affect its temporal envelope as well as its spectral characteristics. Of course, there are many cases where filtering influences a sound’s attack, and where driving forces control the decay, but for a wide variety of dynamic sounds (e.g., percussion), attack time is controlled primarily by the driving force, while decay time is controlled by filtering.

In fact, when we consider sounds from the perspective of a listener, we realize that the “filter” affecting any particular sound is actually a combination of many interacting filters, including the source object (e.g., a violin string) and coupled structures (the violin body) and air masses (the violin’s interior, the room containing both violinist and listener). Embedded in the sound’s decay is information about the characteristics of each of these filter components. The size of a room, for instance, can be estimated from its reverberation time, the materials which cover its walls, by the spectral changes in the decay, etc. Similarly, the geometric shape of struck objects can be discriminated based on the vibrational modes apparent in their acoustic decay (Lakatos *et al.*, 1997). The ability to “parse” a sound in such a way as to recover information about these different characteristics would certainly be advantageous for any organism. Not only would this ability assist in identifying sound sources by their physical characteristics, but it would also help orient the listener in reverberant space.

The term “temporal asymmetry” (Patterson, 1994a) is

used to describe sounds with differing attacks and decays. Such sounds can be useful in any study attempting to describe auditory processes involved in perceiving dynamic stimuli, and especially those involved in processing attacks and decays. For instance, if a tone with an instantaneous attack and a gradual decay is reversed in time, the result is a new tone which shares many characteristics with the first (e.g., Fourier spectrum, peak amplitude, etc.). However, the attack and decay times of the two tones are radically different. In terms of driving forces and filtering, the first tone resembles an impulse delivered to a linear filter; interpreting the second in a similar manner implies a very different filter (broadband, with a short time constant), as well as a very different driving force (gradually increasing). Differential processing of attacks and decays by the auditory system should result in major perceptual differences between the two stimuli.

It has been shown that temporal asymmetry can affect a stimulus’ timbre (Cutting and Rosner, 1974; Rosen and Howell, 1981; Patterson, 1994a), its perceptual timing (Morton and Marcus, 1976; Vos and Rasch, 1981; Gordon, 1987), and even its pitch (Hartmann, 1978). Best known is the study of Cutting and Rosner (1974), who showed that changing the attack time of tonal stimuli resulted in a change in timbral categorization, with subjects describing fast-rising tones as “plucked” and slow-rising tones as “bowed.”

Patterson (1994a, 1994b) introduced temporal asymmetry into the envelope of an amplitude-modulated tone by reversing the order of attack and decay of each cycle of the modulator. In his so-called “damped” tones, each cycle had an instantaneous attack followed by an exponential decay. Simply reversing the modulator in time produced “ramped” tones, with exponential rises and instantaneous decays. Despite the fact that the two stimuli had identical energy spectra, Patterson’s subjects reported that ramped tones had a stronger sinusoidal quality than damped tones, which sounded “hollow” and “percussive.” Clearly, the exponen-

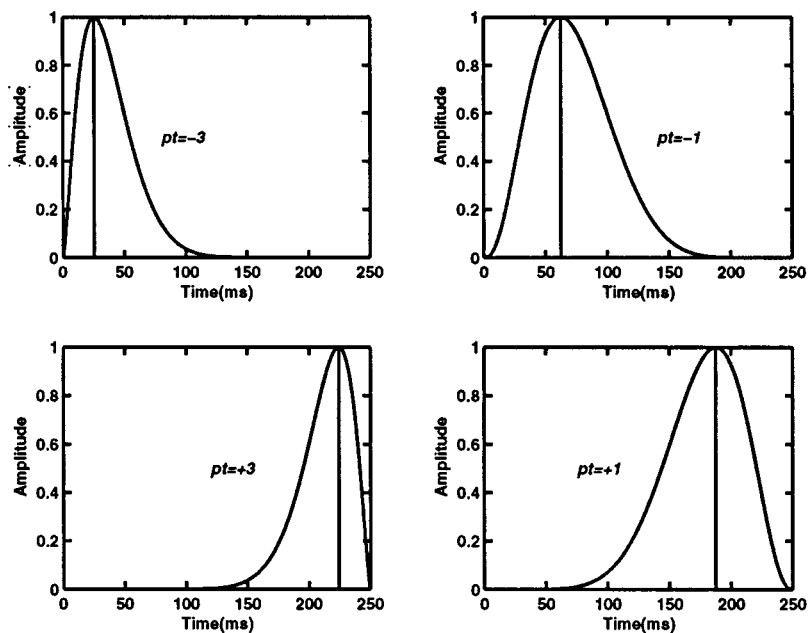


FIG. 1. Gating envelopes used in generating experimental stimuli. The top two plots (with negative values of  $pt$ ) show envelopes for F-S stimuli. On the bottom, with positive values of  $pt$ , are S-F envelopes. Note that the left- and right-side envelope pairs differ in both total energy and rise/fall slope, but that F-S and S-F envelopes within each pair are equal-energy, temporally reversed versions of one another. The four envelopes shown here were used to gate pure tones (in experiments 1 and 3) and broadband noise (in experiment 2).

tial portion of the stimulus was perceived differently based upon its role as attack or decay.

Vos and Rasch (1981) studied the perceptual onset time (the moment at which a stimulus is judged to begin) of temporally asymmetric stimuli. Two such stimuli were generated by gating 20-component harmonic complexes with trapezoidal envelopes having identical decays, but different rise times. Pairs of these stimuli were repeated every 800 ms, and subjects manually adjusted the time between first and second tones to produce perceptual isochrony; that is, rhythmic regularity. From this study, Vos and Rasch (1981) concluded that the perceptual onsets of these stimuli occurred at the moment when their levels exceeded a relative threshold of about 6 to 15 dB below their peak levels. Thus, perceptual onset was strongly related to attack time.

Based on a suggestion by Wessel (1995), Stecker (1995) used a similar method to examine the perceptual onset of smoothly varying 330-Hz tones of 250-ms duration. Attacks and decays were sigmoidal, as opposed to the linear form used in Vos and Rasch (1981). Alternate tones were time-reversed, such that a stimulus with a fast attack and slow decay (hereafter referred to as fast-slow, or F-S) was followed by one with a slow attack and fast decay (slow-fast, or S-F). Each F-S, S-F pair was repeated every 1000 ms. When subjects adjusted the division of this period to produce isochrony, the time from F-S to S-F was shorter than that from S-F to F-S, again suggesting that perceptual onsets were earlier for stimuli with rapid attacks than for those with slow attacks. In terms of the relative threshold model suggested by Vos and Rasch (1981), the results indicate a threshold approximately 15 to 20 dB below peak amplitude.

The S-F and F-S stimuli in Stecker (1995) were equated in loudness using a perceptual adjustment task prior to the temporal judgments. Unexpectedly, S-F stimuli were judged to be louder than F-S stimuli of equal amplitude. The current study was designed to examine this result more explicitly. In experiment 1, the relative loudnesses of F-S and S-F tones were judged in a two-interval comparison (2IC)

procedure. Subsequently, in experiment 2, a magnitude estimation (ME) procedure was used to obtain absolute judgments of loudness for tones and broadband noises. Finally, the role of carrier frequency was examined, again using the ME procedure, in experiment 3.

## I. EXPERIMENT 1

In this experiment, subjects made paired comparisons between the loudness of fixed-level standard tones and variable-level test tones. Both standard and test tones were gated with temporally asymmetric envelopes.

### A. Methods

#### 1. Subjects

Three college-aged, paid subjects with normal auditory thresholds from 250 to 4000 Hz participated in the experiment. Although experienced at listening in psychoacoustics (signal detection and localization) experiments, all subjects were unaware of the purpose of the present study at the time of their participation.

#### 2. Stimuli

The stimulus set consisted of four pure-tone (330-Hz) stimuli, each gated by a different temporally asymmetric envelope, 250 ms in duration. Four envelopes were generated using the equations in the Appendix, which define a function with the following properties: (1) It is smooth at all points in time, with a continuous change in slope. (2) Attack and decay portions are divided by a momentary amplitude peak, with no intervening steady-state period. (3) The degree of temporal asymmetry is described by the peak time of the envelope, which is set using the parameter  $pt$ . As shown in Fig. 1, four values of  $pt$  were used to generate two pairs of envelopes:  $-3/+3$  and  $-1/+1$ . Note that the wider envelopes generated with  $pt = \pm 1$  present greater total energy, shallower rise/fall times, and (by definition) more centralized peak times than those with  $pt = \pm 3$  (see Fig. 1). However,

the comparisons of interest will be between equal-energy F–S and S–F envelopes with the same absolute value of  $pt$ . Carriers and envelopes were multiplied with the phase of the carrier set to 0 deg at the beginning of each envelope. This multiplication was performed in software prior to running the experiment.

### 3. Procedure

A two-interval comparison (2IC) procedure was used to obtain relative loudness judgments among the four stimuli. Trials presented a standard stimulus in the first interval, followed after 300 ms by a test stimulus for comparison. For each trial, a new test stimulus was selected, at random, from among the four experimental stimuli, while the standard remained the same for an entire run of 108 trials. Standards were always presented at a level of 80 dB SPL (levels throughout this paper are of equivalent 1-kHz sinusoid at signal's peak amplitude), while the levels of test stimuli were selected at random from a set of nine possibilities: 70, 74, 76, 78, 80, 82, 84, 86, and 90 dB SPL. Subjects were told to indicate, by button press, which interval contained the louder stimulus. They were specifically instructed to base each judgment on the "total 'sound energy' contained over the duration of the sound," ignoring other aspects, such as sharpness and timbre. Stimuli were delivered via Stax model SR-5 electrostatic headphones in a double-walled sound-proof booth located in the basement of Tolman Hall on the Berkeley campus.

Blocks of 108 trials were organized such that each of the four envelopes was presented as the test tone three times at each of the nine presentation levels, in random order. The standard remained the same for all of the 108 trials within a particular block; a new standard was selected from among the four potential envelopes for each new block. Each subject completed 12 blocks with each of the four standards. The resulting data were analyzed using logistic regression. (SPSS, 1999).

### B. Results

As can be seen in Fig. 2, the range of levels used for test tones was adequate to evoke nearly a full range of potential responses. That is, subjects nearly always judged the test tone as being the louder when presented at 90 dB SPL (10 dB above the standard), and nearly always judged it to be the quieter when presented at 70 dB SPL. The shapes of the loudness functions (relating response probability to presentation level) do not differ among the four stimuli used; differences in loudness between the stimuli can be seen as lateral shifts of the loudness functions. Overall, S–F stimuli were judged louder than the standard significantly more often than F–S stimuli ( $p < 0.05$ ); this is discussed in more detail below. Additionally, the effect of overall stimulus energy ( $pt = \pm 3$  vs  $pt = \pm 1$ ) is also significant; this is to be expected due to the relationship between energy and loudness. There is no significant interaction between envelope type (F–S vs S–F) and stimulus energy. Therefore, subsequent analyses combine stimuli across magnitude of  $pt$ , instead focusing on differences in its sign.

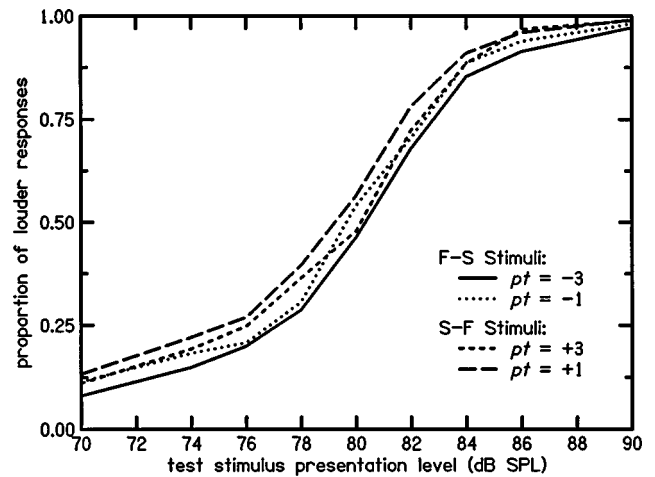


FIG. 2. Results of experiment 1 (2IC procedure), indicating proportion of responses (averaged across three listeners) indicating the test tone to be louder than the standard, as a function of test-tone level. The standard was always presented at 80 dB SPL. Overall, responses indicate greater loudness for S–F tones than for F–S tones. Test tones with  $pt = \pm 1$  were louder than their  $pt = \pm 3$  counterparts, consistent with the energy difference between them. However, the pattern of (F–S vs S–F) results is the same for the two pairs, and the shapes of the loudness functions do not differ among the four stimuli.

Figure 3 shows the same data averaged over presentation level and magnitude of  $pt$ , but plotted as a function of the F–S/S–F envelope characteristic of the standard in each trial. Overall analyses of the response proportions given for each combination of standard and test tone indicate two things: first, as noted above, S–F stimuli were judged as being louder than any particular standard more often than were the corresponding (equal-energy) F–S stimuli. Second, the difference between S–F and F–S versions of the envelopes was greatest when the standard itself was one of the F–S stimuli. Both the main effect of test-tone envelope and the interaction with the envelope characteristic of the standard were statistically significant for both pairs ( $pt = \pm 1$ ,  $pt = \pm 3$ ) of envelopes ( $p < 0.05$ ). Finally, linear regression estimates based on the psychometric functions for one listener (subject CD) indicate the magnitude of the loudness

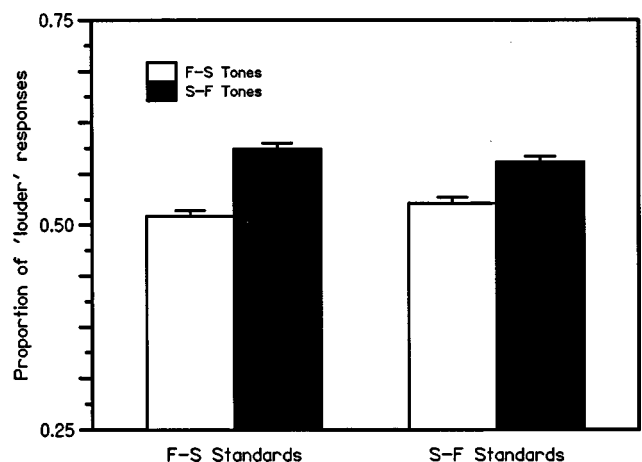


FIG. 3. Results of experiment 1 (2IC procedure), continued. Data are plotted separately for F–S ( $pt < 0$ , left bars) and S–F ( $pt > 0$ , right bars) standards. An overall loudness advantage for S–F test tones appears for both types of standard, but the effect is significantly attenuated for S–F standards.

difference to be approximately 1–3 dB following F–S standards, but less than 1 dB following S–F standards.

### C. Discussion

The main finding of experiment 1 is that S–F stimuli were judged to be louder than F–S stimuli of equal energy. A second finding is that the judgments were influenced by the stimulus “context” immediately preceding the judged tone: the difference in loudness was greatest when judged tones were compared to stimuli with F–S envelopes.

## II. EXPERIMENT 2: MAGNITUDE ESTIMATION

While the main findings of experiment 1 clearly demonstrate that S–F tones are judged to be louder than equal-energy F–S tones, the finding that this difference depends upon the choice of standard is troubling. No theory which assigns loudnesses to single stimuli in a context-free manner can account for this type of finding. Perhaps the context effect indicates a procedural influence, with listeners altering their response criteria based on the identity of the standard. In experiment 2, we sought to reduce this context effect by eliminating the standard from each trial. Instead, we used a magnitude estimation (ME) procedure, where each response indicates an *absolute* judgment for a single stimulus. While the context for relative judgments in the 2IC task consists of the previous stimulus (the standard), context in ME tasks consists of the entire set of stimuli [see, e.g., Marks (1988)].

We adopted a procedure quite similar to that of Stevens (1956), with a few critical differences. In one case, Stevens presented listeners with a standard or “modulus,” set to a fixed level and assigned a magnitude, for instance, “10.” Listeners were instructed to estimate the loudness of subsequent test stimuli by assigning numbers to each, such that the ratio of the response to the modulus equaled the perceived ratio of the two loudnesses. That is, a stimulus 4 times as loud as the modulus should be called “40,” one half as loud should be “5,” etc. In an alternative procedure, Stevens (1956) omitted the modulus but maintained the ratio instructions by requiring listeners to respond with numbers whose ratios matched the relative loudnesses of the test stimuli.

Since we hoped to eliminate the effects of relative judgment in this experiment, we omitted Stevens’ (1956) ratio instructions, instead asking subjects to simply assign numbers which matched their perceived loudnesses, using any scale with which they felt comfortable. In an effort to reduce intersubject variability, however, we included a single “reminder” stimulus at the start of each run. Subjects were told that this reminder, presented at a moderate fixed level, should be called “50,” and that its purpose was to help them anchor their scales of loudness from run to run and day to day. However, subjects were also told that they were free to choose any particular values for their ratings, so long as they tried to be consistent across trials (the same loudnesses should produce the same ratings later in the experiment). As a result, most subjects used scales centered around 50 (e.g., 0–100, 20–80), although a few used very different scales (see Fig. 4).

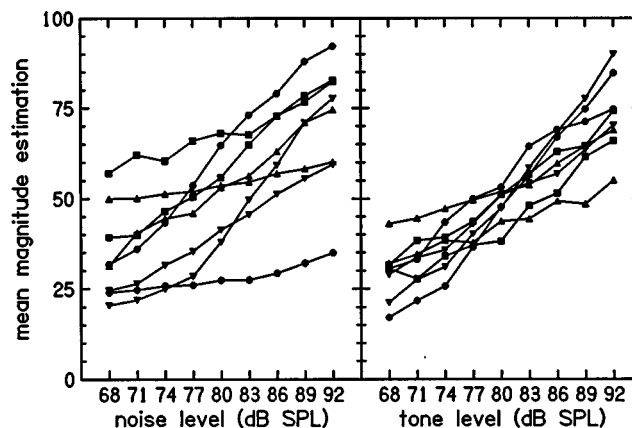


FIG. 4. Intersubject variability in rating data (ME procedure). The mean value of magnitude estimations made by each subject is plotted as a function of stimulus level. Ratings for noise stimuli appear on the left; ratings for tonal stimuli are plotted on the right. Separate lines represent individual subjects, who varied in both the range (indicated by the slopes of their loudness functions) and bias (indicated by their intercepts) of their responses.

We were also interested in whether these results would generalize to stimuli with different characteristics, in terms of tonality, frequency, bandwidth, and regularity. Irino and Patterson (1996), for instance, found that temporal asymmetry for ramped/damped modulation was reduced when using noise, compared to tonal carriers. As such, we ran two conditions in experiment 2: a pure-tone condition with the same stimuli as in experiment 1, and a noise condition using a white-noise carrier which changed from trial to trial.

### A. Methods

#### 1. Subjects

One group of eight undergraduate volunteers, participating to obtain course credit, was included in the pure-tone condition. A second group of eight undergraduates, recruited separately, participated in the noise condition. All subjects reported having normal hearing; however, audiometric testing was not performed in conjunction with the experiment. None was informed about the purpose of the study prior to participation.

#### 2. Stimuli

In the pure-tone condition, stimuli were exactly as described for experiment 1, although their levels were changed slightly (see below). However, carrier and envelope signals were synthesized separately and combined in real time using an analog multiplier (voltage-controlled amplifier), just prior to amplification and stimulus delivery via earphones. Carrier phase (at the envelope’s initial sample) was set to 0 deg.

In the noise condition, the carrier signal was a white (1 Hz–20 kHz  $\pm$  0.1 dB) noise, generated by an analog noise generator. As in the pure-tone condition, this carrier was gated in real time using an analog multiplier, thus causing each trial to contain a different noise sample. The four envelopes (see Fig. 1) used for gating were identical to those used in experiment 1.

### 3. Procedure

At the start of each experimental run, subjects were presented with a “reminder” stimulus. A single envelope (F–S,  $pt = -3$ ) and level (80 dB SPL) was used for this in all conditions. Subjects were instructed to consider the perceived loudness of this stimulus to be “50,” and then rate the magnitude of each of 72 subsequent test stimuli on a scale of their choosing. The purpose of the reminder was to designate to listeners the median presentation level at the start of each run; subjects were not instructed to use it as a reference, but to rate each trial independently. Each trial contained a single stimulus followed by a prompt on the computer screen and the subject’s response. Test stimuli were drawn randomly from the set of four envelopes and nine presentation levels, spaced every 3 dB from 68 to 92 dB SPL. Note that this is a slightly larger range of levels than that used in the 2IC procedure, and that the purpose of randomizing levels is the same as in experiment 1: namely, to provide a source of variation in stimulus loudness. Randomization was constrained to produce an equal number of presentations, within each run, for each envelope and level. Here, stimuli were delivered to subjects via Etymotic ER-4S earphones within a single-walled soundproof booth.

At the prompt, subjects indicated the perceived loudness of the test tone by entering a numerical value on a computer keyboard. As in experiment 1, subjects were instructed to base each judgment on the total sound energy contained over the duration of the sound, ignoring other aspects, such as sharpness and timbre. Each subject finished six runs, including an initial (discarded) “training” run, and runs 2–6, which were used in the subsequent analysis. The entire procedure was identical for the two conditions (pure-tone and noise).

### B. Results

Figure 4 shows the mean magnitude estimations for each subject, as a function of stimulus SPL. Notice that the ordinate is plotted in linear, not logarithmic, coordinates, and that the loudness functions are all approximately linear. Normally, magnitude estimation experiments produce ratio responses, and the obtained loudness functions are power functions. The linear functions shown in Fig. 4 indicate that our subjects did not use ratio scales in estimating loudness; indeed, they resemble functions obtained when the scale is preassigned (Stevens, 1956). Stevens argued that such linear functions resulted from overly constraining the listeners’ responses and that his magnitude estimation method, free of such constraints, revealed the natural judgments to be based on ratios, rather than linear scales. Without discounting the converging evidence in support of Stevens’ power law, we might say that, based on our own subject’s tendency to use linear scales in the absence of specific ratio instructions, the idea that magnitude estimations are inherently ratios may follow from procedural effects rather than any fundamental perceptual law.

Figure 4 also reveals large intersubject variation in both the range and bias of their responses. To facilitate the pooling of data across subjects, we normalized the slope and

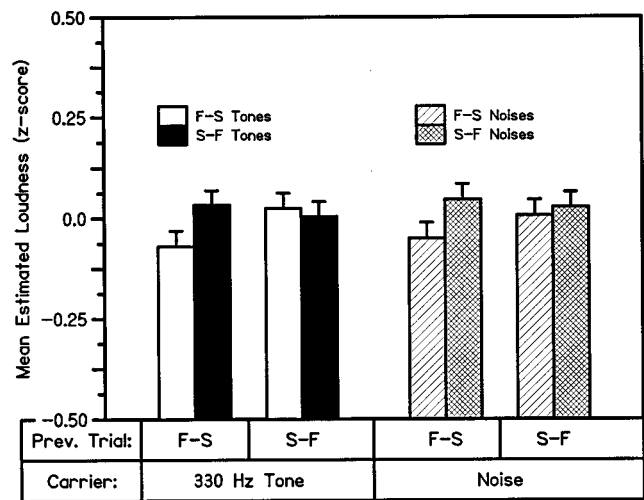


FIG. 5. Results of experiment 2 (ME procedure). The mean loudness ratings for 330-Hz tones (black and white bars) and broadband noises (hatched bars), are plotted by test-tone envelope shape, separately for trials immediately preceded by F–S trials and those preceded by S–F trials. Each listener’s ratings are normalized, using a  $z$  transform, prior to combining data across subjects (see the text). As in experiment 1, an overall loudness advantage is observed for S–F stimuli, and is significantly reduced following previous S–F stimuli. There are no statistically significant differences between the results for tones and noises.

intercept of each subject’s loudness function using the  $z$  transform: each rating ( $y$ ) was transformed into a  $z$  score

$$z = \frac{y - \bar{y}_{\text{subj}}}{\sigma_{\text{subj}}}$$

based on the individual subject’s overall mean ( $\bar{y}_{\text{subj}}$ ) and standard deviation ( $\sigma_{\text{subj}}$ ) of responses (across all presentation levels).  $Z$ -score ratings were treated as raw data points for subsequent analysis by analysis of variance (ANOVA). As stated above, this transformation (which follows from the decision to treat the obtained loudness functions as linear) normalizes both the slope and intercept of the loudness function. While it is not useful for visualizing the relationship between intensity and loudness, by reducing intersubject variability, it allows the pooling of data across subjects and facilitates the comparison of loudness functions for different stimuli. A difference in loudness between F–S and S–F stimuli, for instance, should result in a lateral displacement of the normalized loudness functions, and a resulting difference between the mean  $z$  scores, averaged across SPL, as shown in Fig. 5. Mean  $z$ -score loudness ratings are shown for 330-Hz tones and broadband noises, plotted as a function of the envelope shapes of both the test stimulus and that of the test stimulus on the previous trial (i.e., the local context). This partitioning of the data allows us to examine the influence of prior stimulation, despite the fact that judgments were not intended to incorporate this local context, as they were under the 2IC procedure.

For pure tones, the main effect observed in experiment 1 was reproduced using the ME procedure. Overall, stimuli with S–F envelopes were rated louder than F–S stimuli ( $p < 0.05$ ). In addition, there was a significant interaction ( $p < 0.05$ ) between the loudness ratings for a given stimulus and the envelope shape of the context stimulus. Examination

of the simple effects of test-stimulus envelope for each context reveals a significant loudness difference following F–S ( $p < 0.05$ ), but no significant difference following S–F context stimuli.

There was no significant difference between the results for noise and those for pure tones. As in the pure-tone condition, S–F noise stimuli received significantly higher ratings ( $p < 0.05$ ) than F–S stimuli, and the difference was significantly greater ( $p < 0.05$ ) when preceded by F–S context stimuli.

### C. Discussion

First, the main finding of experiment 1, that S–F tones were louder than F–S tones of equal energy, was reproduced in experiment 2, using the ME procedure. This was true for both tones and noises, indicating that the effect did not depend upon the tonal quality or the spectral makeup of the stimuli used.

Second, although the ME procedure was designed to eliminate the effects of local context, it did not. It is clear that subjects did, to some extent, judge all test stimuli in the context of previous stimuli. Specifically, S–F stimuli were rated louder than F–S stimuli when the previous stimulus had F–S characteristics.

Previously, two kinds of context effect have been reported in magnitude estimation and ratio-scaling experiments. Marks (1988), for instance, demonstrated that manipulating the overall range of presentation levels for each of two stimulus types can systematically alter loudness matches between the stimuli. This does not explain the short-term effect of stimulus context observed in this study, since the effect Marks described reflects differences across the entire stimulus set. Jesteadt *et al.* (1977) demonstrated an influence of response history on magnitude estimations, whereby large responses tend to be followed by large responses and small responses tend to follow small responses (an *assimilative* sequential effect). The current study, however, shows an *assimilative* effect only when both the just-prior stimulus and the test stimulus have F–S envelopes. Responses to S–F test stimuli in this case are not assimilated. Thus, the effect is not simply *assimilative* or *contrastive*, but depends upon the envelope of the stimulus being judged. Additionally, the context effect is *asymmetric*, in that F–S contexts affect subsequent judgments, but S–F contexts do not. These facts, along with the agreement of results from both 2IC and ME procedures, suggest that the observed context effect is not procedural in origin, but may instead relate to the mechanism responsible for the difference between S–F and F–S loudness. We will return to this issue in the general discussion below.

### III. EXPERIMENT 3: EFFECTS OF CARRIER FREQUENCY

Patterson (1994a, 1994b) showed that ramped tones produce a stronger sinusoidal percept than do corresponding damped tones, a finding not unlike that observed in the current study. In addition, he noted that this perceptual asymmetry was diminished for high-frequency ( $> 3$  kHz) carriers, and related that finding to the loss of phase-locking in

auditory-nerve fibers at high frequencies (Patterson, 1994b). He applied the auditory image model (AIM) (Patterson *et al.*, 1995), a modular framework for modeling the early auditory system, to explain his results. While carrier frequency was not a focus of Patterson's modeling work, there are at least three temporally asymmetric processing stages in AIM, each related to frequency in a different manner. These stages are auditory filtering, neural encoding, and temporal integration (Irino and Patterson, 1996).

AIM's representation of basilar-membrane activity includes narrow-band filtering, rectification, and compression (Patterson *et al.*, 1995). Filters slightly mistuned from the carrier frequency of a stimulus respond differently to ramped and damped tones. The output waveform in response to a ramped tone is extended in time, while that to a damped tone is more sharply peaked, with an amplified peak and attenuated decay.<sup>1</sup> Compression, which follows auditory filtering, enhances this difference by "squashing" the peaked response to damped tones. This results in greater overall activity in response to ramped tones. This mechanism is frequency dependent. In general, increasing carrier frequency (without changing the stimulus envelope) will reduce the difference between responses to ramped and damped stimuli. This difference is naturally related to the similarity between the decay of the stimulus and that of the filter's impulse response.

Later stages explicitly incorporate mechanisms of temporal adaptation which are asymmetrical in time, responding rapidly to increases in input level, but lagging behind rapid decreases. The end result of this adaptation is that the exponential attack of a ramped sound has more influence on activity in the model than does the corresponding exponential decay of a damped stimulus.<sup>2</sup> Irino and Patterson (1996) suggest that these mechanisms, rather than the behavior of auditory filters, play the dominant role in producing temporal asymmetry in AIM, for two reasons: first, they exhibit a much larger effect of temporal asymmetry (Irino and Patterson, 1996), and second, they provide a better explanation for temporal asymmetry in nonperiodic sounds, such as damped and ramped noises (Akeroyd and Patterson, 1995). While a comprehensive treatment of these mechanisms is beyond the scope of this paper, it is important to point out that these adaptive mechanisms do not appear to be frequency dependent.

We became interested in AIM as a potential mechanism for producing temporal asymmetry in the loudness of F–S and S–F tones, and posed the following questions: Can AIM predict differences in the *loudness* of F–S and S–F tones? If so, will the predicted difference be affected by carrier frequency? Finally, will the loudness of high-frequency F–S and S–F tones be judged by human listeners in agreement with predictions from AIM?

As it stands, AIM does not generate direct predictions of loudness; however, the software package itself<sup>3</sup> includes a loudness function which calculates a loudness measure based on instantaneous activity in the temporally integrated auditory image. This measure is essentially the mean of activation across frequency channels, and represents the instantaneous loudness of a stimulus as it changes over time. Figure

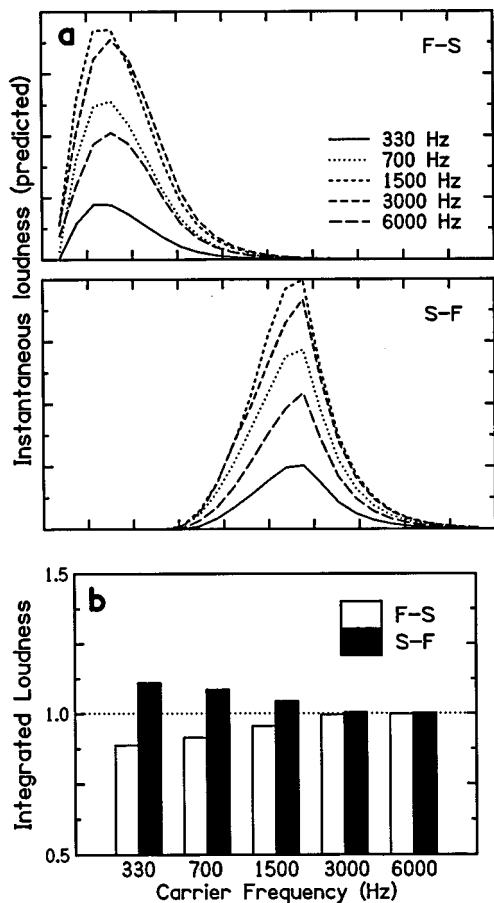


FIG. 6. Instantaneous loudness as a function of time (a) for F–S (upper panel) and S–F (lower panel) sounds. Values are predictions from AIM (Patterson *et al.*, 1995) and the parameter is carrier frequency. There are two points to note: first, middle frequencies (1500 and 3000 Hz) are louder overall than either high or low frequencies. Second, while F–S stimuli have functions which resemble their envelopes (see Fig. 1), S–F stimuli produced continued activity following their offsets. In (b), predicted loudness has been integrated over time, producing a single loudness estimate for each stimulus. Since the overall magnitudes of these estimates vary across frequency, each value was normalized (divided) by the mean of the two estimates for stimuli at its carrier frequency. The normalized mean is indicated by the dotted line. As can be seen in the plot, AIM predicts a larger difference between F–S and S–F stimuli at low frequencies (330–1500 Hz) than at higher frequencies.

6 displays these predictions as functions of time for each stimulus. By integrating these instantaneous estimates across time, we generated overall loudness values for pure-tone signals with envelopes as in experiment 1 and carrier frequencies from 330 to 6000 Hz. These showed that, as we found, AIM predicts greater loudness for S–F than F–S stimuli at low frequencies (330–1500 Hz), but a marked decline in this temporal asymmetry above 3 kHz (Stecker and Hafter, 1998). The purpose of experiment 3 was to test this result further by examining the loudnesses of F–S and S–F tones over a range of carrier frequencies (330–6000 Hz), again using the ME procedure.

## A. Methods

### 1. Subjects

Subjects were a new group of ten undergraduate volunteers, participating to obtain course credit. All reported hav-

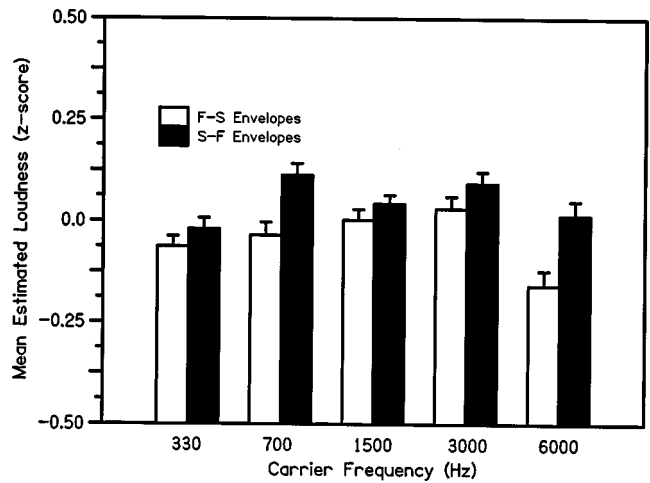


FIG. 7. Results of experiment 3: loudness ratings for F–S and S–F tones across carrier frequency. While AIM (Patterson *et al.*, 1995) predicts an elimination of temporal asymmetry above approximately 3000 Hz, the observed loudness advantage is consistently in favor of S–F tones across carrier frequencies from 330–6000 Hz.

ing normal hearing, although audiometric testing was not done. None was informed of the study’s purpose prior to participating, and none had participated previously in any of the earlier experiments in this study.

### 2. Stimuli

Sinusoidal carriers, 250 ms in duration, were synthesized at 330, 600, 1500, 3000, and 6000-Hz carrier frequencies, using a starting phase of 0 deg. Envelopes were generated using the same procedure described in the Appendix, and were of the same form used in the prior experiments. As in experiment 2, carriers were gated in real time using an analog multiplier (voltage-controlled amplifier) prior to amplification for earphone presentation.

### 3. Procedure

Within each frequency condition, the stimulus presentation and subject response procedures were identical to those used in the ME procedures in experiments 1 and 2. Each subject completed between six and eight experimental runs at each of two randomly assigned carrier frequencies (resulting in four subjects per frequency condition). Since not all subjects completed all eight runs, only runs 2–6 were used for subsequent analysis. A single carrier frequency was used for all stimuli in a single run, including the initial standard-level presentation and all of the test-stimulus presentations. Frequency conditions were not intermixed; all runs at one carrier frequency were completed before moving on to another. Loudness data were analyzed according to envelope type (S–F/F–S) and carrier frequency.

## B. Results

The results of experiment 3 can be seen in Fig. 7. As in the previous experiments, S–F tones were rated as louder than F–S tones ( $p < 0.05$ ). This difference was consistent across all tested carrier frequencies, with no significant interaction between carrier frequency and envelope type.

## C. Discussion

While AIM correctly predicts that S–F tones should be louder than F–S tones, it incorrectly predicts that this difference should be reduced at high frequencies. Based on the AIM literature (Patterson *et al.*, 1995; Patterson, 1994a; Irino and Patterson, 1996; Carlyon, 1996), we assume that the predicted frequency difference arises primarily due to processes involved in auditory filtering. However, Irino and Patterson (1996) suggest that the greatest amount of temporal asymmetry arises later, in the temporal integration stage, which is not frequency dependent; this raises an interesting question: why, if most of AIM's temporal asymmetry is attributable to mechanisms which are *not* frequency dependent, does AIM predict such a large reduction in temporal asymmetry at high frequency? Regardless of the answer to that question, it seems reasonable to assume that AIM could be modified to eliminate the effects of carrier frequency, but we have not undertaken such a modification in our work.

Regardless of the mechanisms producing frequency dependence in AIM, our results show that the difference in loudness between F–S and S–F tones exists across frequencies from 330–6000 Hz. Among other things, this implies that phase-locking in auditory-nerve fibers is not necessary to produce the effect.

## IV. GENERAL DISCUSSION

### A. Main effect of greater loudness for S–F stimuli

The primary finding reported here is the greater loudness of S–F over F–S stimuli, with slow-rise/fast-decay stimuli perceived as being louder than fast-rise/slow-decay stimuli of equal energy. In effect, loudness mechanisms treat the two ends of the stimulus differently based on their roles as either attack or decay. This effect was observed in every experiment, under both two-interval (2IC) and magnitude-estimation (ME) procedures.

We believe that one may rule out several possible mechanisms which might have caused this finding. One is automatic gain control (AGC), which predicts exactly the opposite, specifically that F–S stimuli should have been louder than S–F stimuli. Since AGC mechanisms must utilize estimates of a stimulus' envelope to adjust its level, they naturally respond with some delay. When presented with rapid increases in stimulus level, extra energy is allowed into the system before the AGC can respond. Thus, the output of an AGC will be greater in magnitude for stimuli with fast attacks than for those with slower attacks.

Spectral models which calculate loudness by integrating activity across frequency bands prior to (or in the absence of) compression (Zwicker and Scharf, 1965; Moore and Glasberg, 1996) predict little or no difference between the loudnesses of F–S and S–F stimuli, since no difference exists between their energy spectra. Similarly, temporally symmetric mechanisms such as the autocorrelogram (Slaney and Lyon, 1990) cannot predict a difference between F–S and S–F stimuli [see Patterson and Irino (1998) for a discussion of temporal asymmetry in autocorrelogram models]. Thus, these models are also incompatible with our findings.

Finally, modeling results obtained with AIM (Patterson *et al.*, 1995), show that, while it correctly predicts the main finding of experiments 1 and 2 (though not the effect of context), the model is insufficient for explaining the results of experiment 3. Surprisingly, AIM predicts a reduction or elimination of temporal asymmetry above 3 kHz. With human listeners, the loudness difference between F–S and S–F tones remains for carrier frequencies up to 6 kHz.

We should point out that, although this study has focused on judgments of *loudness*, results from other researchers suggest a more general phenomenon of decay suppression (for lack of a better term) in temporally asymmetric stimuli. Patterson's (1994a, 1994b) initial ramped/damped studies, for instance, demonstrate a *timbral* difference between ramped and damped sinusoids: damped tones produce a weaker "sinusoidal" component, as if the carrier contained in the damped decay were perceptually attenuated. More recently, Schlauch *et al.* (1998) observed that the perceived *duration* of damped tones is shorter than that of corresponding ramped tones, again as if the damped tails were somehow reduced in perceptual strength. Both of these results are congruent with those of the current study; in all three cases, exponential decays seem to be perceptually suppressed, making smaller contributions to the judgments.

Two other phenomena which may be related to the current findings are deserving of mention here. First is the "de-recruitment" phenomenon reported by Canevet and colleagues (Canevet and Scharf, 1990), whereby tones and noises which continuously decrease in level are perceived to decrease in loudness more rapidly than would be expected from the change in level. This would seem to be in agreement with our finding that F–S signals (which are decreasing in level for most of their duration) are less loud, overall, than they should be. However, de-recruitment appears to require stimuli with durations greater than 1 s (Canevet *et al.*, 1999), much longer than any stimuli used in the current study.

Using a similar stimulus, Neuhoff (1998b) recently showed that loudness *change* is judged greater for simple and complex tones increasing in level, relative to tones decreasing by the same amount. While this phenomenon could also be related to our current findings, several key points make comparing the studies directly difficult: most importantly, Neuhoff argues that the perception of loudness change is fundamentally different from the perception of loudness itself (Neuhoff, 1998a) and specifically instructs his listeners to ignore the overall loudness of the stimuli. Second, as in the de-recruitment phenomenon (Canevet and Scharf, 1990), stimulus durations in Neuhoff's experiment were much longer (1.8 s) than in ours. The amount of level change was also limited to 15 dB, resulting in a much slower rate of change. Third, while Neuhoff found the results quite similar for sinusoids and complex tones, he found no such asymmetry between white noises increasing and decreasing in level, whereas the differences in loudness found here were similar for the two carrier types.

### B. Effect of stimulus context

In addition to finding greater loudness for S–F tones and noises, the current data also show significant effects of prior



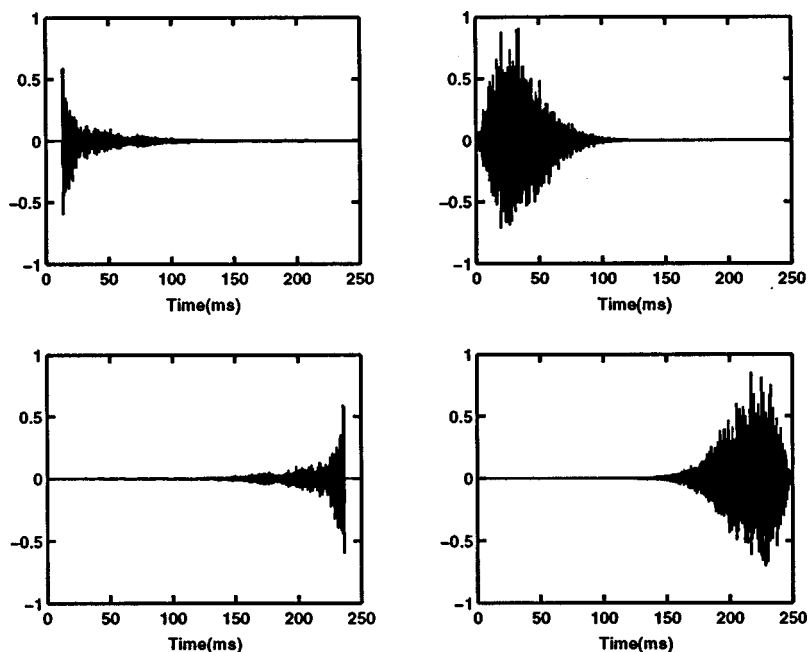


FIG. 8. Decay of reverberant sound. On the left, waveforms of a brick being struck by a hammer in a conference room (Houtsma *et al.*, 1987). On the right, examples of noise stimuli used in experiment 2. The F–S stimulus (top right) has a decay similar to the hammer/brick sound (top left). Playing the hammer/brick recording backwards (bottom left) produces a stimulus similar to the S–F stimuli (bottom right) used in this experiment.

stimulus context. Specifically, preceding the test stimulus with an S–F context reduces or eliminates the loudness difference (i.e., the difference is greatest when stimuli are preceded by F–S contexts). It is important to note that, to our knowledge, no existing models of loudness can account for this finding, as they normally assign loudnesses based on the characteristics of single sounds, independent of previous stimuli. In other words, the context effect we have demonstrated suggests a mechanism with a time constant spanning hundreds of milliseconds or even seconds. Aside from Neuhoff’s explanation of perceptual bias for rising tones (Neuhoff, 1998b, 1998a), none of the models discussed above possesses time constants this long, and so they cannot make predictions regarding the effect. Perhaps models which can accommodate our main finding (AIM, for low frequency carriers) could be extended through the addition of higher-order mechanisms that operate on this longer time scale.

We originally assumed that this effect of stimulus context was related to the direct use of standard-test comparisons in the 2IC procedure; however, a nearly identical effect was obtained using ME in experiment 2. This suggests that the influence of local context on loudness judgments is not methodological in nature. Instead, we argue that this influence reflects the perceptual processing of these stimuli, independent of the judgment task. Specifically, we suggest that our results [along with those of Patterson (1994a, 1994b, 1995) and Schlauch (1998)] point to a general process of decay suppression, and that the effects of local context reflect the natural dynamics of this process.

What would a general mechanism for perceptually suppressing decays be good for? One possibility is that decays are normally more informative about the characteristics of the listening environment (primarily due to reflections, reverberation, and room resonances) than about the sound sources themselves. As shown in Fig. 8, even very short stimuli can produce rather lengthy decays due to room effects. Perhaps the auditory system “parses” the stimulus into direct and

reverberant sound, providing the listener with a cleaner version of the direct stimulus, i.e., the source.

Although the stimuli used here were not generated purposely to resemble reverberated sounds, F–S stimuli are similar to those observed in real rooms, with an initial transient followed by a slower decay. Conversely, an S–F stimulus is more like a room recording played in reverse (Houtsma *et al.*, 1987), which does not provide a sense of reverberation (see Fig. 8). The reduced loudness of F–S relative to S–F stimuli may then be the result of the auditory system compensating for the apparent effects of reverberation in F–S stimuli; eliminating the tail from the judgment of loudness because it is interpreted as an acoustic by-product, rather than a meaningful part of the stimulus being judged. In other words, listeners perceptually “interpret” an F–S stimulus as the output of a system of reverberant filters introduced by the environment. They then make judgments based upon their estimate of the input to that system. The manner in which listeners acquire knowledge of the reverberant system is an important issue which we will discuss below.

Such compensation for environmental effects has been termed “perceptual constancy” and results in the perception of distal “object” properties, rather than proximal stimulus properties (Boring, 1952; Jameson and Hurvich, 1989). For instance, a white piece of paper appears to be the same color in sunlight and fluorescent lighting, despite large changes in the proximal stimulus, including both the magnitude and spectrum of reflected light (Gilchrist, 1977). While perceptual constancies can be very strong and apparently automatic, they are not absolute or instantaneous; Rock (1997) writes of a two-stage process in which a stimulus is initially perceived in a “literal mode” corresponding to its proximal features, and then in a subsequent “constancy mode,” where the visual system has “reinterpreted” the stimulus to match its distal form.

According to Rock (1997), a good example of this process occurs in amodal completion, depicted in Fig. 9(a). The

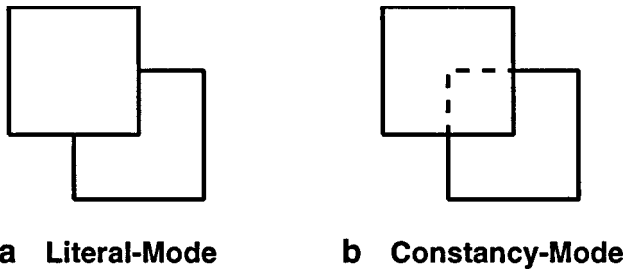


FIG. 9. Amodal completion. The ambiguous figure in (a) can be seen as either a square lying next to an L-shaped object (a literal mode interpretation), or (b) two overlapping squares (a constancy-mode interpretation).

literal-mode solution to this figure sees two objects (one square and one L-shaped) abutting, but not overlapping, while the constancy-mode solution “fills in” the occluded region, resulting in the perception of two overlapping squares. While the literal-mode interpretation is correct, it is superseded by the constancy-mode solution, which represents a more likely state of affairs in the world. Rock does point out that although the literal-mode solution is “superseded” by the constancy mode, both perceptions may in fact be available to other cognitive mechanisms simultaneously. In addition, the formation of a constancy-mode interpretation takes time. Sekuler and Palmer (1992) showed that, for visual displays similar to Fig. 9(a), it requires approximately 200 ms. As argued by Rock (1997), perception remains in the literal mode during this time.

We may distinguish the effects of these two perceptual modes on stimuli in the current study. For instance, in the literal mode the F–S stimulus is perceived as a whole, and so appears equally loud to the equal-energy S–F stimulus. But in constancy mode, the F–S stimulus is heard as composed of two parts, source and reverberation, with loudness based only on the source component. Accordingly, the loudness of the F–S will be reduced in constancy mode. In contrast, the short decay of an S–F stimulus indicates little or no effect of reverberation, and parsing the stimulus in constancy mode produces a “source” which is nearly identical to the stimulus itself. Thus, the loudness of S–F stimuli will be the same in both perceptual modes.

The formation of the constancy-mode interpretation is influenced by previous stimulus exposure. In essence, one perceptual interpretation of the stimulus (matching a particular reverberant environment) is built up, or “primed,” by experience. This facilitates the constancy-mode perception of subsequent stimuli which agree with that interpretation. Initial exposure to an F–S stimulus, for instance, primes the formation of a reverberant solution; this ultimately reduces the loudness of subsequent stimuli which match the primed interpretation. A mismatch between this interpretation and subsequent stimuli (as would occur, for instance, when moving from one room to another, differently arranged, room) causes perception to switch back to literal mode, and a new constancy-mode interpretation must be formed.

We will make two points regarding this explanation before closing. First, it must be noted that, rather than specifying a model, we have provided only a description of a potential mechanism. A “general mechanism for decay

suppression” could take many forms, and it is possible that existing or future psychoacoustic models of temporal asymmetry, such as AIM (Patterson *et al.*, 1995), may provide a framework which accomplishes what we have outlined. The effect of prior stimulation, which we have described as adaptation to changing reverberant environments, could be modeled by incorporating dynamic parameters controlled by adaptive, memory-based, or explicitly cognitive functions.

## V. CONCLUSIONS

In this paper, we have described an effect of temporal asymmetry on the loudness of short-duration (<250 ms) stimuli. In short, stimuli with fast attacks and slow decays (F–S) are heard as less loud than comparable stimuli with slow attacks and fast decays (S–F). This effect is observed across a range of experimental procedures (two-interval, magnitude-estimation) and spectral composition (high- and low-frequency pure tones, wideband noise). Additionally, the degree of effect is modulated by the envelope characteristics of prior stimuli (the local context). Neither of these effects is predicted by spectral mechanisms or AGC. The auditory image model (AIM) (Patterson *et al.*, 1995) correctly predicts the overall loudness advantage for S–F stimuli, but incorrectly predicts that the effect should be eliminated at high frequencies, and makes no prediction regarding the influence of previous stimuli.

We have argued, based upon the results of this and other studies (Patterson, 1994a, 1994b; Akeroyd and Patterson, 1995; Schlauch *et al.*, 1998), for a general mechanism of “decay suppression,” which may act to reduce the perceptual effects of room responses, and that this mechanism acquires a representation (a constancy-mode “solution”) of the reverberant context. It uses this representation for tuning its responses to subsequent stimuli. However, regardless of the actual mechanism responsible, the loudness difference between F–S and S–F stimuli demonstrates that the perception of loudness is sensitive to the shape of a stimulus’ temporal envelope, as well as to the local context of successive stimulus presentations.

## ACKNOWLEDGMENT

This research was supported by the National Institutes of Health (NIDCD Grant No. 00087), which the authors gratefully acknowledge.

## APPENDIX: STIMULUS ENVELOPE GENERATION

Envelopes for these experiments were generated using the following function of time:

$$y(t) = t^{z-1}(1-t)^{w-1},$$

where

$$z = \frac{e^s}{\left(\frac{1/2 + \arctan(pt)}{\pi}\right)^{-1}},$$

and

$$w = \frac{e^s}{z}$$

The parameter  $pt$  is used to control the envelope's peak time. When  $pt = 0$ , the resulting function is temporally symmetric, with equal rise and fall times. Negative values of  $pt$  result in temporally asymmetric functions, with shorter rise times and longer fall times. Exactly the opposite case holds for positive values of  $pt$ . In the limit, where  $pt = \pm\infty$ , the attack or decay becomes instantaneous. Two functions with equal absolute values but opposite sign of  $pt$  are temporally reversed versions of one another.

The parameter  $s$  controls the envelope "sharpness," or overall width. Varying  $s$ , by itself, produces changes in the steepness of both attacks and decays. For all envelopes used in these experiments,  $s$  was set to a value of 3.

<sup>1</sup>This is caused by cancellation in the output of the filter: the abrupt onset of the tone causes the filter to "ring" at its best frequency, while simultaneously passing the decay portion of the stimulus. Since the frequencies of the carrier and impulse response differ, they drift out of phase, and their addition results in a shortened (and attenuated) decay at the filter's output. For ramped tones, the initial ringing response is small relative to the stimulus coming in, and remains so as both grow in magnitude over the attack duration. Following the abrupt offset, the filter continues to ring, causing the output to become extended in time (Patterson, 1994a).

<sup>2</sup>Adaptation in AIM's temporal integration stage affects the frequency of "strobe" generation, which is triggered by local maxima in the input waveform. The strobe threshold adapts over time, responding rapidly to increases, but slowly to decreases in level. With each strobe occurrence, activity in the neural encoding stage is copied into a leaky buffer, providing AIM with a mechanism for temporal integration. Since strobing takes place frequently during the slow-rising portion of a ramped tone, and rarely during the corresponding portion of a damped tone, the ramped tone will produce more activity within the integration buffer.

<sup>3</sup>AIM Release 8.2, 1997, obtained at <ftp://ftp.mrc-apu.cam.ac.uk/pub/aim>

Akeroyd, M. A., and Patterson, R. D. (1995). "Discrimination of wideband noises modulated by temporally asymmetric function," *J. Acoust. Soc. Am.* **98**, 2466–2474.

Boring, E. G. (1952). "Visual perception as invariance," *Psychol. Rev.* **59**, 141–148.

Canevet, G., and Scharf, B. (1990). "The loudness of sounds that increase and decrease continuously in level," *J. Acoust. Soc. Am.* **88**, 2136–2142.

Canevet, G., Scharf, B., Schlauch, R., Teghtsoonian, M., and Teghtsoonian, R. (1999). "Perception of changes in loudness," *Nature (London)* **398**, 673.

Carlyon, R. P. (1996). "Spread of excitation produced by maskers with damped and ramped envelopes," *J. Acoust. Soc. Am.* **99**, 3647–3655.

Cutting, J. E., and Rosner, B. S. (1974). "Categories and boundaries in speech and music," *Percept. Psychophys.* **16**, 564–570.

Gilchrist, A. L. (1977). "Perceived lightness depends on perceived spatial arrangement," *Science* **195**(4274), 185–187.

Gordon, J. W. (1987). "The perceptual attack time of musical tones," *J. Acoust. Soc. Am.* **82**, 88–105.

Hartmann, W. M. (1978). "The effect of amplitude envelope on the pitch of sine wave tones," *J. Acoust. Soc. Am.* **63**, 1105–1113.

Houtsma, A. J. M., Rossing, T. D., and Wagenaars, W. M. (1987). "Effects of echoes," in *Auditory Demonstrations* [Institute for Perception Research (IPO), Eindhoven, The Netherlands].

Irino, T., and Patterson, R. D. (1996). "Temporal asymmetry in the auditory system," *J. Acoust. Soc. Am.* **99**, 2316–2331.

Jameson, D., and Hurvich, L. M. (1989). "Essay concerning color constancy," *Annu. Rev. Psychol.* **40**, 1–22.

Jesteadt, W., Luce, R. D., and Green, D. M. (1977). "Sequential effects in judgments of loudness," *J. Exp. Psychol.* **3**, 92–104.

Lakatos, S., McAdams, S., and Causse, R. (1997). "The representation of auditory source characteristics: Simple geometric form," *Percept. Psychophys.* **59**(8), 1180–1190.

Marks, L. E. (1988). "Magnitude estimation and sensory matching," *Percept. Psychophys.* **43**, 511–525.

Moore, B. C. J., and Glasberg, B. R. (1996). "A revision of Zwicker's loudness model," *Acust. Acta Acust.* **82**(2), 335–345.

Morton, J., and Marcus, S. (1976). "Perceptual centers ( $p$ -centers)," *Psychol. Rev.* **83**(5), 405–408.

Neuhoff, J. G. (1998a). "Perception of changes in loudness-reply," *Nature (London)* **398**, 673–674.

Neuhoff, J. G. (1998b). "Perceptual bias for rising tones," *Nature (London)* **395**, 123–124.

Patterson, R. D. (1994a). "The sound of a sinusoid: Spectral models," *J. Acoust. Soc. Am.* **96**, 1409–1418.

Patterson, R. D. (1994b). "The sound of a sinusoid: Time-interval models," *J. Acoust. Soc. Am.* **96**, 1419–1428.

Patterson, R. D., and Irino, T. (1998). "Modeling temporal asymmetry in the auditory system," *J. Acoust. Soc. Am.* **104**, 2967–2979.

Patterson, R. D., Allerhand, M. H., and Giguere, C. (1995). "Time-domain modelling of peripheral auditory processing: A modular architecture and software platform," *J. Acoust. Soc. Am.* **98**, 1890–1894.

Rock, I. (1997). *Indirect Perception* (MIT Press, Cambridge, MA).

Rosen, S. M., and Howell, P. (1981). "Plucks and bows are not categorically perceived," *Percept. Psychophys.* **30**, 156–168.

Schlauch, R., Ries, D., DiGiovanni, J., Elliot, S., and Campbell, S. (1998). "Intensity discrimination of ramped and damped tones," in *16th International Congress on Acoustics* (Acoustical Society of America, Seattle).

Sekuler, A., and Palmer, S. (1992). "Perception of partly occluded objects: A microgenetic analysis," *J. Exp. Psychol.* **121**(1), 95–111.

Slaney, M., and Lyon, R. (1990). "A perceptual pitch detector," in *ICASSP90, 1990 International Conference on Acoustics, Speech, and Signal Processing* (IEEE, New York), pp. 357–360.

SPSS (1999). *SPSS: Statistical package for the social sciences*, Version 9.0, Chicago, IL.

Stecker, G. C. (1995). "The effect of context on perceptual center," First-year research paper, Dept. of Psychology, U.C. Berkeley.

Stecker, G. C., and Hafter, E. R. (1998). "Temporal asymmetry and loudness: Frequency and duration effects," in *Proceedings of the 21st ARO Mid-winter Meeting* (Association for Research in Otolaryngology, St. Petersburg, FL).

Stevens, S. S. (1956). "The direct estimation of sensory magnitudes—loudness," *Am. J. Psychol.* **69**(1), 1–25.

Vos, J., and Rasch, R. (1981). "The perceptual onset of musical tones," *Percept. Psychophys.* **29**(4), 323–335.

Wessel, D. (1995). Personal communication.

Zwicker, E., and Scharf, B. (1965). "A model of loudness summation," *Psychol. Rev.* **72**(1), 3–26.

# Perceptual classification based on the component structure of song in European starlings

Timothy Q. Gentner<sup>a)</sup> and Stewart H. Hulse

*Johns Hopkins University, Department of Psychology, Baltimore, Maryland 21218*

(Received 15 March 1999; revised 17 September 1999; accepted 14 March 2000)

The ability to recognize individuals based on their vocalizations is common among many species of songbirds. Examining the psychological and neural basis of this functionally relevant behavior can provide insight into the perceptual processing of acoustically complex, real-world, communication signals. In one species of songbird, European starlings (*Sturnus vulgaris*), males sing long and acoustically complex songs composed of small stereotyped note clusters called motifs. Previous studies demonstrate that starlings are capable of individual vocal recognition, and suggest that vocal recognition results from the association of specific motifs with specific individuals. The present study tests this possibility by examining how variation among the motifs that comprise a song affect its discrimination and classification. Starlings were trained, using operant techniques, to associate multiple songs from a single male starling with one response, and songs from four other male starlings with another response. The level of stimulus control exerted by motif variation was then measured by having subjects classify three sets of novel song bouts in which motifs from the training songs were systematically recombined. The results demonstrate a significant, and approximately linear, relationship between song classification and the relative proportions of familiar motifs from different singers that compose a bout. The results also indicate that the motif proportion effects on song classification are primary to retroactive interference in the recall for specific motifs, and independent of any biases due to the syntactic organization of motifs within a bout. Together, the results of this study suggest that starlings organize the complex vocalizations of conspecifics by memorizing large numbers of unique song components (i.e., motifs) that are then associated with different classes. Because individual starlings tend to possess unique motif repertoires, it is likely that under natural conditions such classes will correspond to individual identity. Thus, it is likely that perceptual processing mechanisms similar to those described by the results of the present study form the basis for individual vocal recognition in starlings. © 2000 Acoustical Society of America. [S0001-4966(00)04806-2]

PACS numbers: 43.66.Gf, 43.80.Lb [DWG]

## INTRODUCTION

Most songbirds studied to date are able to recognize conspecific (i.e., same species) individuals based on their songs. Song-based vocal recognition is well documented among male territorial songbirds (see Stoddard, 1996 for review), and more recent examples of this behavior have been observed in both female songbirds (Wiley *et al.*, 1991; Lind *et al.*, 1997; O'Loughlen and Beecher, 1997) and in nonterritorial songbirds (Gentner and Hulse, 1998). The function of vocal recognition varies depending on behavioral context, helping to modulate the complex social dynamics among territorial males, and perhaps influencing mate choice and/or fidelity in females. Among some nonterritorial species vocal recognition is thought to function in the modulation of dominance hierarchies (Eens, 1997). From a proximate standpoint, the widespread capacity for vocal recognition observed among songbirds implies a number of constitutive cognitive processes. First, the animal must be able to discriminate one song from another. Second, the animal must form associations between certain songs and external referents, such as the individuals singing certain songs or loca-

tions in space from which certain individuals sing. For species that sing multiple songs, the problem is made more difficult as either multiple songs must be associated with the same referent, or some categorically unique feature must be linked to the set of songs sung by a given individual. In the present study, we examine the first two of these underlying cognitive processes, by investigating how variation in the component structure of male European starling, *Sturnus vulgaris*, song affects its discrimination and classification.

Male starlings tend to present their songs in long episodes of continuous singing referred to as bouts. Song bouts, in turn, are composed of much smaller acoustic units referred to as motifs (Adret-Hausberger and Jenkins, 1988; Eens *et al.*, 1991), which in turn are composed of still smaller units called notes. Notes can be broadly classified by the presence of continuous energy in their spectrogram representations, and although several notes may occur in a given motif, their pattern is usually highly stereotyped between successive renditions of the same motif. One can thus consider starling song as a sequence of motifs, where each motif is an acoustically complex event. The number of unique motifs that a male starling can sing (i.e., his repertoire size) can be quite large, and consequently different song bouts from the same male are not necessarily composed of the same set

<sup>a)</sup>Electronic mail: tim@droz.uchicago.edu

of motifs. This broad acoustical variation in their song provides several potential cues that starlings might use when learning to recognize the songs of an individual conspecific, and while maintaining that recognition over time. One straightforward recognition mechanism is the association of specific motifs with specific singers. Although some sharing of motifs does occur among captive males (Hausberger and Cousillas, 1995; Hausberger, 1997), the motif repertoires of different males living in the wild are generally unique (Adret-Hausberger and Jenkins, 1988; Eens *et al.*, 1989, 1991; Chaiken *et al.*, 1994; Gentner and Hulse, 1998). Thus, learning which males sing which motifs can provide a discriminative cue for song classification.

Data from operant studies of individual vocal recognition in starlings support the idea that recognition is based at the level of the motif. Starlings trained to recognize individual conspecifics by one set of song bouts can readily generalize correct recognition to novel song bouts from the same singers (Gentner and Hulse, 1998; Gentner *et al.*, 2000). However, when these novel song bouts have no motifs in common with the training songs, and when song exposure outside of the operant apparatus is restricted, recognition falls to chance (Gentner *et al.*, 2000). Likewise, starlings trained to discriminate among pairs of motifs will reverse the discrimination when transferred to the same motif sung by the opposite individual, and perform at chance when transferred to novel motifs sung by the training singers (Gentner, 1999). This failure to generalize correct recognition to songs composed of novel motifs, or to single novel motifs, is inconsistent with the use of individually invariant source and/or filter properties (so called “voice characteristics”) for individual vocal recognition in starlings.

Based upon the above results, it appears that starlings learn to recognize the songs of individual conspecifics by attending to information contained at (or below) the level of the motif, and then by associating distinct sets of motifs with individual singers. If this is true, then once recognition is learned it should be possible to control it systematically by varying the proportions of motifs in a given bout that come from two “vocally familiar” males. That is, if a song bout contains more familiar motifs from male A than male B, then the bout ought to be recognized more often as having been sung by male A than by male B, and vice versa. The manipulations detailed in the present study provide a test of this hypothesis.

## I. GENERAL METHODS

### A. Subjects

Eight male European starlings served as subjects for this study. Subjects were wild-caught on a farm 30 miles North of Baltimore, MD. Prior to the start of the study subjects were housed in a mixed sex aviary, along with ~50 other starlings, in large flight cages that each contained three to six birds of the same sex. During the course of the study, the subjects were housed in individual cages in a single sex aviary. The photoperiod in the aviaries followed the natural seasonal change in sunrise and sunset times at this latitude. Throughout the course of the experiment, all birds were

maintained on a restricted diet of Purina Start and Grow (Purina) such that each subject weighed 85% of their *ad libitum* body weight prior to testing each day. The birds had access to water at all times. All of the subjects were naive to the stimuli and operant procedures used in this experiment. The same eight subjects participated in each of the three experiments, and the sequence of experimental procedures (1, 2, 3, see later in this article) was the same for each subject.

### B. Apparatus

Experimental sessions were conducted inside a sound-attenuating test chamber (IAC Model AC-3) fitted with a custom-made operant response panel (see Gentner and Hulse, 1998, for complete description). Briefly, the panel contained three horizontally aligned response buttons (keys) centered above an opening through which the subjects gained controlled access to a food hopper. Two 10 W incandescent lamps provided indirect illumination of the test chamber. Stimuli were presented through a speaker (Bose model 101) mounted above and behind the response panel, out of view of the subject. A PC-type computer controlled the stimulus presentation, monitored response contingencies, and collected the data for each session. Prior to testing, the maximum sound level of the acoustic stimuli inside the test chamber was set at  $70 \pm 2$  dB (A) SPL. The same apparatus was used in each of the three experiments.

### C. Stimuli

#### 1. Song recording

Recordings of six male European starlings were used to generate all the stimuli for this experiment. The procedures for obtaining digital song recordings from male starlings and for manipulating those songs on a computer have been detailed elsewhere (Gentner and Hulse, 1998). Briefly, a minimum of 0.5 h of song was recorded from each of the five males, while housed individually in a large sound-attenuating chamber. During recording, males had visual and auditory access to a female starling. The same female was used to induce song from all the males. All the songs were recorded on digital audiotape (16 bit, 44.1 kHz) using the same microphone (Sennheiser ME66-K6), downloaded to a computer, and high-pass filtered at 250 Hz to remove extraneous low-frequency background noise. Of the six males recorded, only five were used to generate the stimuli for the baseline training. The songs of the sixth male were used only as the “unfamiliar” songs during experiment one (see below). None of males whose songs were used to generate the stimuli served as subjects in the operant testing. The males we recorded for these stimuli were taken from an area in which the birds used as subjects were also obtained. However, the subjects were caught nearly three years after the stimulus birds. General familiarity with the training songs is possible, but unlikely, as subjects caught at this site a year earlier failed to recognize these same songs (Gentner *et al.*, 2000). Regardless of this possibility, the effects reported here are based only on the classifications learned in the base-

TABLE I. Baseline stimulus set configurations showing correct responses to the different bouts from each singer. Letters denote singer identity, numbers denote specific song bouts from that singer (e.g., A1 refers to bout number “1” from male “A”). Note that songs A1, A2, B1, and B2 occur in both set 1 and set 2

	Peck left (INDIV)	Peck right (MULT)
Set 1	A1, A2, A3, A4, A5, A6, A7, A8,	B1, B2, C1, C2, D1, D2, E1, E2
Set 2	B1, B2, B3, B4, B5, B6, B7, B8	A1, A2, C3, C4, D3, D4, E3, E4

line training, and any facilitated recognition due to stimulus preexposure would not bias the results except to increase the acquisition rate of the baseline task.

## 2. Baseline training stimuli

We used two different stimulus sets for the baseline discrimination. Half of the subjects were trained with each set. Each stimulus set consisted of eight exemplars sampled from the songs of a single bird, and eight exemplars sampled from the songs of four other birds (two per bird). Each exemplar was  $15 \pm 0.5$  s of continuous song taken from either the beginning, middle, or end of a song bout. Table I shows the detailed configurations of the two baseline stimulus sets. Many of the exemplars sampled from the beginning of a song bout included whistles, along with other “warble” motifs [i.e., “variable” motifs, rattles, and high-frequency motifs; see Adret-Hausberger and Jenkins (1988) and Eens (1991) for nomenclature of motif classes]. Those sampled at later time points in a bout comprised only warble song motifs. Earlier data indicate that recognition is easily learned with this length of a song bout sample, and is unaffected by the relative position of the sample within a bout and the broader motif classes it may or may not contain (Gentner and Hulse, 1998).

## D. Procedure

### 1. Shaping

Subjects were trained to work the operant apparatus through a combination of auto-shaping and shaping procedures that represented successively closer approximations to the baseline training task (see Gentner and Hulse, 1998, for a full description). The baseline training procedure used a one-interval choice design (Hulse, 1995; Macmillan and Creelman, 1991). In the task, a peck to the center key initiated a trial by starting the playback of a randomly selected stimulus exemplar. After the exemplar played, pecks to either the left or the right key led to positive reinforcement or punishment depending on the key with which that stimulus was associated. Half of the stimuli were associated with the right key and the other half were associated with the left key (see Table I). Correct responses (e.g., a left key peck following a stimulus associated with the left key) were reinforced with 2.5-s access to the food hopper. Incorrect responses (e.g., a left key peck following a stimulus associated with the right key) were punished with a 6-s time-out during which the house lights were extinguished and the food hopper re-

mained inaccessible. In addition to causing a time-out, an incorrect response to a playback stimulus initiated a correction trial sequence in which the same exemplar repeated on all subsequent trials until the bird responded either appropriately or not at all. The interval between successive trials was 2 s. In the event that a bird failed to respond within 5 s following the completed presentation of a given exemplar, the trial ended without a time-out, and the computer waited for a center key peck to begin the next trial. During the first 12 s of each stimulus presentation responses to the keys had no effect, thereby increasing the amount of song to which subjects were exposed. This, in turn, helped limit potential biases for the animal to solve the task using only the first portion of each exemplar. The first response following this 12-s “observation-period” was reinforced.

## 2. Baseline training

The subjects’ task was to classify the songs of an individual male starling on one key, and those of four other conspecific males on the other key. For example, one subject was reinforced for pecking the left key each time it heard a song from bird A, and for pecking the right key each time it heard a song from either bird B, C, D, or E. Another bird was reinforced for pecking the left key each time it heard a song from bird B and for pecking the right key each time it heard a song from bird A, C, D, or E. In each case, we refer to the key associated with the songs of the single bird as the *individual* (INDIV) key, and the key associated with the songs from multiple birds as the *multiple* (MULT) key. Likewise, the songs associated with each key are referred to as “INDIV” and “MULT” stimuli respectively (see Table I). The association of the INDIV and MULT stimuli with either the left or right response key was counterbalanced across subjects. Two different stimulus sets were used to train the baseline task (see Table I). Half the subjects ( $N=4$ ) were trained with one set, and half ( $N=4$ ) were trained with the other; thus, each subject was exposed to only one baseline training set.

The distribution of correct responses is dependent upon the extent to which the subjects are capable of recognizing the stimuli associated with each of the two response keys. To solve this task subjects must discriminate among the different stimulus exemplars and associate each one with the proper response. It is important to note, however, that the baseline training task does not require subjects to employ a categorical solution strategy, such as “this is (or is not) a song from bird A.” Rather, the subjects only need to learn that one “set” of songs is associated with one response, and perhaps that the complement to this set is associated with the other response. Importantly, abstract category formation with respect to the baseline stimuli is not a prerequisite to the experimental manipulations that follow the baseline training. When we refer to classification or recognition in discussing the results of these experiments, we mean explicitly that the subjects discriminated a stimulus exemplar and correctly associated it with an operant response. The present study examines these discriminative and associative behaviors.

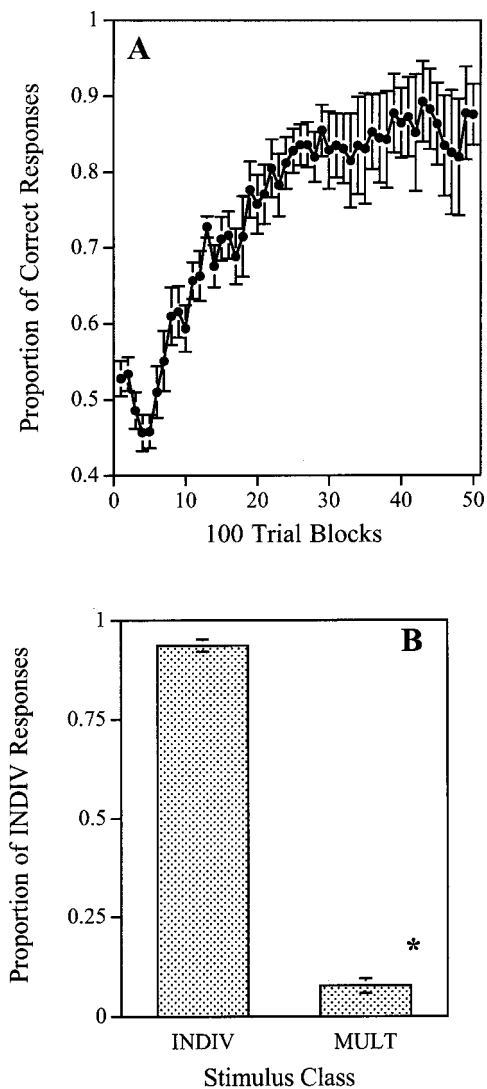


FIG. 1. (a) Baseline acquisition curve. Mean proportion of correct responses made to all the baseline stimuli over the first 50 blocks of trials. (b) Asymptotic performance on the baseline task. Mean proportion of responses made to the key associated with the INDIV stimuli following the presentation of both INDIV and MULT baseline exemplars. (\* Significant difference between the means,  $p < 0.0001$ .)

### E. Statistical analysis

For each subject, we recorded a single response (or lack thereof) and the stimulus presented on every trial. Each response was coded as either “correct” or “incorrect,” depending upon the particular key with which the exemplar presented on a given trial was associated. Alternatively, responses could be coded as either “INDIV” or “MULT” depending upon the key that was pecked. Using these two coding strategies, performance could be expressed as the probability of responding correctly to any stimulus event [as in Fig. 1(a)], or as the probability of making either an INDIV or MULT response to any stimulus event [as in Fig. 1(b)]. Each subject’s performance during the initial baseline training was examined in blocks of 100 trials. Performance during each of the separate experiments was analyzed as a single block of trials that contained all of that subject’s responses to the test and baseline stimuli. Performance on the baseline task was assessed against a 95% confidence interval around

chance, and with a one-way repeated measures analysis of variance (ANOVA) to detect any significant difference between the responses made to INDIV and MULT baseline stimuli. The significance level ( $\alpha$ ) was set at 0.05, two tailed. Data from correction trials (i.e., the nonrandom trial immediately following an incorrect response) and trials in which the subject made no response were not included in any of the statistical analyses, but did contribute to the total number of baseline acquisition trials shown in Fig. 1(a). Data from the test conditions were analyzed using multi-way repeated-measures ANOVAs,  $t$ -tests, and Fisher’s PLSD *post hoc* comparisons.

### F. Baseline training results and discussion

All of the subjects quickly learned the baseline task, discriminating between the songs of an individual male starling (INDIV stimuli) and those of four other male starlings (MULT stimuli). As a group, the subjects were responding above 70% correct after 768 trials [approximately 1200 trials including correction and no-response trials, see Fig. 1(a)]. Figure 1 shows the acquisition curve for baseline training, along with the mean levels of performance at asymptote. The subjects’ performance improved significantly over the first 50 blocks of training [ $F(6,49) = 16.042$ ,  $p < 0.0001$ ] and there was no significant difference in the rate of acquisition between subjects trained on the two stimulus sets [ $F(1,49) = 1.064$ , NS]. The mean ( $\pm$  SEM) percent correct score over the last five blocks of baseline trials was  $93.6 \pm 1.6$  for the INDIV stimuli and  $92.3 \pm 1.8$  for the MULT stimuli. Likewise, at asymptote [Fig. 1(b)], subjects were correctly classifying the stimuli into their appropriate classes (INDIV and MULT). This is reflected in the significant difference between responses (either left or right key pecks) to the stimuli in the INDIV class and those in the MULT class [ $F(1,7) = 726.857$ ,  $p < 0.0001$ , see Fig. 1(b)].

The results of the baseline training demonstrate that subjects were able to discriminate accurately among and classify correctly the different exemplars in each of the two training sets. Moreover, within each training set, we have established two discriminable classes of songs, those associated with the INDIV response and those associated with the MULT response. In the experiments that follow, we examine the perceptual mechanisms underlying this classification by testing the hypothesis that it is based on the varying proportions of familiar motifs in a song bout. Although the baseline data alone do not demonstrate that the subjects categorized the stimuli according to individual identity, the robust discrimination and classification is a sufficient basis for the manipulations that follow. The extent to which their behavior in the operant apparatus is related to the more ecologically relevant task of individual vocal recognition is taken up in the general discussion.

## II. EXPERIMENT 1: SERIAL POSITION EFFECTS

### A. Introduction

The hypothesis that song classification is based on the varying proportions of familiar motifs in a song bout is, in essence, a claim that motifs are perceptually distinct and

memorable auditory events. For recognition to occur in this manner, starlings must parse song bouts into their constituent motifs and then recall with which response each motif has been associated. Under such conditions, motifs that are more easily heard or those that are more easily remembered would exert greater control over their behavior. Therefore, prior to a direct test of our hypothesis, we consider in experiments 1 and 2 how such a motif-based classification system is affected by other cognitive processing mechanisms that, in turn, might influence the salience of specific motifs.

One way that specific motifs might elicit differential control in song classification is through their location in the sampling interval. Because the motifs in a song bout are sung in serial order over time, the recall of specific motifs, and thus decisions about how to classify a song bout at any given time point in the sampling interval, may be affected by constraints in the memory of the classification system. Among humans, the ability to recall a list of words often appears as a characteristic U-shaped function. This is the well-known serial-position effect in which items placed at the beginning and end of a list are more easily recalled than those in the middle (Baddeley, 1990)—the so-called “primacy” and “recency” effects, respectively. Recency effects are well established in the animal learning literature, but the existence of primacy effects remains more equivocal (see Deacon and Rawlins, 1995; Wright and Rivera, 1997). To the extent that starlings are subject to serial-position effects during song classification, motifs occurring at either the beginning or the end of a sampling interval may exert differential control over their decision to consider that sample (i.e., that bout) as belonging to one class or another. In the first experiment, we investigate how serial-position effects modulate the song classification behavior learned in the baseline training.

## B. Methods

### 1. Stimuli

The test stimuli for experiment 1 were constructed from four pairs of baseline stimulus exemplars. Two pairs of exemplars were chosen from each baseline training set, and the same two pairs were used to test all animals trained with that set. Each pair comprised one INDIV exemplar and one MULT exemplar (set 1: A1–D1 and A2–E1; set 2: B1–A2 and B2–D3, see Table I). Each exemplar was chosen at random with the constraint that it had been accurately classified (above 90% correct) by all subjects during the last five blocks of the baseline training. Using digital editing software (SoundDesigner II, Digidesign-Avid) each baseline exemplar was divided into three ~ 5-s segments, taking care that the point of division between any two adjacent segments fell at a natural transition between motifs. The test stimuli were made by combining the song segments from each pair of INDIV and MULT exemplars according to the six patterns shown in Fig. 2. Given the constraints that the original relative position of each song segment (either first, second, or third) is always maintained, the six patterns in Fig. 2 account for all possible three-element permutations. In addition to the 12 stimuli constructed using the two pairs of INDIV and MULT baseline exemplars (six test stimuli per pair), we constructed

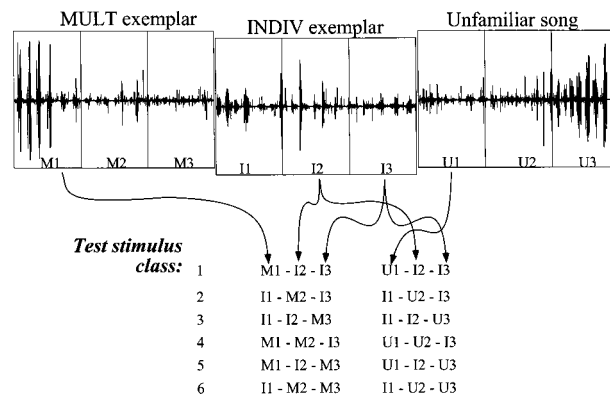


FIG. 2. Strategy for constructing the experiment 1 test stimuli from two baseline stimulus exemplars and an unfamiliar song bout. Non-overlapping 5-s song segments were taken from each of the three stimuli and recombined following the six patterns shown.

two more sets of stimuli by substituting 5-s song segments taken from the song bouts of a novel (i.e., unfamiliar) male starling (see Sec. I C 1). In all, there were 24 different stimuli for each of the two baseline stimulus sets. Each test stimulus was ~ 15 s long.

The test stimuli used in experiment 1 varied along three dimensions: (1) the location at which a song segment from one singer was inserted into an INDIV exemplar—either the beginning, middle, or end; (2) the source of those inserted song segments—either from a familiar (i.e., MULT) or unfamiliar singer; and (3) the number of song segments inserted into a given exemplar—either 1 or 2.

In addition to examining primacy and recency effects in vocal recognition, the test stimuli for the first experiment also allow us a brief look at the effect that variation in the overall proportion of INDIV and MULT motifs in a particular song bout has on the classification of that bout. For half of the test stimuli one segment from a MULT song (or unfamiliar) bout was combined with two segments from an INDIV bout. For the other half of the stimuli, one segment from an INDIV song bout was combined with two segments from a MULT (or unfamiliar) bout. To the extent that song classification is under the control of the familiarity associated with each of the motifs in a bout, one would expect the test stimuli that contain more INDIV segments than MULT (or unfamiliar) segments to elicit significantly more responses to the INDIV key than those for which the proportion of MULT (or unfamiliar) and INDIV motifs is reversed.

### 2. Procedure

Once performance on the baseline stimuli reached an asymptote—as assessed by five consecutive 100-trial blocks in which the total percent correct during each block was above 85%—the rate of reinforcement for correct responses was reduced from 100% to 70%. The observation period was then set equal to the length of the stimulus on a given trial, and performance was again allowed to reach an asymptote. During an experimental session, all 24 test stimuli derived from a given baseline stimulus set were presented randomly without replacement on 50% of the trials until the pool of test stimuli was exhausted. Every 48 trials the pool of test



stimuli was replenished. During each session, subjects continued to hear the baseline exemplars on the remaining 50% of the trials. All responses to the test stimuli were reinforced with food 70% of the time, regardless of which key the subject pecked, but only if the subject pecked a key. “No-responses” to the test stimuli were never reinforced or punished. Each animal completed one 2-h session per day until it had responded at least 50 times to each test exemplar.

Our method of partial nondifferential reinforcement for the test stimuli served two purposes. First, because reinforcement was independent of the response choice on any given trial, subjects could not learn to associate a particular response with a given test stimulus. Therefore, provided that accurate classification of the baseline stimuli is maintained during the experiment, the proportion of right and left key pecks made to the different test stimuli (i.e., the response generalization gradient) reflects the feature space used to solve the baseline task. That is, systematic variation along relevant perceptual dimensions among the test stimuli should lead to corresponding variation in the probability with which a subject responds to either key. Second, because the subjects continued to receive reinforcement for responses to the test stimuli, responding to them did not extinguish. This allowed us to examine responses to a very large number of test stimuli over an extended period.

### C. Results

The subjects maintained very accurate discrimination of the baseline stimuli over the course of the first experiment, correctly responding to the INDIV and MULT baseline stimuli ~90% of the time. The probability of pecking the INDIV key in response to an INDIV baseline exemplar was significantly different from the probability of making that same response to one of the MULT baseline exemplars [ $F(1,7) = 547.062, p < 0.0001$ ]. At the same time, not all of the test stimuli were responded to in the same manner. Figure 3 shows the proportion of responses made to the INDIV key following the presentation of the six different test stimuli along with the baseline stimuli for reference. The differences among the mean proportion of INDIV responses associated with each of the six different types of test stimuli were significant [ $F(5,35) = 6.854, p < 0.0001$ ]. There were no significant differences between subjects trained on different baseline stimuli, in either the overall level of response to the test stimuli [ $F(1,6) = 0.420, NS$ ] or the pattern of responses distributed among the different test stimuli [ $F(5,30) = 0.115, NS$ ]. Thus, for subsequent analyses, we pooled the data for all subjects.

Most of the variation among responses to the test stimuli was due to the number of INDIV song segments that were present in a given test stimulus, and whether or not an inserted song segment was familiar (i.e., from a MULT exemplar) or unfamiliar. These main effects are shown in Fig. 4. The proportion of pecks to the INDIV key was significantly higher for the test stimuli that had only one MULT or unfamiliar song segment inserted in them, than for those test stimuli that had two MULT or unfamiliar song segments inserted [ $F(1,7) = 220.115, p < 0.0001$ ]. In addition, the proportion of INDIV responses made to the test stimuli that had

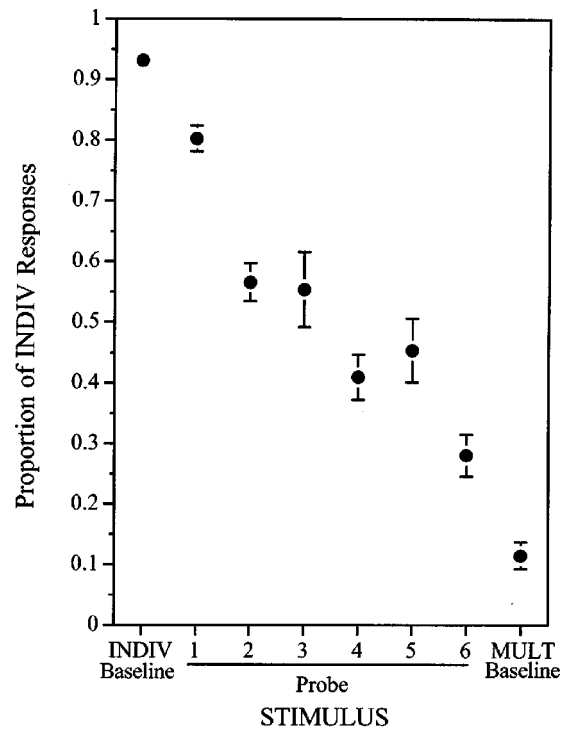


FIG. 3. Mean ( $\pm$ SEM) proportion of INDIV key responses as a function of stimulus class for experiment 1. The means show the data for all the exemplars in each test stimulus class (see Fig. 2) and the baseline stimuli.

an unfamiliar song segment (or segments) inserted in it were significantly higher than those for the test stimuli that had a MULT song segment (or segments) inserted in them [ $F(1,7) = 1333.353, p < 0.01$ ]. This familiarity effect was most pronounced among the test stimuli in which two song segments had been inserted, as shown by the significant in-

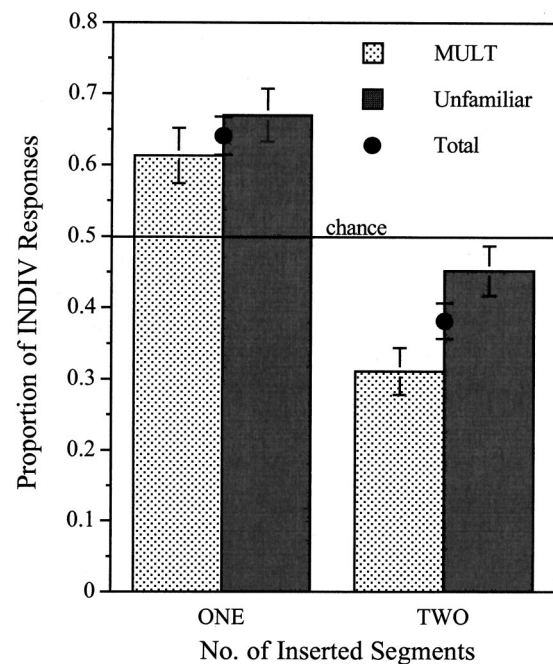


FIG. 4. Mean ( $\pm$ SEM) proportion of INDIV responses as a function of the test stimuli in experiment 1. Means show the data for all exemplars in which either one or two MULT song segments were inserted, and those in which either one or two unfamiliar song segments were inserted.

teraction between the number of inserted song segments and the familiarity (either MULT or unfamiliar) of those inserts [ $F(1,7) = 19.545$ ,  $p < 0.01$ ; see Fig. 4].

In contrast to the above results, varying the location at which a song segment was inserted into a test stimulus produced a limited effect on recognition. From Fig. 3, one can see that insertions at the first position (test stimuli 1 and 6, see Fig. 2) had smaller effects on classification than did insertions at the second and third positions. *Post hoc* comparisons among the different test stimuli show that responding to the type 1 test stimuli was significantly higher than that to types 1 and 2 ( $p < 0.05$ , Fisher's PLSD), whereas all comparisons among test stimuli 4–6 were not significant.<sup>1</sup>

## D. Discussion

The stimuli used in the first experiment were designed to test the idea that motifs at different locations in a sampling interval (e.g., those in the first few seconds) have differential effects on the discrimination and subsequent classification of a song bout. Because serial-position effects are well documented in both human and nonhuman memory tasks (Baddeley, 1990; Deacon and Rawlins, 1995; Wright and Rivera, 1997), we reasoned that similar effects in songbirds might make certain motifs at either the beginning or end of a sampling interval easier to remember. Thus, these motifs may exert greater control during classification decisions that follow the presentation of a song bout. The results of the first experiment suggest, however, that this is not the case. Varying the location of inserted motifs within a song bout had relatively little effect on classification. Among the test stimuli in which only a single segment from a MULT or unfamiliar song was inserted (nos. 1–3, see Fig. 1), insertions at the first segment had the smallest effect on recognition. This is inconsistent with the notion that primacy effects exert control over song discrimination behavior, because such effects should render insertions at the first portion of the song more salient, and thus have a large effect on classification. In addition, the fact that we observed no difference between substitutions made at the second and third segments is inconsistent with a strict criterion for recency effects. However, substitutions at both the second and third positions did have more substantial effects on recognition than substitutions at the first position. Therefore, one might consider this a modest demonstration of recency effects. The trend for more terminal motifs to exert greater control over discrimination in the present task may reflect the ability for starlings to hold sensory traces of different motifs in a short-term memory store. If, as is reasonable to assume, this memory store decays over time and is subject to retroactive interference, those motifs having the closest temporal proximity to the point at which the recognition choice is measured will exert the greatest control. As a caveat to the observed serial-position effects, note that our method of baseline training, in which subjects were forced to withhold responses until most of the stimulus had been presented, may have biased the observed lack of proximity effects. A better test for general primacy effects in songbirds would examine the recall for each motif independently, rather than as a function of a single classification judgment weighted by all the motifs in a

sampling interval. For our immediate purposes, however, it was necessary to quantify serial-position effects with respect to performance in the present task.

In contrast to the modest position effects observed during the first experiment, much more dramatic and general effects on discrimination behavior were elicited by variation in the proportion of familiar motifs composing a bout. The subjects were better at correctly classifying test stimuli in which two out of the three song segments were from the INDIV singer than those in which one out of three song segments was from the INDIV singer. The tendency to respond based on motif proportions was observed for all of the test stimuli, regardless of the exact sequence in which the song segments composing each exemplar were presented, and is consistent with the idea the starlings learn to classify conspecific songs by associating sets of motifs with different singers.

Finally, the fact that test stimuli composed of unfamiliar and INDIV motifs were easier to classify than those composed of MULT and INDIV motifs suggests that subjects recognized motifs in all the baseline training songs. Had subjects simply been listening for only INDIV motifs (or MULT motifs) in a pool of otherwise unclassified “noise,” the differences between the MULT and unfamiliar song segments would not have affected the classification of the test stimuli. As it was, hearing a MULT song segment in one of the test exemplars was more distracting than hearing an unfamiliar song segment in the same location (see Fig. 4). Thus, it appears that subjects had learned to recognize a substantial portion of the motifs from all, or at least many, of the baseline exemplars, both INDIV and MULT.

## III. EXPERIMENT 2: MOTIF SYNTAX EFFECTS

### A. Introduction

Another way in which some motifs may exert comparatively stronger (or weaker) control over classification responses is through their syntactic relationship to other motifs in a given bout. Earlier experiments have demonstrated that starlings are sensitive to the syntactical structure of familiar motif sequences in male song bouts, such that randomizing the order of motifs in a familiar bout lowers the recognition of that bout (Gentner and Hulse, 1998). The syntactic structure of a song bout may influence the perception of its constituent motifs in two different ways. First, the position of each motif relative to the other motifs in that bout may be important. For instance, the leading motifs in a sequence may facilitate, or “prime,” the recognition of trailing motifs. Receivers may also learn to recognize larger “chunks” or “sub-sequences” of motifs as single objects, such that violating the sequence of motifs in the chunk would lead to decreased recognition of the constituent motifs. Second, the position of each motif relative to the overall temporal structure of the song bout may be important. Because the motifs in a bout are presented over an extended period of time, when the subjects are trying to recognize a given exemplar, they may listen for specific motifs at specific points in the sampling interval. Violations of the second, more global, syntax rules may affect the perception of individual motifs

TABLE II. The 12 motif patterns used to generate the test stimuli for experiment 2. Letters and numbers denote particular motifs drawn from the INDIV or MULT baseline exemplar, respectively. Patterns marked with an asterisk have more motifs from one or the other baseline exemplars, and were presented in the form shown below and in another form (not shown) in which the first two motifs in the sequence (denoted by the underscore) were removed. The notation used here should not be confused with that used in Table I.

		Baseline stimuli pair																	
INDIV:	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	Q	r	
MULT:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
		Test stimuli																	
Stimulus type	Sequence																		
I	a	2	c	4	e	6	g	8	i	10	k	12	m	14	o	16	q	18	
II	1	b	3	d	5	f	7	h	9	j	11	l	13	n	15	p	17	r	
III*	a	b	3	4	e	f	7	8	i	j	11	12	m	n	15	16	q	r	
IV*	1	2	c	d	5	6	g	h	9	10	k	l	13	14	o	p	17	18	
V	a	b	c	4	5	6	g	h	i	10	11	12	m	n	o	16	17	18	
VI	1	2	3	d	e	f	7	8	9	j	k	l	13	14	15	p	q	r	
VII	a	8	c	18	e	3	g	15	i	16	k	9	m	2	o	7	q	17	
VIII	11	b	4	d	1	f	5	h	13	j	6	1	14	n	10	p	12	r	
XI*	a	b	4	10	e	f	16	5	i	j	14	2	m	n	8	7	q	r	
X*	15	9	c	d	13	1	g	h	12	3	k	1	11	17	o	p	18	6	
XI	a	b	c	18	13	8	g	h	i	15	10	6	m	n	o	5	14	2	
XII	9	16	4	d	e	f	12	7	11	j	k	1	1	17	3	p	q	r	

and thus recognition of the song. In the second experiment, we vary both these syntactical features in an attempt to make some of the motifs in a given bout more or less salient, and thereby bias classification of the test stimuli.

## B. Methods

### 1. Stimuli

For the stimuli used in experiment 2, we selected one pair from exemplars from each baseline stimulus set (set 1: A3 and C2; set 2: B3 and C3, see Table I). As in the first experiment, each pair comprised one INDIV exemplar and one MULT exemplar that had been very accurately recognized (above 90% correct) by all subjects at asymptotic performance on the baseline procedure. The same pair of stimuli was used for all the subjects trained with each baseline set. In addition, the baseline exemplars chosen for experiment 2 were different than those used to generate the stimuli for experiment 1. Table II shows the patterns used to generate the test stimuli for experiment 2. We purposely chose pairs of INDIV and MULT baseline exemplars that contained the same number of motifs, and parsed each exemplar, using digital editing software (SoundDesigner II, Digidesign-Avid), according to the natural divisions between motifs. Each test stimulus was created by combining half of the motifs from an INDIV exemplar with half of the motifs from the MULT exemplar with which it had been paired. Each test stimulus was ~15 s long.

In all, 12 different primary types of test stimuli were created. As shown in Table II, test stimulus types I and II contained the same number of INDIV and MULT motifs, but the sequence of motifs was arranged so that no two INDIV or MULT motifs appeared in immediate succession. At the same time, the relative position of each motif with respect to its original position in the baseline exemplar was maintained. Stimulus types I and II are complements of one another, such

that all of the motifs omitted in type I are contained in type II, and vice versa. Stimulus types III–VI followed a similar pattern to that of I and II except for the following difference. In stimulus types III and IV the motifs from the INDIV and MULT exemplars were organized into consecutive runs of two motifs, so that two motifs from the INDIV exemplar were followed by two motifs from the MULT exemplar. In stimulus types V and VI, the motifs from each of the INDIV and MULT baseline exemplars were organized into consecutive runs of three. Again, for all stimulus types I–VI the relative position of each motif with respect to its original position in the baseline exemplar was maintained. Stimulus types VII–XII follow the general pattern of types I–VI, respectively, except that the relative position of the MULT motifs (with respect to the original baseline stimulus) was not maintained. For example, in type VII stimuli the sequence of INDIV motifs is identical to that for type I, but the sequence of MULT motifs is dramatically altered from its original order in the baseline exemplar (see Table II). The motif placed at each location in types VII–XII was chosen at random (without replacement) from all the MULT motifs in a given baseline exemplar, with the constraint that a given motif could not occupy the same location in the both the baseline and test exemplar. In addition, none of the MULT motifs in stimulus types VII–XII were in an appropriate position relative to other MULT motifs from that same song (e.g., the motif sequence 4–5–6 does not appear anywhere in stimulus type XI or XII).

If the original syntax of motifs is important, then recognition should improve as one looks from stimulus type VII to XII, because these syntax cues have been removed from the MULT motifs and progressively strengthened among the INDIV motifs. In addition, if subjects are expecting to hear certain motifs at certain points in the presentation of an exemplar (e.g., near the beginning), then recognition of the

INDIV motifs ought to be better for types I–VI than for types VII–XII because these cues are missing from the MULT motifs in the later stimuli. Lastly, for stimulus types III, IV, IX, and X the sequences shown in Table II bias the test stimuli so that exemplars in these classes have more INDIV motifs than MULT motifs (or vice versa). For these stimuli, we created an additional version of each exemplar in which the first two motifs were omitted so that the proportions of INDIV and MULT motifs in each exemplar were equal. In all, each subject was exposed to 16 different test stimuli created from a particular pair of INDIV and MULT baseline exemplars.

## 2. Procedure

Following completion of the first experiment, all subjects were returned to the baseline procedure (at 70% reinforcement) and allowed to stabilize at asymptotic performance. Subjects then began the second experiment. The procedure for experiment 2 was identical to that described for experiment 1, except that different test stimuli were presented. During this experiment, one subject developed a strong position bias for all the stimuli, and his data were excluded from the analysis.

## C. Results

The subjects maintained accurate classification of the baseline stimuli over the duration of the second experiment. They responded correctly to the INDIV and MULT baseline stimuli ~90% of the time (mean for INDIV =  $89.5\% \pm 3.8\%$ , mean for MULT =  $91.7\% \pm 1.6\%$ ). Accordingly, the difference between the proportion of INDIV responses made to the INDIV and the MULT baseline stimuli was significant [ $F(1,6) = 249.281, p < 0.0001$ ]. At the same time, the mean proportion of INDIV responses to the test stimuli ( $62.6\% \pm 1.3\%$ ) was significantly less than that for the INDIV baseline stimuli ( $t = -23.075, p < 0.0001$ , two-tailed), and significantly greater than that for the MULT baseline stimuli ( $t = 36.035, p < 0.0001$ , two-tailed). This fact suggests that subjects recognized the test stimuli as different from the baseline stimuli. In addition, the overall level of responding to the test stimuli was not significantly different among the two groups of subjects trained on different baseline stimulus sets [see Table I;  $F(1,5) = 0.382$ , NS], and the distribution of responses among the different test stimuli did not differ between these groups [ $F(19,95) = 1.649$ , NS]. Thus, for subsequent analysis, the data were pooled across all subjects.

Although the test stimuli were treated differently than the baseline stimuli, the subjects appear to have responded to all the test stimuli in the same manner. We observed no significant differences among the mean proportion of INDIV responses made to the different test stimuli [ $F(19,114) = 0.711$ , NS].

## D. Discussion

The second experiment examined song classification when the syntax of motifs composing a song bout varied in two ways. First, the position of each motif relative to other motifs from the same singer was varied. Second, the position of motifs relative to the overall temporal structure of a bout

was varied. In both these cases, we failed to find any significant effect on recognition performance. The subjects generally responded in the same way to all of the test stimuli by pecking the INDIV key approximately 63% of the time. The test stimuli for this second condition each contained an equal number of INDIV and MULT motifs, but varied according to the syntactic structure of the MULT motifs in a given exemplar. We hypothesized that as syntactical cues were removed from the MULT motifs, they would become less distracting, and it would therefore bias the subjects to respond to the test stimulus as if it was an INDIV bout. The data are not consistent with this hypothesis. None of the syntactic variations among the test stimuli elicited any corresponding variation in the subjects' responses.

Although we found no support for the role of syntactical cues in the present context, it should be pointed out that this is a very strong test of that hypothesis. In fact, earlier data do support the role of syntactical cues in more straightforward recognition tasks where starlings are better at recognizing naturally ordered strings of motifs compared to randomly ordered strings (Gentner and Hulse, 1998). The present task was more complex in that we attempted to use the absence of syntactical cues to make some of the motifs in a bout less distracting. In this context, where sets two of motifs are, in a sense, competing for a classification decision, removing syntactical cues from one set of motifs has no detectable effect on the ability to recognize the other motifs in that bout.

## IV. EXPERIMENT 3: MOTIF PROPORTION EFFECTS

### A. Introduction

As detailed in the general introduction, the existing evidence suggests that starlings rely on variation in the proportions of familiar motifs in order to classify conspecific songs. In the third experiment, we test this hypothesis directly by constructing a large corpus of stimuli in which the percentage of familiar motifs taken from two different training singers is systematically varied from song bouts composed of 10% singer 1 motifs and 90% singer 2 motifs to bouts composed of 90% singer 1 motifs and 10% singer 2 motifs. We then observe how the subjects classify these composite songs. The results of the first two experiments provide important information regarding the effects that serial position and motif syntax have on the classification of conspecific song. In the third experiment, this information is used to build appropriate controls into the set of test stimuli.

### B. Methods

#### 1. Stimuli

For the stimuli used in experiment 3, we again selected two pairs of baseline exemplars, one pair from each set of baseline stimuli (set 1: A4 and B2; set 2: B4 and E3, see Table I). Each pair contained one INDIV baseline exemplar and one MULT baseline exemplar. The baseline exemplars chosen for experiment 3 were different than those used in the first two experiments, and the same pairs of baseline stimuli were used for all subjects trained on a given set (see Table I). Each of the four chosen baseline exemplars was parsed into

TABLE III. Examples of motif patterns used to generate the test stimuli for experiment 3. Different letters and numbers denote unique 2–3 motif segments taken from one of the INDIV or MULT baseline exemplars, respectively. Note that for all of the sequence patterns, the original position of each segment in the baseline exemplar has been maintained. The notation used here should not be confused with that in Table I.

		Baseline stimuli								
INDIV:	a	b	c	d	e	f	g	h	i	j
MULT:	1	2	3	4	5	6	7	8	9	10
		Test stimuli								
Stimulus class	Sequence									
I	a	b	3	d	e	f	g	h	i	j
I	a	b	c	d	e	f	g	8	i	j
II	a	2	c	4	e	f	g	h	i	j
II	a	b	c	d	e	6	g	h	9	j
III	a	2	3	d	5	f	g	h	i	j
III	a	b	c	d	e	6	g	8	i	10
IV	1	2	c	4	5	f	g	h	i	j
IV	a	b	c	d	e	6	7	h	9	10
V	1	2	3	4	5	f	g	h	i	j
V	a	b	c	d	e	6	7	8	9	10
VI	a	b	3	d	e	6	7	8	9	10
VI	1	2	3	4	5	f	g	8	i	j
VII	a	2	c	4	e	6	7	8	9	10
VII	1	2	3	4	5	6	g	h	9	j
VIII	a	2	3	d	5	6	7	8	9	10
VIII	1	2	3	4	5	6	g	8	i	10
IX	1	2	c	4	5	6	7	8	9	10
IX	1	2	3	4	5	6	7	h	9	10

ten smaller segments which themselves contained two to three whole motifs. Using digital editing software, the segments from each pair of baseline exemplars were then combined in varying proportions, so that each test stimulus contained ten segments. Table III shows representative patterns for the different classes of test stimuli. Nine different classes of test stimuli were created in which the percentage of INDIV and MULT motifs in a given exemplar varied from 90% to 10% in increments of 10%. For example, in the class I test stimuli the ratio of INDIV:MULT segments was 9:1, in the class II stimuli the ratio of INDIV:MULT segments was 8:2, and in the class III stimuli the ratio of INDIV:MULT segments was 7:3. This pattern of combining song segments from the baseline exemplars meant that for classes I–IV and VI–IX, either the MULT or INDIV motifs, respectively, made up the minority portion. These eight classes of test stimuli were further subdivided into two types of exemplars, those in which the minority motifs occurred somewhere within the first five segments (i.e., the first half of the bout), and those in which the minority motifs occurred somewhere within the last five segments. The remaining class of test stimuli (class V: 50% INDIV–50% MULT) were composed of five segments from the INDIV baseline exemplar followed by five segments from the MULT baseline exemplar, or vice versa. Each test stimulus was ~ 15 s long.

The total number of possible permutations for this stimulus set is very large. Therefore, for each stimulus class, a subset of all the possible permutations was selected to achieve a reasonable number of test stimuli that met the criteria outlined above and also covered the range of possible permutations in that class. For stimulus classes I, IV, VI, and IX, six different exemplars per class were constructed. For

classes II, III, VII, and VIII, 12 different exemplars per class were constructed. There were only two exemplars in class V. Each subject was exposed to 74 different test stimuli. For all of the test stimuli in experiment 3, the location of each segment with respect to its position in the original baseline exemplar was maintained (see Table III).

## 2. Procedure

Following completion of the second experiment, all subjects were returned to the baseline procedure (at 70% reinforcement) and allowed to stabilize at asymptotic performance. Subjects then began the third experiment. The procedure for experiment 3 was identical to that described for experiment 1, except that different test stimuli were presented.

## 3. Statistics

The statistical analysis followed that given in the general methods, with one exception. In order to balance the structure of the models used to analyze the interaction between the different test stimulus classes and the structure of the stimuli with each class, some of the data from classes with more than six exemplars were not used. For the larger test stimulus classes (II, III, VII, and VIII), we selected six stimuli at random from each of the classes and analyzed the responses across all classes with a repeated measures ANOVA. Identical tests using different exemplar samples from the larger classes yielded results similar to those reported below. The separate repeated measures ANOVAs run on the stimuli within each class used data from all the exemplars.

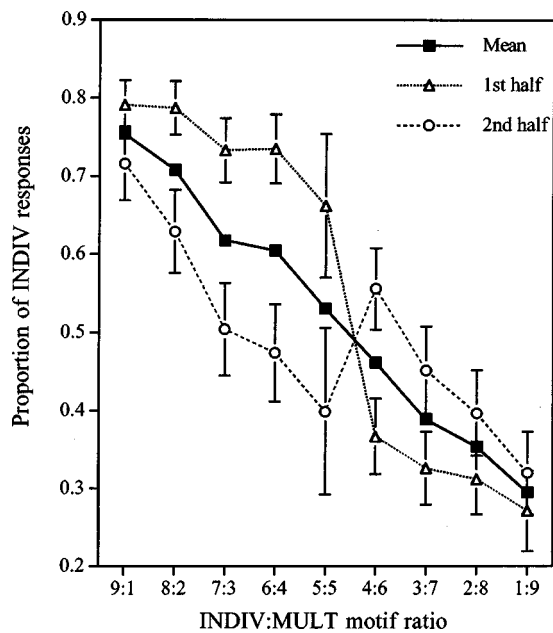


FIG. 5. Mean ( $\pm$ SEM) proportion of INDIV responses as a function of test stimulus class in experiment 3. The mean shows the data from all exemplars in each class (see Table III). Each class is labeled according to the ratio of INDIV:MULT motifs that comprise its exemplars. Data are further divided according to the location of the minority motifs (either the first or last half of the exemplar, see Table III).

### C. Results

All of the subjects maintained very accurate classification of the baseline stimuli over the duration of the third experiment, making correct responses to the INDIV and MULT baseline stimuli  $\sim 91\%$  of the time (mean for INDIV =  $91.1\% \pm 1.7\%$ , mean for MULT =  $90.7\% \pm 2.8\%$ ). Accordingly, the mean proportion of INDIV responses made to the INDIV baseline stimuli was significantly greater than the proportion made to the MULT baseline stimuli [ $F(1,7) = 568.468$ ,  $p < 0.0001$ ]. From Fig. 5, it is apparent that the proportion of INDIV responses made to each class of test stimuli varied along with the proportions of INDIV and MULT motifs present in each stimulus. The variation among these responses to the different test stimuli was significant [ $F(73,438) = 12.870$ ,  $p < 0.0001$ ]. As in the first two experiments, performance between the two groups of subjects trained on the different baseline stimulus sets did not differ significantly in either the overall level of response to the test stimuli [ $F(1,6) = 0.117$ , NS], or in the pattern of responding observed among the different test stimuli [ $F(73,438) = 0.759$ , NS]. Subsequent analyses pooled the data from all subjects.

To examine more subtle patterns in the responses made to the different classes of test stimuli we analyzed the variance across stimulus test classes I–IV and VI–IX using a single repeated measures ANOVA (see Sec. IV B). Again, the probability of making an INDIV response varied significantly among the test stimuli [ $F(7,49) = 47.819$ ,  $p < 0.0001$ ]. We observed no difference in the mean proportion of INDIV responses made to exemplars with substitutions in the first half of the bout and those with substitutions in the last half of the bout [ $F(1,7) = 1.584$ , NS]. However,

there was a significant interaction between the location of a substitution and the eight different classes of test stimuli in this analysis [ $F(7,49) = 9.995$ ,  $p < 0.0001$ ]. This last finding suggests that the location of a substitution (i.e., the minority motifs, see Table III) was important, but only in some of the test stimulus classes (see Fig. 5). To explore this effect further, we ran separate repeated measures ANOVAs on all the test stimuli in each of the different classes. For classes II–VII, the location of the minority motifs had a significant effect on the mean proportion of INDIV responses made to a given stimulus. For classes II–V, subjects were significantly more likely to make an INDIV response when the MULT motifs appeared in the first half rather than the second half of the stimulus. For classes VI and VII, this trend was reversed, with subjects more likely to make an INDIV response when INDIV motifs appeared in the second half rather than the first half of the stimulus. These results can be understood as position effects, and are consistent with the modest recency effects observed in the first experiment. Recall that in experiment 1, motifs in the terminal portion of the sampling interval exerted greater control over the subjects' responses than did those in the initial portion. The serial-position effects for experiment 3 can be seen as the difference between the dotted and dashed curves in Fig. 5, and the motif proportion effect as the slope of each curve.

### D. Discussion

In the third experiment, we examined how recognition was effected by variation in the relative proportions of familiar motifs composing a song bout. To do this, we created several sequences of motifs (i.e., song bouts) in which the relative proportions of motifs from two of the baseline-training singers were systematically varied. We then observed the pattern of responses to these bouts. Classification of the test stimuli was strongly controlled by the proportions of familiar motifs in each exemplar. When a bout contained more motifs from the INDIV singer, the subjects tended to classify that song as an INDIV bout. When a bout contained more motifs from the MULT singer, subjects tended to classify that song as a MULT bout. As Fig. 5 shows, the generalization gradient between INDIV and MULT responding closely follows the variation in relative proportions of motifs from the two singers contained in each exemplar. In addition, within each class of test stimuli, all of the different permutations elicited similar proportions of INDIV responses. This last fact suggests that no single motif exerted any greater control over recognition than did any other. By extension, then, it appears that large subsets of motifs were associated with each response.

The data from the third experiment also provide converging support for the manner in which motifs at different locations of the sampling interval control recognition. In general, motif substitutions made in the later half of a bout exerted greater control over recognition than those in the first half of the bout. As with the recency effects observed in the first experiment, those seen in experiment 3 are likely due to the short delay between motifs presented in the second half of the sampling interval and the point at which the subject makes a choice on any given trial. Interestingly, these serial-

position effects are not consistent across the different test stimulus classes, but instead are centered around those exemplars in which the INDIV:MULT ratio is closest to one, and at a maximum when the ratio equals one (see Fig. 5). This pattern suggests two conclusions. First, subjects appear to integrate acoustic information over the time span of several motifs. Combined with the negative results from experiment 2, in which we observed no differences in responding to any of the test stimuli, we can estimate a lower bound on this integration window of approximately four to six motifs. Because even those exemplars in which a single minority segment appeared in the first half of the stimulus, and those with insertions at the first position in experiment 1, were treated differently than the baseline stimuli, the upper bound to the integration window appears to be at least 10 s. Second, because the serial-position effects are inversely related to the bias in motif proportions (see Fig. 5), and at a maximum when the motif proportion bias is minimized (INDIV:MULT ratio=1), the stimulus control they elicit appears to be secondary to that elicited by the ratio of familiar motifs.

## V. GENERAL DISCUSSION

The results of this study suggest that when starlings are compelled to classify conspecific songs, they do so by memorizing large numbers of unique song components (i.e., motifs), and then organize subsets of these motifs into separate classes. As a cognitive strategy, classifying songs according to their component structure represents a parsimonious method of dealing with these complex acoustic signals under more natural conditions. Because individual starlings tend to possess unique motif repertoires, disjoint sets of motifs will generally correspond to individual identity. Therefore, attending to the motif structure will capture a significant portion of the individual variation in the signal, which, of course, is a requisite to individual vocal recognition. Given this potential source for individual variation, it is tempting to consider the possibility that perceptual processing mechanisms similar to those described here serve as the basis for individual vocal recognition, at least in starlings. Although we suggest that this is likely to be the case, it is important to remember that individual vocal recognition requires more than the discrimination and classification of song. Like most real-world behaviors, individual vocal recognition is likely to involve complex multi-modal processing mechanisms that integrate nonarbitrary visual and auditory information, as well as a level of categorical reference that we have not addressed here. Nonetheless, it is difficult to imagine how individual vocal recognition might occur without the discrimination and classification of song, and our data suggest a likely behavioral mechanism for such cognition.

Although the operant task detailed here is an idealization of vocal recognition in more natural contexts, the two behaviors share a functional equivalence in that they require subjects to associate specific acoustic events with external referents. Under natural conditions, these referents may be locations in space or the visual images of individual conspecifics, whereas in the operant task, the referent is a response button at a given location. We designed the operant contin-

gencies so that classification covaried with individual identity, with the intent that the subjects would engage functional vocal recognition mechanisms to solve the task. In fact, previous data support this idea by suggesting that behaviors observed under similar conditions involve substantially more than the rote memorization of arbitrary stimuli. Correct classification acquired with these same stimulus sets will readily transfer to novel song bouts containing both familiar motifs, and those composed of entirely novel-motifs sung by the training singers (Gentner and Hulse, 1998). Whereas the transfer to songs containing novel instances of familiar motifs holds for all starlings tested to date (Gentner and Hulse, 1998; Gentner *et al.*, 2000), the transfer to songs composed of entirely novel motifs sung by the training singers obtains only when the subjects have had experience with the singers prior to any exposure in the operant apparatus (Gentner *et al.*, 2000). In other words, individual vocal recognition can be demonstrated in starlings, but only after the subjects have had direct experience with the songs that different birds sing. Therefore, when individual vocal recognition has been established in an operant context, it appears to be essentially a problem in the classification of familiar songs. It does not appear to be mediated by the use of more global acoustic properties imparted to all, or some subset of, the motifs sung by a given individual (so called "voice-characteristics"). The results of the present study strongly suggest that the discrimination and classification of familiar song is controlled by the relative proportions of motifs from different singers comprising a bout.

Based on the results of this study, we can begin to outline a minimum set of cognitive processes that are likely to serve as the basis for individual vocal recognition in starlings. First, starlings need to be able to dissociate strings of motifs (i.e., songs) from irrelevant background noise and the songs of other individuals occurring simultaneously. This is analogous to the cocktail party effect in human listeners (Cherry and Taylor, 1954), and is a problem in auditory stream segregation (Bregman, 1990; Wisniewski and Hulse, 1997; MacDougall-Shackleton *et al.*, 1998). Once a relevant stream has been segmented from others, that song must be parsed into its constituent motifs, and then sets of motifs that occur in a single bout must be grouped into separate classes. Novel motifs occurring in subsequent song bouts must then be assigned to different classes, perhaps based on their proximity to already familiar motifs. Only after a suitable "library" of motif classes has been established can accurate recognition take place. Within this context, decisions about which singer produced a given song bout are then based on judgments about the relative proportions of motifs from each class that occur within some sampling interval. These putative processes require a mechanism for consolidating neural representations of each motif into a long-term memory store, and suggest a memory retrieval/decision mechanism (or mechanisms) in which similarity judgments between new and stored motifs are made. The results of this study suggest that the later mechanism is integrating similarity judgments made for many motifs in order to reach a decision about individual identity. Although the proximate mechanisms that give rise to these, and in fact most, cognitive processes remain

largely unknown, we hope that the present study will provide at least some of the groundwork necessary for their subsequent investigation.

## ACKNOWLEDGMENTS

The work reported here was supported by NSF Grant No. IBN 9317868 to Stewart Hulse. We thank three anonymous reviewers and D. Wesley Grantham for helpful comments on earlier drafts of the manuscript.

<sup>1</sup>Recall that for each of the two baseline stimulus sets, two pairs of INDIV and MULT exemplars were used to construct the test stimuli for experiment 1. This dimension in the design provided a within-subject control for effects restricted to a single exemplar from the baseline stimulus set, as opposed to those involving more general stimulus control. Including exemplar information in the analysis yielded a significant effect for the serial-position test stimuli [ $F(2,14) = 8.841, p < 0.005$ ]. Closer inspection of the data revealed that this effect was due to a drop in performance when a single substitution was made at the second segment in one of the test stimuli (type 2, Fig. 1). Rather than arguing for reliable serial-position effects, this suggests that in this particular exemplar, motifs located in the central region exerted greater control over the discrimination than did motifs at this location in the other test stimuli.

Adret-Hausberger, M., and Jenkins, P. F. (1988). "Complex organization of the warbling song in starlings," *Behaviour* **107**, 138–156.

Baddeley, A. (1990). *Human Memory: Theory and Practice* (Allyn and Bacon, Needham, MA).

Beecher, M. D. (1988). "Kin recognition in birds," *Behav. Genet.* **18**, 465–482.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge).

Chaiken, M. L., Böhrner, J., and Marler, P. (1994). "Repertoire turnover and the timing of song acquisition in European starlings," *Behaviour* **128**, 25–39.

Cherry, E., and Taylor, W. (1954). "Some further experiments upon the recognition of speech, with one and with two ears," *J. Acoust. Soc. Am.* **26**, 554–559.

Deacon, R. M. J., and Rawlins, J. N. P. (1995). "Serial position effects and duration of memory for non-spatial stimuli in rats," *J. Exp. Psychol.* **21**(4), 285–292.

Eens, M. (1997). "Understanding the complex song of the European starling: An integrated approach," *Adv. Study Behav.* **26**, 355–434.

Eens, M., Pinxten, R., and Verheyen, R. F. (1989). "Temporal and sequential organization of song bouts in the starling," *Ardea* **77**, 75–86.

Eens, M., Pinxten, R., and Verheyen, R. F. (1991). "Organization of Song

in the European Starling—Species-Specificity and Individual Differences," *Belgian J. Zool.* **121**(2), 257–278.

Gentner, T. Q. (1999). "Behavioral and neurobiological mechanisms of song perception among European Starlings," unpublished Ph.D. Thesis, Johns Hopkins University.

Gentner, T. Q., and Hulse, S. H. (1998). "Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*," *Anim. Behav.* **56**, 579–594.

Gentner, T. Q., Hulse, S. H., Bentley, G. E., and Ball, G. F. (2000). "Individual Vocal Recognition and the Effect of Partial Lesions to HVc on Discrimination, Learning, and Categorization of Conspecific Song in Adult Songbirds," *J. Neurobiol.* **42**(1), 117–133.

Hausberger, M. (1997). "Social influences on song acquisition and sharing in the European starling (*Sturnus vulgaris*)," in *Social Influences on Vocal Development*, edited by C. Snowden and M. Hausberger (Cambridge U.P., Cambridge).

Hausberger, M., and Cousillas, H. (1995). "Categorization in birdsong: From behavioral to neuronal responses," *Behav. Processes* **35**(1–3), 83–91.

Hulse, S. H. (1995). "The discrimination-transfer procedure for studying auditory perception and perceptual invariance in animals," in *Methods in Comparative Psychoacoustics (Vol. 10)*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhäuser Verlag, Basel, Switzerland).

Lind, H., Dabelsteen, T., and McGregor, P. K. (1997). "Female great tits can identify mates by song," *Anim. Behav.* **52**, 667–671.

MacDougall-Shackleton, S. A., Hulse, S. H., Gentner, T. Q., and White, W. (1998). "Auditory scene analysis by European starlings (*Sturnus vulgaris*): Perceptual segregation of tone sequences," *J. Acoust. Soc. Am.* **103**, 3581–3587.

Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide* (Cambridge U.P., Cambridge).

O'Loughlen, A. L., and Beecher, M. D. (1997). "Sexual preferences for mate song types in female song sparrows," *Anim. Behav.* **53**, 835–841.

Stoddard, P. K. (1996). "Vocal recognition of neighbors by territorial passerines," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Comstock/Cornell, Ithaca).

Wiley, R. H., Hatchwell, B. J., and Davies, N. B. (1991). "Recognition of individual males' songs by female Dunnocks: a mechanism increasing the number of copulatory partners and reproductive success," *Ethology* **88**, 145–153.

Wisniewski, A. B., and Hulse, S. H. (1997). "Auditory scene analysis in European starlings (*Sturnus vulgaris*): Discrimination of song segments, their segregation from multiple and reversed conspecific songs, and evidence for conspecific song categorization," *Journal of Comparative Psychology* **111**(4), 337–350.

Wright, A., and Rivera, J. (1997). "Memory for auditory lists by rhesus monkeys," *J. Exp. Psychol.* **23**(4), 441–449.



# Observer weighting of monaural level information in a pair of tone pulses

Mark A. Stellmack and Neal F. Viemeister

*Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455*

(Received 17 September 1999; revised 4 January 2000; accepted 3 March 2000)

A correlational analysis was used to assess the relative weight given to the levels of two monaurally presented tone pulses for interpulse intervals (IPIs) ranging from 2–256 ms. In three different experimental conditions, listeners were instructed to discriminate the level of the first pulse, the level of the second pulse, or the difference between the levels of the two pulses. The level of the target pulse was chosen randomly and independently from trial to trial from a Gaussian distribution. The level of the nontarget pulse was either fixed at 75 dB SPL or varied in the same manner as the level of the target. In the tasks in which one pulse was to be ignored, listeners gave increasing weight to the nontarget component as IPI decreased. Listeners weighted the level information in the pulses appropriately only when the IPI approached 256 ms. When the listeners were instructed to compare the pulse levels to one another, two of three listeners weighted the levels optimally at all IPIs, while the third listener did so only at the longest IPI. For the two listeners who weighted the pulses optimally, a minimum in performance was achieved at IPIs around 16–32 ms. Intensity discrimination thresholds were also measured for one pulse in the presence of a second fixed pulse for IPIs of 2–256 ms. Thresholds were higher in all the two-pulse conditions relative to a one-pulse condition, and were dependent on the level of the nontarget pulse but not on IPI. The results indicate that level information is integrated to some extent over fairly long durations, but not in a manner that is consistent with simple temporal integration. © 2000 Acoustical Society of America.

[S0001-4966(00)02506-6]

PACS numbers: 43.66.Mk, 43.66.Fe [SPB]

## INTRODUCTION

Auditory information is integrated or accumulated over time, a process which is exhibited in a number of ways. In particular, detection thresholds decrease with increasing signal duration, and temporal resolution is limited by temporal integration, such that rapid changes that occur in an acoustical signal are combined or “smoothed over” by the auditory system. The “resolution-integration paradox” refers to the fact that studies of temporal resolution in the auditory system (e.g., Viemeister, 1979; Forrest and Green, 1987) generally produce estimates of around 3 ms for the integration time, while studies of the detection of signals as a function of duration yield integration times on the order of hundreds of ms (e.g., Green, 1960; Green and Swets, 1966). Often, this paradox is explained by assuming that there are two systems with different time constants, and that the system that is used depends on the particular listening task (Green, 1985; de Boer, 1985).

Viemeister and Wakefield (1991) proposed a multiple-looks hypothesis to attempt to account for long-term temporal integration in a system with a short integration time. The idea is that the auditory system accumulates many short-term “looks” at the signal, retains those looks in memory, and selectively uses them as the basis for decisions. In order to “selectively use” the looks, a listener might give equal consideration to the information in all looks (e.g., in detecting a long-duration signal), the listener might only attend to particular looks while ignoring or discarding others (e.g., in detecting a signal at a particular temporal position in a noisy background), or the listener might compare looks to one another (e.g., in identifying a temporal pattern of loudness

changes). (These hypothetical listening strategies represent the ideal strategy for each situation.) The manner in which the looks are utilized is given by a weighting function that describes the relative influence of each look on a decision. In order to account for actual temporal integration data for the detection of long-duration stimuli, Viemeister and Wakefield proposed that listeners apply a weighting function in which early portions of the signal are given relatively little weight, while the relative weight increases for later portions of the signal. Consistent with the notion that a long-duration stimulus is effectively processed as many short-term looks, Viemeister and Wakefield found that detection thresholds for a signal consisting of two 10-ms tone pulses that are 100 ms apart are unaffected by the level of a 50-ms noise burst temporally centered between the tone bursts.

Buus (1999) examined the multiple-looks hypothesis using an analysis in which he computed the perceptual weight given to a number of discrete short-duration pulses that comprised a signal to be detected. In a two-interval forced-choice task, the signal consisted of 25-ms, 1-kHz tone pulses (Buus *et al.*, 1996). On each trial, the levels of the signal pulses in the signal interval were selected randomly and independently from a distribution with a particular mean level and a standard deviation of either 3 or 6 dB, in separate conditions. In order to compute the weight given to a particular pulse, or temporal look, the data were sorted into ten categories based upon the level of that look on each trial. Then  $d'^2$  was plotted as a function of the average squared intensity of the looks in each of the ten categories. This was done separately for each temporal look. Thus for a signal consisting of six tem-

poral segments, six different plots of  $d'^2$  vs intensity-squared were produced. Buus showed that the squared perceptual weight given to a particular look is proportional to the slope of the line through these data for the given look. In other words, if a particular look is given a large perceptual weight, its level on a given trial will have a large influence on the detectability of the signal as a whole, and  $d'^2$  will vary to a large extent with intensity-squared, resulting in a relatively large slope value for the line relating  $d'^2$  to intensity-squared. On the other hand, if a particular look is given no perceptual weight, then detectability of the signal will not depend upon the level of that look on any given trial, and  $d'^2$  will be unrelated to intensity-squared (i.e., the function will have a slope of zero).

Buus found that detection of the signal was based on an optimal combination of the information in the individual temporal looks. Each look was given approximately equal weight by the listeners in detecting the signal, which is appropriate given that each look was equally informative overall because each look had the same distribution of levels across trials. This result suggested that detectability of a long-duration signal is based upon an optimal combination of the information in a number of shorter segments of the signal, a result that Buus concluded was consistent with the general form of the multiple-looks hypothesis described by Viemeister and Wakefield (1991). It should be noted, however, that equal weights also would be computed for the individual pulses of the signal if energy were integrated over the entire duration of the signal, that is, if the entire duration of the signal actually were treated as a single "look" by the auditory system. (In addition, somewhat different results were obtained in a separate condition in which a second off-frequency masker was comodulated with the on-frequency masker.) Similarly, in a sample-discrimination paradigm, Lutfi (1990) asked listeners to discriminate changes in the frequency, level, or duration of a sequence of pulses. Using a COSS (conditional on single stimulus) analysis to compute observer weights (Berg, 1989), Lutfi also measured essentially equal weights for all pulses, which for this task, as in the detection task of Buus (1999), is the optimal weighting strategy.

One implication of the multiple-looks hypothesis is that the listener accumulates information about a number of short-term looks in memory and the listener then can access selectively the information in the individual looks depending on the task at hand. In other words, while Lutfi (1990) and Buus (1999) showed that the listener gives equal weight to the individual looks in a situation where all looks contribute equally to the "signal," the multiple-looks hypothesis would also predict that the listener can ignore the information in uninformative temporal segments of a stimulus. For example, as described earlier, Viemeister and Wakefield (1991) showed that detection thresholds for a signal consisting of two 1-kHz tone pulses were unaffected by increases or decreases in the level of noise during the time between the two pulses. This seems to support the notion that relatively short, temporally discrete elements of a stimulus can be used or ignored by the listener as necessary.

One question that arises is: In a task in which not all

temporal segments of the stimulus are relevant, does the pattern of perceptual weights indicate that listeners are able to ignore the irrelevant information? Gilkey and Robinson (1986) computed weights for individual temporal segments of a stimulus interval in which listeners were instructed to detect a 500-Hz tonal signal in a broadband noise masker. The 100-ms signal was approximately temporally centered in the 148-ms masker, so that in a signal-bearing interval (of the single-interval, yes-no task) there was a masker fringe preceding and trailing the signal. This allowed the authors to compute weights for temporal segments of the stimulus that did not contain the signal as well as those that contained both signal and noise. Generally, it was observed that listeners either gave little weight to the leading masker fringe, indicating that they ignored the initial temporal segment of the stimulus, or listeners gave slightly negative weight to the leading masker fringe, indicating that they compared the noise-alone segment to later segments that might potentially contain the signal. For all listeners, the perceptual weights rose during the first half of the signal, while the weights were inconsistent across listeners for the second half of the signal. Rather large weights were computed for some listeners for the final segment of the interval, which also always consisted of masker alone (the trailing fringe). These results give some indication that listeners use, or are not able to ignore, temporal portions of the stimulus that never contain the signal. In addition, the weights computed by Gilkey and Robinson gradually changed during the duration of the stimulus, while Buus (1999) computed weights that were essentially equal for the entire stimulus duration. The source of these differences is not clear. The maskers in both experiments were of longer duration than the signals, resulting in temporal fringes preceding and following the signals. One difference between these studies is that Gilkey and Robinson used a 100-ms tone as the signal while Buus used a number of 25-ms tone pulses which essentially resulted in an amplitude-modulated signal. Furthermore, the weights computed by Buus were based on the intensity of only the signal in each temporal segment of the stimulus. The intensity of the masker in each temporal segment, an additional source of variability in the computation of weights, was ignored.

Lutfi (1992) performed a sample-discrimination experiment in which only one tone of a 10-tone sequence was identified as the target. A different sequence was presented in the two intervals of each trial, and listeners were instructed to select the interval in which the level of the target was chosen from an incremented distribution of levels. In a separate condition, listeners made similar judgments about the frequencies of the tones. Each tone in the sequence was 30 ms in duration, rectangularly gated, with no gaps between the tones. In this task, the optimal weighting strategy is one in which the listener gives no weight to the nine nontarget tones. Lutfi found that the specific form of the weighting function depended upon the relative jitter applied to the target and nontarget tones. The target tone was given the greatest weight when its jitter was larger than that of the nontarget tones, and all tones were weighted nearly equally when the jitter of the target tone was smaller than that of the nontarget tones. In most cases, many of the nontarget tones were given

nonzero weight. There was a tendency for the final tone in the sequence, a nontarget tone, to be given slightly greater weight than the other nontarget tones.

Another type of task for which the optimal weighting strategy would involve unequal weights over time is one in which the listener is required to make judgments as to the relative intensity of temporal segments of a signal, or, equivalently, a task requiring the listener to judge the temporal order of pulses with different levels (Hirsh, 1959; Babkoff and Sutton, 1963; Ronken, 1970; Green, 1973). In one such experiment, Green (1973) asked listeners to judge the temporal order of two 1-ms tonal pulses that differed in intensity by 10 dB. He measured percent correct as a function of the temporal separation of the pulses for frequencies of 1, 2, and 4 kHz. Green found that for all frequencies performance peaked at separations around 2–4 ms, decreased as the separation increased from 4 to 32 ms, then increased once again as the separation was made larger than 32 ms. Similar results were reported by Babkoff and Sutton (1963). Green concluded that the nonmonotonic change in performance with increasing temporal separation indicates that two distinct temporal processes are involved in the discrimination of the intensity of sequentially presented tone pulses.

Schlauch *et al.* (1999) performed a series of experiments designed to assess the ability of listeners to access level information in discrete temporal elements of an auditory stimulus. In each trial, a train of either three or four noise bursts was presented to listeners. The levels of successive bursts alternated between 60 and 80 dB SPL, with the level of the first burst chosen at random. Listeners were instructed to indicate whether the first burst was a low- or high-level burst, and performance was measured as a function of the presentation rate of the bursts. In a second experiment, listeners were instructed to count the number of bursts in a train of equal-level noise bursts that varied in their rate of presentation from trial to trial. No feedback was given in either experiment. Schlauch *et al.* reported that both experiments yielded similar results: listeners were able to perform both tasks for low stimulus presentation rates (large temporal spacings), but performance dropped as presentation rate increased (temporal spacing decreased). The authors estimated the upper limit of temporal resolution as measured by these tasks as about 20 Hz, corresponding to a temporal separation (silent period) of about 25 ms, which is in good agreement with the data on temporal order judgment discussed above.

Experiment 1 examines in more detail the issue of the perceptual weight given to discrete temporal elements of an auditory stimulus, more specifically, the weighting strategies adopted by listeners when a stimulus is composed of a number of temporally discrete pulses and listeners are asked to attend to the pulses in a number of different ways. The question to be addressed is: Do listeners alter their weighting strategies in an intensity discrimination task when the relevance of short-duration segments of the stimulus is varied?

## I. EXPERIMENT 1: PERCEPTUAL WEIGHTS

In one set of conditions of this experiment, listeners are asked to discriminate changes in the level of one temporal

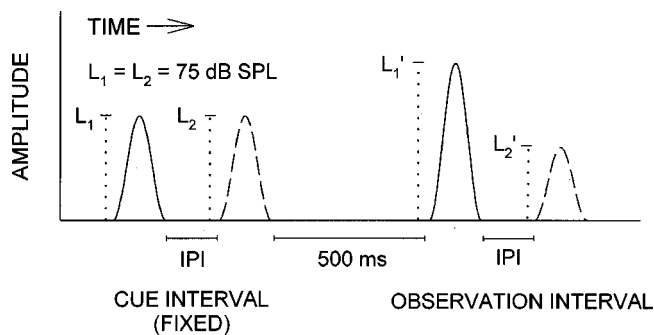


FIG. 1. A representation of a trial in which the levels of both the first and second pulses were varied. The first interval (the cue) consisted of two pulses with levels of 75 dB SPL. The second interval (the observation interval) consisted of two pulses with randomly and independently chosen levels. (All  $L$ s are in units of dB SPL.) The IPI was fixed for an entire block of trials. Because the levels of the pulses were selected randomly and independently from trial to trial in certain experimental conditions, the changes in level of the two pulses (relative to the cue) might be consistent with or opposite to one another on any given trial in those conditions.

segment of the stimulus (the target) while ignoring level changes in another temporal segment (the nontarget). In a case such as this, the optimal weighting strategy is one in which all perceptual weight is given to the target and a weight of zero is given to the nontarget. In another condition, listeners are asked to compare the level of the first pulse to the second pulse and to indicate which pulse has a higher level. In this case, the optimal weighting strategy is one in which weights of equal magnitude but opposite sign are given to the pulses. Weighting strategies are assessed as a function of the temporal separation between the pulses. Feedback is provided to listeners in each condition.

## A. Methods

All stimuli were composed of tone pulses (described below) and were presented over headphones to the listener's left ear. Each trial, an example of which is shown in Fig. 1, consisted of two intervals (resulting in a cued single-interval task). In the first interval (the cue), the first and second pulses were both presented at their mean level. The pulses in the cue were separated by the interpulse interval (IPI) which was determined by the specific condition that was being run. The second interval (observation interval) consisted of two pulses that were separated by the same IPI as the cue. The IPI for the cue and observation interval was the time between the offset of the first pulse and the onset of the second pulse. Data were collected for IPIs of 2, 8, 16, 32, 64, 128, and 256 ms.

The first and second pulses in the observation interval had levels that were randomly and independently chosen or that were fixed at the mean level on each trial, depending on the particular experimental condition. When the levels were randomly chosen, they were chosen from a Gaussian distribution with a mean of 75 dB SPL. Thus when the levels of both the first and second pulses were randomly chosen, the pulses might change level in the same or opposite directions relative to the cue on a trial-by-trial basis.

All pulses were 1-kHz pure tones, 10 ms in duration, with 5-ms raised cosine on–off ramps (0-ms steady state). The level of each pulse was defined to be the level in dB SPL

of a continuous tone with the same peak amplitude. The cue and observation interval of each trial were separated by 500 ms of silence. The IPI and instructions to the listener were fixed during each 50-trial block. Twenty blocks were presented to listeners for every IPI in each experimental condition so that 1000 responses were obtained for each. In all of the conditions described below, feedback that was consistent with the particular experimental condition was given to the listener after each trial.

Listeners were instructed to use the information in the first and second pulses in a particular way in each experimental condition. In one set of conditions, the first pulse was identified as the target. In this case, the listener was instructed to indicate whether the first pulse increased or decreased in level relative to that of the cue by pressing one of two keys on a computer keyboard. The listener was instructed to ignore the second pulse in both intervals when the first pulse was the target. In one condition, the level of the second pulse in the observation interval was fixed at its mean level (75 dB SPL), and in a separate condition the level of the second pulse was selected randomly from a Gaussian distribution with the same parameters as that of the first pulse. In this latter condition, the first and second levels were selected independently. For convenience, these separate conditions were called “target-1/fixed-2” and “target-1/variable-2.”

In another set of conditions, the second pulse was identified as the target and the listener was instructed to indicate whether it increased or decreased in level relative to the cue while ignoring the first pulse. As above, the nontarget (first pulse) level was fixed at its mean level in one condition (“target-2/fixed-1”), and in a separate condition (“target-2/variable-1”) the nontarget level in the observation interval was selected from a Gaussian distribution with the same parameters as that of the target.

In a final condition, the difference condition, the listeners were instructed to compare the two pulses in the observation interval to one another and to indicate whether the second pulse was higher or lower in level (“louder” or “quieter”) than the first pulse. In this condition, the levels of the two pulses in the observation interval were always selected randomly and independently. The fixed-level cue, consisting of two pulses at the same IPI as the observation interval, was presented in this condition even though it may not have been particularly useful to the listener. This was done so that the stimuli in this condition were statistically identical to those in the target-1/variable-2 and target-2/variable-1 conditions. The difference condition was run at the same IPIs as the previous conditions.

In all conditions, when a pulse’s level varied from trial to trial, the level was selected randomly from a Gaussian distribution with a mean of 75 dB SPL and a standard deviation (in dB) chosen to yield approximately 90% correct for the baseline intensity discrimination task for a single pulse in isolation, determined by trial and error. To achieve this level of performance, the standard deviation was set to 4 dB for listener S1 while it was set to 8 dB for the remaining four listeners. For conditions in which the levels of both the first and second pulse were variable, the levels were sampled in-

dependently from the Gaussian distribution. The levels were limited to a range of  $\pm 2.5$  standard deviations from the mean in order to avoid excessively high levels, producing overall ranges in level of 20 dB for listener S1 and 40 dB for the remaining four listeners. If the randomly chosen levels would produce a stimulus for which a correct response was undefined (0-dB change in level for the target, or equal levels for the two pulses in the difference condition), new levels were randomly chosen for the pulses.

As a baseline measure of performance, percent correct was also measured for a single pulse. In this case, the cue (a single pulse at the mean level) was presented in the first interval of each trial, followed by a single pulse with its level selected in the same manner as those in the variable level conditions described above. Listeners were instructed to indicate whether the second pulse was higher or lower in level than the cue.

All stimuli were generated digitally on a NeXT computer and converted to analog signals using the NeXT’s 16-bit D/A converters set to a rate of 44.1 kHz. The pulses were presented monaurally over Sony MDR-V6 headphones to listeners seated in an IAC sound-attenuating chamber. Listeners entered their responses on the computer keyboard and visual feedback was provided after each trial via the computer monitor.

Listeners were run during 2-h sessions during which approximately 15–20 50-trial blocks were completed. As stated previously, 1000 trials were run for each IPI in each condition. All trials of a particular condition were run before trials of another type were presented.

There were five listeners in the target-1 and target-2 conditions, and three of those listeners performed in the difference condition. One listener was the first author (S1). All remaining listeners were undergraduate volunteers who were paid to participate in the experiment. The first author was experienced in listening to these types of stimuli and required little training in order to become familiar with the task. The remaining four listeners had little or no experience in these types of listening tasks and were run for several thousand trials before data collection began. In addition, listeners were run in one or more blocks of practice trials in each condition before data were collected. All listeners ran the conditions in the following order: target-1/fixed-2, target-1/variable-2, target-2/fixed-1, target-2/variable-1. Within each condition, stimuli were presented from the longest IPI to the shortest IPI. The difference condition was run last, also from the longest to shortest IPI. Several randomly chosen conditions were repeated at the end of the experiment and no changes in performance or weighting strategies were observed, indicating that there were no order effects present.

## B. Results

### 1. Computation of observer weights

The data in the present experiments were analyzed using a correlational analysis (Richards and Zhu, 1994; Lutfi, 1995). In this model, the listener’s responses are based on a

decision variable that is the weighted sum of the changes in level of the individual components (in this case, the first and second pulses),

$$D = w_1(\Delta L_1) + w_2(\Delta L_2) + \varepsilon, \quad (1)$$

where the  $w$ 's are the observer's weights for the level changes between intervals (in dB) of the first and second pulses (as indicated by the subscripts), and  $\varepsilon$  is an additive error term encompassing all variability unaccounted for by the weighted levels. The value  $\Delta L$  for each pulse is the change in level of each pulse across intervals (e.g.,  $\Delta L_1 = L'_1 - L_1$  from Fig. 1), or, equivalently, the difference between each pulse and the mean pulse level. Responses are assumed to be based on the value of  $D$  such that listeners respond

$$\begin{aligned} &\text{‘quieter’ if } D < k; \text{ ‘louder’ if } D > k, \\ &\text{‘quieter’ or ‘louder’ randomly if } D = k, \end{aligned} \quad (2)$$

where  $k$  is the value of the criterion adopted by the listener. In the absence of response bias,  $k$  will be equal to zero, the mean change in level of each pulse across all trials. Equation (2) represents rescaling of the continuous internal decision variable  $D$  as a dichotomous response variable. The relative weight given to the level information of a given component [in Eq. (1)] is proportional to the correlation between the trial-by-trial levels for that component and the listener's binary responses.

In the results reported here, the *magnitudes* (absolute values) of the correlation coefficients (rather than the actual correlation coefficients) were normalized to sum to 1.00, and the *sign* of each original correlation coefficient was retained. This was done in order to deal with situations in which negative correlations were obtained. When negative correlations are obtained, normalizing the actual correlation coefficients produces weights beyond the range of  $-1.00$  to  $+1.00$ . Normalizing the magnitudes of the correlation coefficients more accurately reflects the relative weight given to each component across conditions. The sign of each weight then simply indicates whether the listener's responses were consistent with (+) or opposite to (−) the direction of the level change of that component.

## 2. Interpretation of observer weights

One can consider the weighting strategy used by an ideal observer that performs the task optimally, that is, a “noiseless” observer that maximizes percent correct. In these experimental tasks, every trial has an unambiguously correct response based on either the level change of the target or the difference between the levels of the two pulses. Thus an ideal noiseless observer that weights the information optimally for a given task will achieve 100% correct. When a human observer fails to achieve 100% correct, one can evaluate the extent to which imperfect performance results from (1) inappropriate weighting of the levels of the two pulses or (2) the influence of additional sources of variability other than the sum of the weighted levels.

The listener is said to use a nonoptimal weighting strategy when the relative weights for that listener differ from the

ideal weighting strategy for a given condition. Given the method of normalizing the observer weights used here, the weights will fall in the range from  $-1.00$  to  $+1.00$ . When the first or second pulse is identified as the target and the listener is instructed to ignore the nontarget pulse, the ideal weighting strategy (which maximizes percent correct) is one in which the observer gives the target pulse a weight of  $+1.00$  and the nontarget pulse a weight of  $0.00$ . Such a weighting strategy will produce 100% correct for the ideal noiseless observer. To the extent that weight is given to the nontarget pulse, performance will fall below 100% correct. When the listener is instructed to respond to the level of the second pulse relative to that of the first (the difference condition), the ideal weighting strategy is one in which the pulses are weighted equally but with opposite signs, that is, the weight for the second pulse is  $+0.50$  and the weight for the first pulse is  $-0.50$ . This weighting strategy will produce 100% correct for an ideal observer in the difference condition.

Even if the human observer uses the optimal weighting strategy, performance (percent correct) will be adversely affected by the presence of other sources of variability [ $\varepsilon$  in Eq. (1)] that are unrelated to the sum of the weighted levels, such as the discriminability of the level changes in the individual pulses or interference in the memory trace of one pulse produced by another pulse. However, the error term cannot be computed directly because the listener's responses are the result of generating dichotomous responses using Eq. (2) from values of the continuous variable  $D$  in Eq. (1).

Although the error term cannot be computed directly, its relative influence across experimental conditions can be estimated indirectly. An estimate of the relative influence of  $\varepsilon$  in Eq. (1) indicates how well the weighted combination of levels alone accounts for the listener's responses. One way in which the relative influence of the error term across conditions can be estimated is by computing the percentage of the listener's responses that can be predicted by the listener's weights.<sup>1</sup> In order to calculate this percentage, the computed weights were inserted into Eq. (1) and  $\varepsilon$  was set to zero. Simulated responses were obtained by inserting into Eq. (1) the actual levels that were presented to the listener on each trial and using Eq. (2) to generate a response.<sup>2</sup> The percentage of simulated responses that agreed with the listener's actual responses (the percentage of responses predicted by the weights) was then computed. This value is inversely related to the relative contribution of  $\varepsilon$  in the decision process. If  $\varepsilon = 0$ , then 100% of the responses can be predicted by the weights [when Eq. (2) is used to generate dichotomous responses from the continuous variable  $D$ ]. As  $\varepsilon$  becomes large, the percentage of responses predicted by the weights will drop to a lower limit of 50%. Furthermore, the particular weighting strategy that is adopted by the listener becomes irrelevant when the weights predict a very low percentage (near 50%) of the responses. In such cases, the weights have essentially no predictive value (responses are uncorrelated with the levels of either pulse). A comparison of the influence of  $\varepsilon$  across conditions indicates which conditions are influenced to a greater or lesser extent by factors other than the weighted level changes. In practice, it is unlikely that the

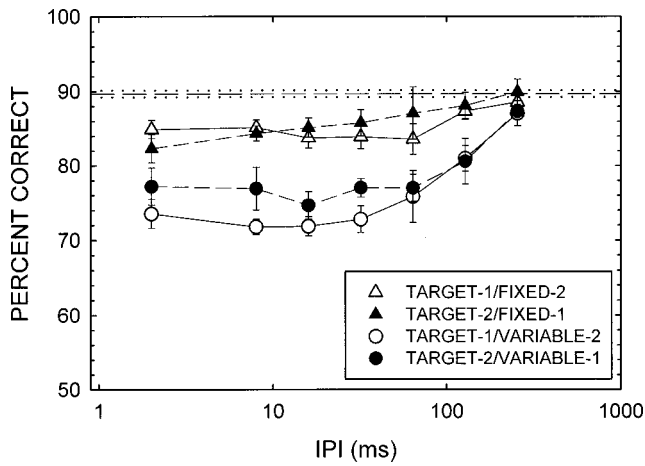


FIG. 2. Percent correct as a function of IPI in the two target-1 conditions and the two target-2 conditions. Percentages have been averaged across five listeners. Error bars represent standard errors of the mean. The dashed line at the top of the figure represents mean percent correct for the same five listeners in the single-pulse condition. The dotted lines above and below the dashed line represent  $\pm 1$  standard error of the mean.

weights will account for 100% of the responses in a given condition simply because the levels of the pulses are chosen randomly from trial to trial and there is some probability in each trial that the level of a particular pulse will be indistinguishable from its mean level. (See Stellmack *et al.*, 1999 for a more detailed discussion of the interpretation of the weights.)

### 3. Target-1 and target-2 conditions

Because the patterns of results for the five listeners were quite similar in the target-1 and target-2 conditions, the results were averaged across listeners. Figure 2 shows percent correct as a function of IPI in all of the target-1 conditions (open symbols) and target-2 conditions (filled symbols). Triangles represent the conditions in which the level of the nontarget pulse was fixed, and circles represent the conditions in which the level of the nontarget pulse varied. The error bars indicate standard errors of the mean. The dashed and dotted horizontal lines near the top of the panel indicate the average percent correct and  $\pm 1$  standard error of the mean, respectively, for these listeners in the baseline (single pulse) condition.

On average, percent correct was lower when two pulses were present relative to performance in the single-pulse condition except for the longest IPI (256 ms), where performance was essentially equal across conditions. In addition, there was a substantial decrement in performance when the level of the nontarget pulse varied from trial to trial (circles) relative to conditions in which the level of the nontarget pulse was fixed (triangles) for all but the longest IPI.

In the top panel of Fig. 3, percent correct scores from the target-1/variable-2 condition (open symbols) and target-2/variable-1 condition (filled symbols) have been replotted from Fig. 2. The middle panel of Fig. 3 shows the average normalized target weights as a function of IPI for the same conditions. (For the conditions in which the nontarget level is *fixed*, the relative weights are undefined.) Because the weights were normalized to sum to 1.00, the nontarget

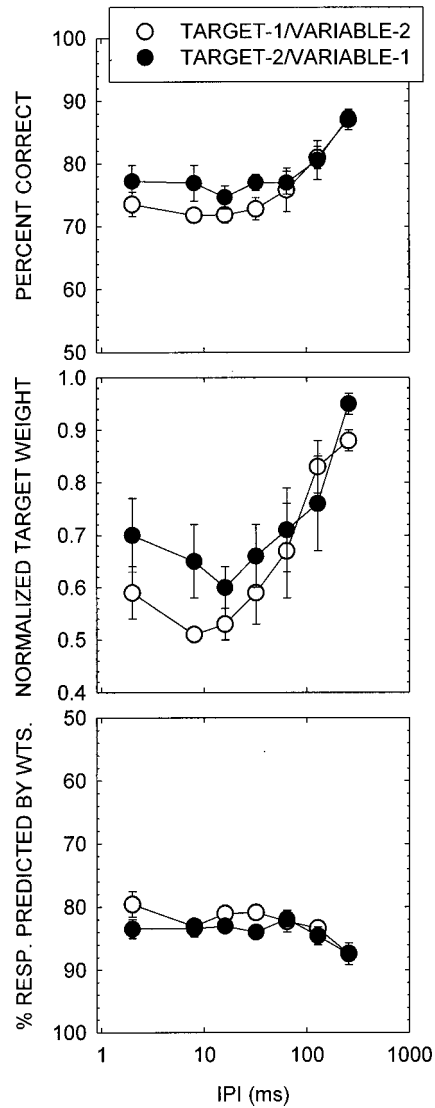


FIG. 3. In the top panel, percent correct is replotted from Fig. 2 for the target-1/variable-2 and target-2/variable-1 conditions. For the same conditions, the middle panel shows the target weight and the bottom panel shows the percentage of responses predicted by the weights, both as a function of IPI. All data are averaged across five listeners. Error bars represent standard errors of the mean.

weights are simply 1.00 minus the target weights that are plotted in the middle panel of Fig. 3. Although negative weights are possible, all average weights for both the target and nontarget pulses were positive in these conditions. The bottom panel of Fig. 3 shows the percentage of responses predicted by the weights for the same conditions. The ordinate in the bottom panel has been inverted to reflect the inverse relationship between the percentage of responses predicted by the weights and the relative influence of variance due to factors other than the weighted sum of the pulse levels. Thus higher relative amounts of unexplained variance are represented toward the upper portion of the panel.

The pattern of weights in the middle panel of Fig. 3 is similar to the pattern of performance shown in the top panel. High percent correct values in the top panel correspond to high target weights in the middle panel, with a similar correspondence for lower values. Percent correct was slightly

higher for the target-2 condition (filled symbols) relative to the target-1 condition (open symbols) for IPIs of 32 ms and less, and the same is true for the average target weights. In contrast, the percentage of responses predicted by the weights (in the bottom panel) remains relatively constant across IPI with a slight increase at the largest IPI. In other words, *relative* performance across these conditions is more closely associated with the relative weights given to the two pulses and is not influenced to a great extent by differences in the amount of additional unexplained variance across IPI.

Recall that the optimal weighting strategy that would produce the maximum percent correct is one in which the target is given a weight of 1.00. The target weight approaches but does not reach 1.00 at the largest IPI, indicating that changes in the level of the nontarget pulse influenced responses at all IPIs. As IPI increases from about 16 ms, the target weight becomes larger (closer to optimal). Target weights were at a minimum for IPIs around 8–16 ms. For these intermediate IPIs, the greatest nontarget weights were computed. The weight given to the target was slightly higher at the shortest IPI (2 ms).

#### 4. Difference condition

In the difference condition, two pulses with randomly and independently chosen levels were presented in the observation interval of each trial, and listeners were instructed to indicate whether the level of the second pulse was higher or lower than that of the first pulse. The results for the difference condition were less consistent across listeners than those for the previous conditions, so the data are presented individually in Fig. 4. Data for listeners S1 and S2 were qualitatively similar to one another, and are plotted in the left-hand column of Fig. 4 (in the same format as the data shown in Fig. 3). Data for listener S3 were somewhat different and are plotted separately in the right-hand column. In the middle panel of each column, the filled symbols are weights for the second pulse and the open symbols are weights for the first pulse. Different symbol types represent different listeners, as shown in the figure legend. Due to unavailability, S3 did not provide data for the smallest IPI (2 ms).

For S1 and S2 (left-hand column of Fig. 4), percent correct followed the same pattern in the difference condition as in the previous conditions (shown in Fig. 3). Best performance was obtained at the largest IPI (256 ms) and a minimum was observed at intermediate IPIs (32 ms for S1, 16 ms for S2). Recall that the optimal weighting strategy for the difference condition is one in which the weights are equal in magnitude but with opposite sign (+0.50 and -0.50). In contrast to the results for the target-1 and target-2 conditions, the weighting strategies for these two listeners (middle panel) remained fairly constant on average and nearly optimal across IPIs (although the second pulse was given slightly greater weight, as shown by the relative magnitudes of the weights), while the variation in the percentage of responses predicted by the weights (bottom panel) was more similar to the variation in percent correct. As the relative amount of unexplained variability increased, percent correct decreased. It appears that in the difference condition, performance

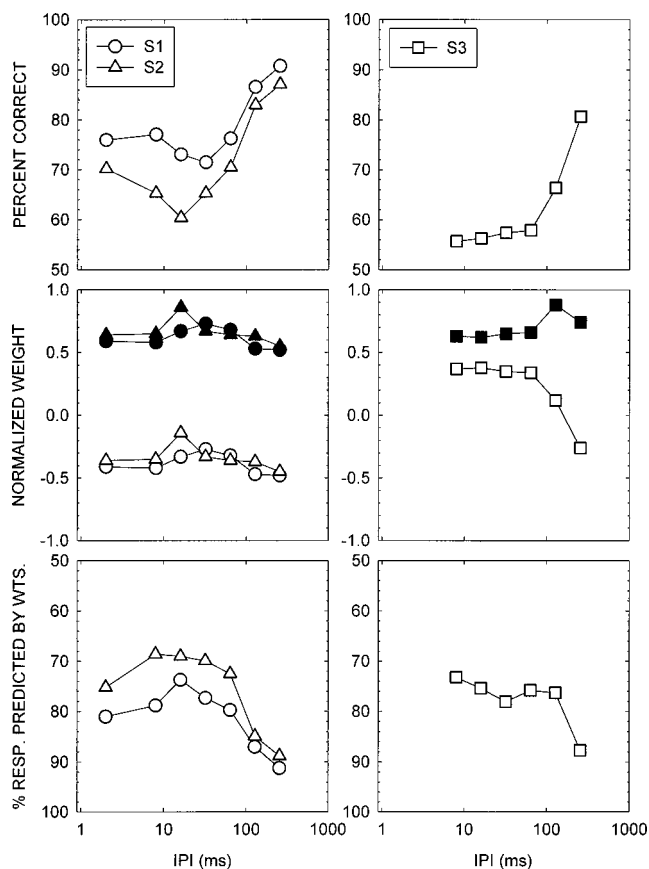


FIG. 4. The results of the difference condition, plotted in the same format as Fig. 3. Results are shown for listeners S1 and S2 in the left-hand column and for listener S3 in the right-hand column. In the middle panels of both columns, open and filled symbols represent weights for pulse 1 and pulse 2, respectively. Each symbol type represents results for a different listener, as indicated in the legend.

across IPI was less dependent upon the particular weighting strategy adopted by the listener and more influenced by additional, unexplained sources of variability than in the target-1 and target-2 conditions.

Listener S3 (right-hand column of Fig. 4) provided a slightly different pattern of results. Percent correct was at a maximum for the largest IPI, as for the other two listeners, and quickly dropped to near chance as IPI decreased. Weights of opposite sign were obtained only for the largest IPI, although slightly larger weight was given to the second pulse in all conditions, as was seen for the other two listeners. It appears that at the shorter IPIs, S3 had substantial difficulty with the difference task and adopted a strategy of merely responding to the change in overall level between intervals, with the second pulse weighted slightly more heavily.

#### C. Discussion

The results indicate that, over the range of IPIs tested here, listeners were unable to completely ignore an adjacent temporal segment of a stimulus that was qualitatively similar to the target segment when performing an intensity discrimination task. For 10-ms, 1-kHz pulses separated by intervals from 2 to 256 ms, the nontarget pulse influenced the responses of the listener. Listeners appeared to have the most

difficulty ignoring the nontarget pulse at intermediate IPIs (around 8–16 ms), with better performance at longer and shorter IPIs. The target weight was slightly greater when the second pulse was the target (compared to when the first pulse was the target).

The difference task is similar to the tasks requiring temporal order judgments that were described in the Introduction, although here the levels of the pulses were randomized from trial to trial. For listeners S1 and S2, there is some indication that percent correct is lowest for IPIs around 16–32 ms, consistent with the results of Babkoff and Sutton (1963) and Green (1973), but the nonmonotonicity in performance across IPIs is not nearly as pronounced as in the previous studies. Green (1973) observed that the nonmonotonicity was most evident in the data of unpracticed subjects. The large number of trials run in the present study may have served to reduce this effect.

Note that the stimuli in the difference condition were statistically identical to the stimuli in the target-1 and target-2 conditions with variable nontargets. Listeners S1 and S2 (and S3 to some extent) altered their weighting strategies depending upon the specific task at hand. Slightly greater positive weight was given to the target pulse in the target-1 and target-2 conditions, with relatively small positive weight given to the nontarget (Fig. 3). In the difference condition, weights of more similar magnitude but opposite sign were given to the two pulses (Fig. 4). Although the weights for only two of the three listeners varied in a manner consistent with the task instructions across all IPIs, this provides support to the notion that the levels of the individual pulses were treated as individual “looks” by the auditory system. If the level information were completely integrated across pulses, the weights would have been approximately equal and with the same sign across all conditions.

In the target/variable conditions, relative performance across IPIs was primarily determined by the weighting strategy adopted by the listeners. Decreases in percent correct were associated with increases in the weight given to the irrelevant nontarget pulse. In contrast, for two of the three listeners in the difference condition, relative performance was not strongly related to the listeners’ weighting strategy. The weighting strategy was nearly fixed and close to optimal across IPI. Decreases in performance were mirrored by increases in additional, unexplained variability in responses. (For the third listener in the difference condition, performance declined as both the weighting strategy became non-optimal and the variability increased.) Thus it appears that listeners adopt a weighting strategy that is closer to optimal when they are required to make decisions based on the entire stimulus rather than on certain temporal segments of the stimulus.

Green (1973) suggested that performance in the discrimination of temporal order exhibits a nonmonotonicity because two different mechanisms are at work at short and long IPIs, with neither mechanism extremely effective at intermediate IPIs. At very short IPIs, listeners perceive a slight qualitative difference between stimuli consisting of a soft–loud tone sequence versus a loud–soft sequence. Apparently, with brief IPIs listeners perceive the two pulses as a single

auditory event, with the change in loudness across the pulses producing a qualitative difference depending on the direction of the change. (Note that there are no long-term spectral differences between the time-reversed soft–loud and loud–soft tone sequences used by Green when one considers the spectrum of the entire tone sequence.) At intermediate IPIs, these qualitative differences become more difficult to hear as the pulses become more distinct. At long IPIs, the listener can clearly perceive two distinct pulses and the ordering of their intensities can be easily followed.

The idea that intensity information in a sequence of tone pulses is processed as a number of short-term looks does not imply that the listener necessarily perceives the pulses as individual auditory events. The fact that weights of opposite sign and nearly equal magnitude were obtained for two listeners at short IPI’s indicates that the levels of the two pulses effectively were processed independently. The effect is similar to that in situations involving the detection of amplitude modulation at high modulation frequencies. Although “roughness” might be perceived rather than discrete fluctuations in intensity (Wright and Dai, 1999), the fact that the fluctuations can be detected at all indicates that at some level of processing the fluctuations are resolved by the auditory system. When the task requires listeners to selectively process a particular peak in the fluctuating waveform, as in the present target/variable task or Schlauch *et al.* (1999), the inability to perceptually segregate the fluctuations limits their ability to do so.

One reason that listeners may have tended to perceptually fuse the stimuli in the present experiment at short IPIs is that the temporal segments of the signal were qualitatively similar (having the same frequency and duration). It is possible that listeners would be more able to selectively process the temporal segments of a stimulus if they were more qualitatively dissimilar, for example, if they were tones of different frequencies or durations. The implication is that interactions in the processing of intensity information in the individual pulses occur at some more central level of processing, after the intensities of the individual looks are extracted at more peripheral levels. This is, in effect, the basic premise of the multiple-looks hypothesis.

The results also may have been influenced by the type of cue that was used, specifically, two pulses fixed in level at the mean level of the variable pulses. Perhaps the two-pulse cue influenced the listeners to perceptually combine the two pulses in both intervals even at the larger IPIs. Perhaps a one-pulse cue would encourage the listener to hear out the target pulse more readily, or perhaps presenting only the observation interval on each trial with no cue would produce such results.

To examine these possibilities, some of the conditions of experiment 1 were repeated with no cue and with a cue consisting of only one pulse at the mean pulse level. The target-1/variable-2 conditions of experiment 1 were repeated with these different cues for IPI= 2, 32, and 256 ms. The results are shown in Fig. 5 for listeners S1 and S4 in the same format as Fig. 3. The observation interval of each trial was identical to those of experiment 1, while the cue consisted of either one pulse (filled squares) or two pulses (“time 2,”



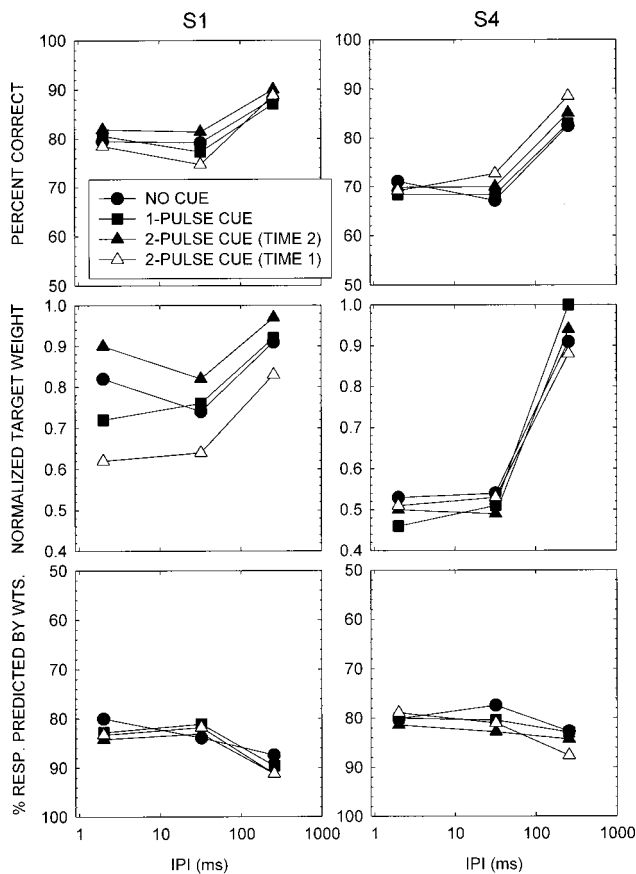


FIG. 5. Data from two listeners for selected conditions of experiment 1, repeated with different types of cues, as indicated in the legend. Results are plotted in the same format as Fig. 3. Data for the two-pulse cue condition at time 1 were gathered for experiment 1; all remaining data were collected at a later date.

filled triangles), or the observation interval was presented with no cue (filled circles). The two-pulse-cue condition was rerun because these data were gathered several months after the original data. The original two-pulse-cue data for these listeners from experiment 1 ("time 1," open triangles) are replotted for comparison.

It appears that the specific type of cue, including the complete lack of a cue, has little effect on the results. For S4, the results for no cue and a one-pulse cue are essentially the same as those for the two-pulse cue at both times 1 and 2. S1 produced the highest target weights for the two-pulse-cue condition at time 2, and identical, slightly lower target weights for the no-cue and one-pulse-cue conditions at IPI = 32 and 256 ms. All of the target weights for S1 measured at time 2 were slightly larger than those measured at time 1, representing a practice effect similar to that documented by Leek and Watson (1984). This type of practice effect was not seen for S1 when some of the conditions were repeated at the time of the initial data collection. In spite of the larger target weights over time for S1, both listeners still show the same general result as the original data: the largest target weights were obtained for IPI = 256 ms with substantially lower target weights at IPI = 2 and 32 ms.

A result that is not predicted by the multiple-looks hypothesis is the observation that a second nontarget pulse affected intensity discrimination performance even when the

level of the nontarget was fixed (Fig. 2, triangles). In this condition, listeners were required simply to indicate the direction of a level change occurring during the observation interval without the possible additional burden of ignoring a potentially conflicting level change, or of attributing two level changes to their respective temporal positions. The decrease in overall performance in the target/fixed conditions suggests that just-noticeable differences (JNDs) for intensity may be poorer (larger) for a tonal pulse in the presence of a second, fixed tonal pulse of the same frequency and duration. Such an effect was observed by Plack *et al.* (1995). An increase in intensity JNDs in the presence of a fixed nontarget would be predicted if intensity information were simply integrated across the two pulses, but the patterns of weights seen in the various conditions described above, in particular the fact that different patterns were obtained depending on the specific task, suggests that level information in the two pulses was processed independently, at least to some extent. To explore this issue further, an additional experiment was performed to examine the effect on intensity JNDs of a fixed nontarget pulse.

## II. EXPERIMENT 2: INTENSITY JNDs

In this experiment, intensity JNDs are measured for a tonal pulse in the presence of a second, temporally discrete tonal pulse with a fixed level. Both pulses have the same frequency and duration. Intensity JNDs are measured as a function of the time between the two pulses (interpulse interval, IPI), ranging from 2 to 256 ms. JNDs are measured in conditions in which the level of the fixed pulse is -10, 0, or 10 dB relative to the pedestal level of the target pulse (in separate conditions). Intensity JNDs are also measured for a single pulse in isolation. If intensity simply is integrated across the pulses, then, relative to the single-pulse condition, intensity JNDs must increase by 0.41, 3.0, and 10.4 dB [ $10 \log(\Delta I/I)$ ] for relative nontarget levels of -10, 0, and 10 dB, respectively, in order to maintain a particular value of the Weber fraction.

### A. Methods

A two-interval, forced-choice procedure was used in which one interval of each trial contained the target pulse at its pedestal level and the other interval contained the target pulse at its pedestal level plus an increment. The listener's task was to select the interval containing the incremented target pulse. In conditions involving a nontarget pulse, it was presented in both intervals at a fixed level. In separate conditions, the level of the nontarget pulse was fixed at -10, 0, or 10 dB relative to the pedestal level of the target (i.e., 65, 75, or 85 dB SPL). JNDs were measured for the first pulse as target and, separately, for the second pulse as target. Intensity JNDs were also measured for a single pulse in isolation. The pedestal level of the target was 75 dB SPL in all conditions.

JNDs were measured adaptively using a three-down one-up tracking rule that estimated the 79.4% correct point on the psychometric function (Levitt, 1971). In each adaptive run, the step size of the increment was 4 dB, in units of

$10 \log(\Delta I/I)$ , until four reversals were obtained, after which the step size was changed to 2 dB until a total of 12 reversals were obtained. The average of the increments at which the last eight reversals occurred was calculated as the JND estimate for each adaptive run. Four such estimates were averaged to obtain a final estimate of the listener's intensity JND in each condition.

Both the target and nontarget pulses had the same parameters as those used in the previous experiment: 1-kHz tones, 10 ms in duration, with 5-ms on-off ramps (no steady state). Within each interval, the target and nontarget pulses were separated by the IPI, which ranged from 2–256 ms and which was defined as the time between the offset of the first pulse and the onset of the second pulse. The two intervals of each trial were separated by 500 ms of silence.

Four listeners from the previous experiment (all except S3) participated in this experiment. All conditions of the experiment were run in a pseudorandom order chosen by the experimenter.

## B. Results and discussion

The difference between the intensity JNDs for each two-pulse condition and the one-pulse condition was computed for each listener in units of  $10 \log(\Delta I/I)$ . The form of these differences was similar for all listeners and thus were averaged across listeners. Figure 6 shows these average differences, with the top and bottom panels representing data for conditions in which the first or second pulse, respectively, was the target. The symbol type indicates the level of the nontarget pulse with respect to the level of the target pedestal. The points plotted separately on the left-hand side of the panels indicate the threshold differences that would be predicted if perfect integration across the two pulses occurred.

Despite substantial variability in the actual size of JND differences between subjects, it can be seen that JNDs increased with increasing nontarget level roughly independent of IPI. In most cases, particularly for the  $-10$  and  $+10$  dB nontargets, the actual JNDs could not be predicted by assuming perfect temporal integration of the pulses. JNDs were higher than predicted for the  $-10$  dB nontargets and they were lower than predicted for the  $+10$  dB nontargets. Although JNDs depended upon the level of the fixed nontarget component, the fact that they did not change with IPI over the range studied here indicates that changes in weighting strategy across IPI that were seen in experiment 1 cannot be accounted for by differential sensitivity to intensity changes at different IPIs.

Plack *et al.* (1995) presented results for a similar task (their experiment 1), in which intensity JNDs were measured for a 30-ms tone in the presence of a second 30-ms masker tone that either preceded or followed the signal for IPIs ranging from 12.5 to 200 ms. They also found that thresholds were higher in all cases for the two-pulse condition relative to a one-pulse condition, but thresholds decreased slightly with decreasing IPI. No such decrease was found in the present results. In the Plack *et al.* experiment, the level of the nontarget pulse was 80 dB SPL and the level of the target pedestal was 50 dB SPL. It is possible that as the IPI was

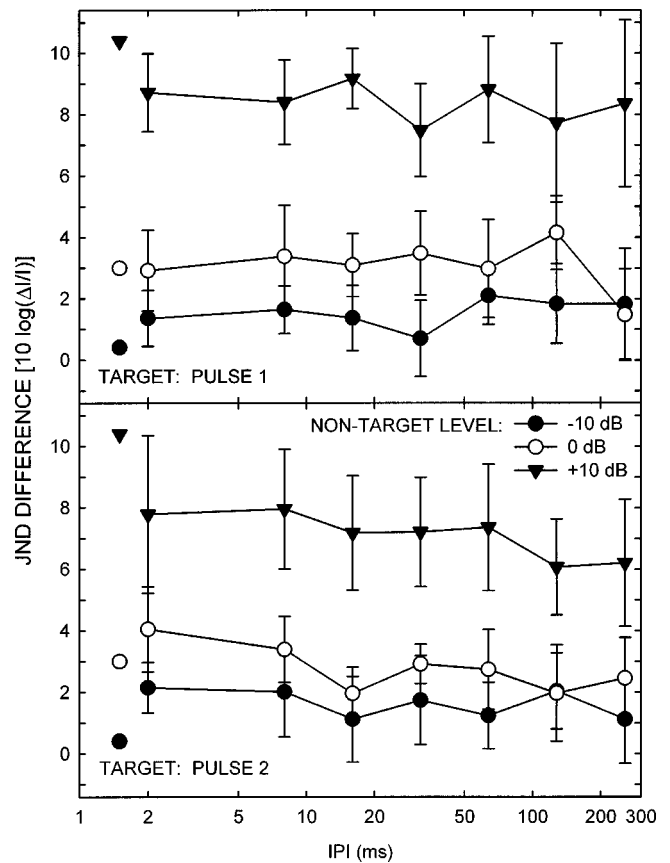


FIG. 6. The differences between intensity discrimination thresholds in the two-pulse conditions and the threshold for one pulse, plotted as a function of IPI. Different symbol types indicate the level of the nontarget pulse with respect to the level of the target pedestal, which was always 75 dB SPL. The upper panel shows data for conditions in which the first pulse was the target, and the lower panel shows data for the second pulse as the target. Each separate symbol at the left of the panels indicates the threshold increase that would be predicted for each nontarget level assuming perfect temporal integration of the pulses. The data are averaged across four listeners. Error bars represent standard errors of the mean.

reduced in their experiment, the pedestal and increment were sufficiently masked by the nontarget pulse that the intensity discrimination task became more like a detection task. Consistent with this explanation, in subsequent experiments, Plack *et al.* found that intensity discrimination thresholds increased as the level of a second pulse that was 12.5 ms from the target was reduced. It should be noted, however, that the stimuli used by Plack *et al.* were more complicated than those in the present experiment, with a third pulse more temporally remote from the target.

Plack *et al.* accounted for their results in terms of a referential coding hypothesis, whereby the listener can make use of the second fixed pulse as an intensity reference and can encode the intensity of the target pulse in terms of its intensity relative to the reference rather than in terms of its absolute intensity. Indeed, in the present experiment, when the nontarget pulse was equal in level to the target pedestal, on each trial the listener was presented with two pairs of pulses, or three pulses of equal level and one pulse with an incremented level. It would seem that in the equal level (0-dB) condition, the listener could have used the fixed pulses as additional references, essentially getting additional

“samples” of the pedestal level. However, even at the longest IPI (256 ms), thresholds in the two-pulse condition were elevated relative to the one-pulse condition, which itself contains two pulses separated by 500 ms on each trial. This suggests that intensity discrimination thresholds in a one-pulse condition, in which one pulse is presented in each of two intervals, would increase as the time between the intervals is decreased from 500 to 256 ms.

A number of studies have addressed the topic of interference between two intervals in an intensity discrimination task as a function of the time between the intervals (Tanner, 1961; Sorkin, 1966; Taylor and Smith, 1975). In general, it has been found that performance in a two-interval, two-alternative forced-choice task is maximal for an interstimulus interval (ISI) of about 500 ms. At shorter ISIs, it is assumed that the close temporal proximity of the intervals creates interference in short-term memory, limiting processing of either interval. For ISIs longer than about 500 ms, the memory trace of the first interval presumably decays or becomes increasingly corrupted by noise (Kinchla and Smyzer, 1967) which limits the accuracy of comparisons between the intervals. None of these hypotheses nor the present results imply simple temporal integration across stimuli for temporal separations greater than a few milliseconds. Rather, it is suggested that interference occurs between the information provided in the discrete temporal intervals at a central level of processing.

### III. GENERAL DISCUSSION

#### A. Multiple-looks hypothesis

Experiment 1 measured the weighting functions that describe the manner in which listeners use short-term looks at the level of a stimulus in making decisions about changes in level at different temporal positions. In general, different weighting functions were computed when listeners were instructed to make different decisions about the stimuli even though the stimuli were statistically identical across trials. This indicates that listeners do in fact have access to level information in brief segments of a stimulus, and that listeners can make optimal use of that information to different degrees depending upon the temporal separation of the stimulus segments. The results of the difference condition and those of Lutfi (1990) and Buus (1999) suggest that it is easier for listeners to adopt an optimal weighting strategy when the task requires listeners to utilize the information in all of a group of short-term looks that are closely spaced in time. Listeners appear to have more difficulty selectively ignoring some of those looks, as in the target/variable conditions of experiment 1 and in Lutfi (1992). This may be true only for situations in which the looks are qualitatively similar, given the results of Viemeister and Wakefield (1991), which showed that changing the level of a noise burst temporally positioned between two tone bursts has no effect on detection thresholds for the two tone bursts.

As implied above, given that different weighting strategies are observed for different tasks, and that there is some variability in weighting strategies across listeners, it is possible that weighting strategies also would be affected by the

type of signal. For example, the present experiments and those of Buus (1999) and Viemeister and Wakefield (1991) utilized signals consisting of temporally discrete tonal pulses. Although listeners in most cases used a nearly optimal weighting strategy when all of the discrete pulses were to be used in performing a listening task (e.g., detection of the pulses or comparison of the pulse levels), nonoptimal strategies were observed when listeners were instructed to base decisions on one pulse while ignoring the other (as in the target/variable conditions). Although detection of a signal consisting of a series of tone pulses yielded optimal weighting strategies, it is possible that a nonoptimal weighting strategy would be used when the task requires the listener to detect a stimulus consisting of a single long-duration tone. This was observed by Gilkey and Robinson (1986) and posited by Viemeister and Wakefield (1991) to account for previous temporal integration data. In addition, the weighting strategy is influenced by the relative statistics of the parameters of the stimulus elements (Berg, 1990; Lutfi, 1992) and by the characteristics of off-frequency maskers (Buus, 1999). It is clear that no single weighting strategy can account for performance with all stimuli and listening tasks, but that the weighting of short-duration segments of a stimulus can account for a substantial portion of responses in many specific tasks.

#### B. Comparison to binaural data

In an experiment that was a binaural analog to experiment 1 of the present study, Stellmack *et al.* (1999) measured weighting functions for two temporally separated clicks with different interaural delays. In that study as in the present paper, listeners were instructed to respond to the information in only one stimulus element or to compare the stimulus elements to one another. Although the experiments were performed with mostly different listeners and slightly different stimuli, a comparison of the data from the two experiments suggests some possible similarities between the weighting functions measured in the two tasks. For example, in the binaural task in which the interaural delays of both clicks varied from trial to trial, percent correct on average was lowest for an intermediate range of interclick intervals (8–32 ms). In addition, as in the present study, there was evidence that performance was limited by two different factors at short and long interclick intervals: the precedence effect at short interclick intervals (see Litovsky *et al.*, 1999, for a recent review), and an inability to discern the temporal order of the interaural delays at longer interclick intervals, a limitation found in a number of sensory modalities by Hirsh and Sherrick (1961). Additional data would need to be gathered in both the monaural and binaural conditions from the same listeners in order to establish whether weighting functions at the larger range of IPIs might be driven by a common limitation on the ability to judge temporal order.

#### C. Conclusions

The weighting functions obtained in experiment 1 suggest that the auditory system treats long-duration stimuli as if they are composed of many short-term looks, with the infor-

mation in each look weighted in a particular way depending upon the task at hand. It appears that when information-bearing, qualitatively similar stimulus elements are separated by less than about 16 ms, listeners adopt a weighting strategy that is closer to optimal when required to make decisions based on a combination of all of the looks rather than on a subset of the looks. Listeners display an increasing ability to ignore particular temporal segments of the stimulus with increasing temporal spacing above about 16 ms.

Intensity JNDs for one pulse are elevated when a second fixed-level pulse is present, but cannot be predicted by assuming perfect temporal integration across the pulses. The amount of JND elevation is dependent upon the level of the fixed-level pulse but not on the temporal spacing of the pulses over the range from 2 to 256 ms. This indicates that the various weighting strategies measured in experiment 1, which differed across temporal separations, are not attributable to decreased sensitivity to level changes in one pulse when a second pulse is present.

## ACKNOWLEDGMENTS

The authors would like to thank the associate editor (Dr. Sid P. Bacon) and two anonymous reviewers for their careful reviews and helpful suggestions. The authors also thank Dr. Martin E. Rickert and Dr. Magda Wojtczak for fruitful discussions of the material herein. This work was supported by Research Grant No. R01 DC 00683 from the National Institute on Deafness and Communication Disorders, National Institutes of Health.

<sup>1</sup>Using the equations of Richards and Zhu (1994), one can estimate the relative magnitudes (in proportion to the relative weights) of the variance of the continuous decision variable ( $\sigma_D^2$ ) and the variance of the additive internal noise process ( $\sigma_e^2$ ). From these values, one can estimate the proportion of variance in the listener's responses attributable to the additive noise process ( $\sigma_e^2/\sigma_D^2$ ), and the proportion of variance attributable to the sum of the weighted level changes ( $1 - \sigma_e^2/\sigma_D^2$ ). A graph describing the latter quantity as a function of IPI is extremely close in form to that describing the proportion of responses predicted by the weights as a function of IPI, which is used in the present paper. Because relative comparisons are made across conditions in the present manuscript, either measure would provide the same information. The proportion of responses predicted by the weights was also chosen for its intuitive appeal.

<sup>2</sup>The criterion,  $k$ , in Eq. (2) was estimated as described by Richards and Zhu (1994). In general, the estimated criteria were very close to zero (the unbiased value of the criterion). When the estimated criterion was used to predict the listener's responses, very few additional responses were predicted compared to when an unbiased criterion ( $k = 0$ ) was assumed. In the most extreme case, 33 additional responses out of 1000 were correctly predicted when the estimated criterion was used. In most cases, fewer than ten additional responses were predicted using the estimated criterion, indicating that listener's responses for all practical purposes were unbiased.

Babkoff, H., and Sutton, S. (1963). "Perception of temporal order and loudness judgments for dichotic clicks," *J. Acoust. Soc. Am.* **35**, 574–577.  
 Berg, B. G. (1989). "Analysis of weights in multiple observation tasks," *J. Acoust. Soc. Am.* **86**, 1743–1746.  
 Berg, B. G. (1990). "Observer efficiency and weights in a multiple observation task," *J. Acoust. Soc. Am.* **88**, 149–158.  
 Buus, S. (1999). "Temporal integration and multiple looks, revisited:

Weights as a function of time," *J. Acoust. Soc. Am.* **105**, 2466–2475.  
 Buus, S., Zhang, L., and Florentine, M. (1996). "Stimulus-driven, time-varying weights for comodulation masking release," *J. Acoust. Soc. Am.* **99**, 2288–2297.  
 de Boer, E. (1985). "Auditory time constants: A paradox?," in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, Berlin), pp. 141–158.  
 Forrest, T. G., and Green, D. M. (1987). "Detection of partially filled gaps in noise and the temporal modulation transfer function," *J. Acoust. Soc. Am.* **82**, 1933–1943.  
 Gilkey, R. H., and Robinson, D. E. (1986). "Models of auditory masking: A molecular psychophysical approach," *J. Acoust. Soc. Am.* **79**, 1499–1510.  
 Green, D. M. (1960). "Auditory detection of a noise signal," *J. Acoust. Soc. Am.* **32**, 121–131.  
 Green, D. M. (1973). "Temporal acuity as a function of frequency," *J. Acoust. Soc. Am.* **54**, 373–379.  
 Green, D. M. (1985). "Temporal factors in psychoacoustics," in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, Berlin), pp. 122–140.  
 Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).  
 Hirsh, I. J. (1959). "Auditory perception of temporal order," *J. Acoust. Soc. Am.* **31**, 759–767.  
 Hirsh, I. J., and Sherrick, Jr., C. E. (1961). "Perceived order in different sense modalities," *J. Exp. Psychol.* **62**, 423–432.  
 Kinchla, R. A., and Smyzer, F. (1967). "A diffusion model of perceptual memory," *Percept. Psychophys.* **2**, 219–229.  
 Leek, M. R., and Watson, C. S. (1984). "Learning to detect auditory pattern components," *J. Acoust. Soc. Am.* **76**, 1037–1044.  
 Levitt, H. (1971). "Transformed up-down methods in psychophysics," *J. Acoust. Soc. Am.* **49**, 467–477.  
 Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999). "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633–1654.  
 Lutfi, R. A. (1990). "Informational processing of complex sound. II. Cross-dimensional analysis," *J. Acoust. Soc. Am.* **87**, 2141–2148.  
 Lutfi, R. A. (1992). "Informational processing of complex sound. III: Interference," *J. Acoust. Soc. Am.* **91**, 3391–3401.  
 Lutfi, R. A. (1995). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," *J. Acoust. Soc. Am.* **97**, 1333–1334.  
 Plack, C. J., Carlyon, R. P., and Viemeister, N. F. (1995). "Intensity discrimination under forward and backward masking: Role of referential coding," *J. Acoust. Soc. Am.* **97**, 1141–1149.  
 Richards, V. M., and Zhu, S. (1994). "Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients," *J. Acoust. Soc. Am.* **95**, 423–434.  
 Ronken, D. A. (1970). "Monaural detection of a phase difference between clicks," *J. Acoust. Soc. Am.* **47**, 1091–1099.  
 Schlauch, R. S., DiGiovanni, J. J., and Ries, D. R. (1999). "Cognitive limitations on auditory temporal resolution," Association for Research in Otolaryngology Mid-Winter Meeting, February 1999.  
 Sorkin, R. D. (1966). "Temporal interference effects in auditory amplitude discrimination," *Percept. Psychophys.* **1**, 55–58.  
 Stellmack, M. A., Dye, Jr., R. H., and Guzman, S. J. (1999). "Observer weighting of interaural delays in source and echo clicks," *J. Acoust. Soc. Am.* **105**, 377–387.  
 Tanner, Jr., W. P. (1961). "Physiological implications of psychophysical data," *Ann. (N.Y.) Acad. Sci.* **89**, 752–765.  
 Taylor, M. M., and Smith, S. M. (1975). "Monaural detection with contralateral cue. V. Interstimulus interval in MDCC and amplitude discrimination," *J. Acoust. Soc. Am.* **57**, 1500–1511.  
 Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364–1380.  
 Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal integration and multiple looks," *J. Acoust. Soc. Am.* **90**, 858–865.  
 Wright, B. A., and Dai, H. (1999). "Detection of sinusoidal amplitude modulation at unexpected rates," *J. Acoust. Soc. Am.* **104**, 2991–2996.

# Integration of monaural and binaural evidence of vowel formants

Michael A. Akeroyd<sup>a)</sup> and A. Quentin Summerfield

*MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, United Kingdom*

(Received 28 May 1999; revised 16 November 1999; accepted 4 March 2000)

The intelligibility of speech is sustained at lower signal-to-noise ratios when the speech has a different interaural configuration from the noise. This paper argues that the advantage arises in part because listeners combine evidence of the spectrum of speech in the across-frequency profile of interaural decorrelation with evidence in the across-frequency profile of intensity. To support the argument, three experiments examined the ability of listeners to integrate and segregate evidence of vowel formants in these two profiles. In experiment 1, listeners achieved accurate identification of the members of a small set of vowels whose first formant was defined by a peak in one profile and whose second formant was defined by a peak in the other profile. This result demonstrates that integration is possible. Experiment 2 demonstrated that integration is not mandatory, insofar as listeners could report the identity of a vowel defined entirely in one profile despite the presence of a competing vowel in the other profile. The presence of the competing vowel reduced accuracy of identification, however, showing that segregation was incomplete. Experiment 3 demonstrated that segregation of the binaural vowel, in particular, can be increased by the introduction of an onset asynchrony between the competing vowels. The results of experiments 2 and 3 show that the intrinsic cues for segregation of the profiles are relatively weak. Overall, the results are compatible with the argument that listeners can integrate evidence of spectral peaks from the two profiles.

© 2000 Acoustical Society of America. [S0001-4966(00)02206-2]

PACS numbers: 43.66.Pn, 43.66.Ba, 43.71.Es, 43.71.An [DWG]

## INTRODUCTION

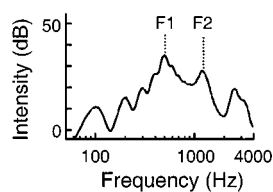
The intelligibility of speech is sustained at lower signal-to-noise ratios (SNRs) when the speech has a different interaural configuration from the noise than when both have the same configuration (e.g., Licklider, 1948; Levitt and Rabiner, 1967). The advantage is usually referred to as the binaural intelligibility level difference (BILD). It ranges in size up to about 8 dB, depending on the speech stimuli and the configurations that are compared (e.g., Blauert, 1983). The BILD arises because the difference in interaural configuration allows listeners to extract binaural cues to the spectrotemporal structure of the speech signal, and these cues supplement cues obtained from monaural analysis. At high SNRs the auditory representation of the across-frequency profile of intensity—the monaural excitation pattern—contains evidence of the spectral structure of the speech (e.g., Glasberg and Moore, 1990). In this paper it is argued that, as the SNR is reduced, the representation of the across-frequency profile of interaural decorrelation—conveniently termed the “binaural excitation pattern”—contains this evidence (Culling and Summerfield, 1995). Thus, the intelligibility of speech would be sustained across a range of SNRs if listeners could combine evidence from the two excitation patterns. The experiments reported below explored the extent to which such combination is possible.

The potential benefits of combining evidence in this way

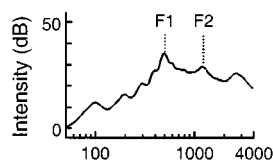
are illustrated in Fig. 1. The stimulus was a voiced vowel synthesized with a fundamental frequency of 100 Hz (Klatt, 1980), with the first formant (F1) at 508 Hz and the second formant (F2) at 1240 Hz.<sup>1</sup> It is similar to the British-English monophthongal vowel in the word “heard.” It was defined as a two-channel stimulus in which the waveform in the right channel started 400  $\mu$ s before the waveform in the left channel. It was analyzed in quiet (top row) or after the addition of diotic white noise (lower three rows). The SNR was progressively reduced to  $-20$  dB by decreasing the level of the signal. Minus twenty dB is about the lowest SNR at which the members of small sets of synthetic vowels are identifiably distinct from each other when presented with an interaural time difference (ITD) of 400  $\mu$ s in diotic noise (Culling *et al.*, 1994). The left column of Fig. 1 contains monaural excitation patterns calculated from the right channel of the stimulus; the right column contains binaural excitation patterns calculated from both channels. The procedures used to calculate the excitation patterns are described below in Sec. A. In the monaural excitation patterns, peaks corresponding to the first two formants of the vowel are visible at an SNR of 0 dB. As the SNR is reduced the peaks become less distinct, and are not evident at  $-20$  dB. In comparison, the peaks become more evident in the binaural excitation patterns as the SNR is reduced; they are particularly clear at  $-20$  dB. Consequently, the formant frequencies would be best determined from the monaural excitation pattern at high SNRs and from the binaural excitation pattern at lower SNRs. At intermediate SNRs, the vowel would be identified most accurately if the formant information from the two excitation patterns was combined.

<sup>a)</sup>Present address: Surgical Research Center, Department of Surgery (Otolaryngology) and Center for Neurological Sciences, University of Connecticut Health Center, Farmington, CT 06030. Electronic mail: michael@neuron.uchc.edu

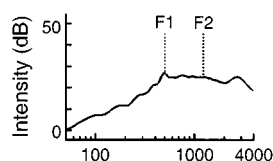
## Monaural excitation patterns



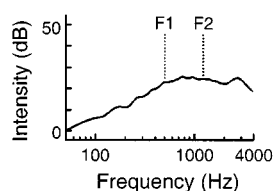
Vowel in quiet



SNR = 0 dB



SNR = -10 dB



SNR = -20 dB

## Binaural excitation patterns

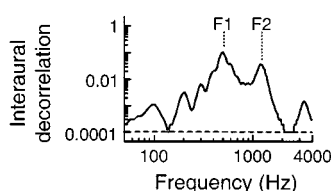
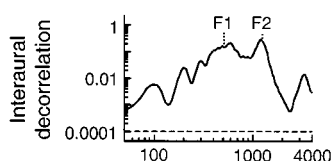
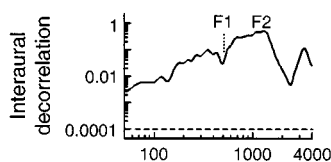
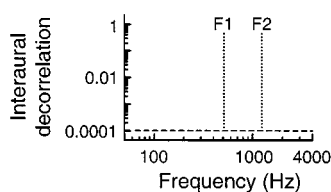


FIG. 1. Monaural and binaural excitation patterns for a synthetic exemplar of a voiced vowel with an ITD of  $+400 \mu\text{s}$ . The top row contains excitation patterns for the vowel in quiet. The remaining rows contain excitation patterns calculated after the addition of a diotic white-noise masker at signal-to-noise ratios of 0, -10, and -20 dB. The labels ‘F1’ and ‘F2’ mark the frequencies of first and second formants.

The advantages of integrating evidence from the two excitation patterns should not be restricted to the identification of steady-state sounds. Integration would also be beneficial in the analysis of connected speech. Connected speech is a time-varying signal in which the amplitudes of the spectral components may vary from moment to moment over a range of at least 30 dB. At some points in time and frequency, therefore, the instantaneous SNR favors the monaural excitation pattern, but at other points it favors the binaural excitation pattern. In this case, as with steady-state sounds, intelligibility would be maximized by combining evidence from the monaural and binaural excitation patterns.

The present experiments involved steady-state sounds. Experiment 1 tested the hypothesis that listeners can combine evidence of the spectral structure of vowels from the monaural and binaural excitation patterns. In this experiment, the first formant of a vowel was defined in one excitation pattern, while the second formant was defined in the other excitation pattern. Experiments 2 and 3 investigated the extent to which such integration is optional or mandatory. In these experiments, listeners were required to attend selectively to two formants defined in one excitation pattern in the presence of competing formants defined in the other excitation pattern. The results from the three experiments show that listeners can combine evidence of formants from the two excitation patterns, but find it more difficult to attend selectively to formants in one pattern in the presence of competing formants in the other pattern. Overall, the results are compatible with the idea that listeners preferentially combine evidence of the spectral structure of speech from the two

excitation patterns, in accordance with the account of the BILD proposed above.

### A. Monaural and binaural excitation patterns

The monaural excitation patterns illustrated in Fig. 1 were calculated using a standard model of monaural auditory analysis (Patterson *et al.*, 1995). The stimulus was initially filtered by a bandpass filter representing the transfer function of the middle ear. It was then separated into a set of overlapping frequency channels using a gammatone auditory filterbank. The outputs of the frequency channels were half-wave rectified. Their intensity was measured and plotted as a function of the center frequency of the channels in dB.

We define the binaural excitation pattern as the across-frequency profile of interaural decorrelation.<sup>2</sup> The binaural excitation patterns illustrated in Fig. 1 were calculated using an extension to the monaural model. The left and right waveforms of the stimulus were initially filtered by the bandpass filter representing the transfer function of the middle ear. Each of them was separated into two sets of overlapping frequency channels using matched gammatone auditory filterbanks. The outputs of each channel were half-wave rectified. Next, the interaural correlation was calculated by first measuring the maximum cross-product of the left and right outputs  $l(t, f)$  and  $r(t, f)$  of a frequency channel  $f$ , as a function of an internal time delay  $\Delta t$  ranging from  $-5000$  to  $+5000 \mu\text{s}$  applied to one output, and then dividing by the average power of those outputs<sup>3</sup>

TABLE I. Format frequencies used in the experiments. Within the constraint that frequency was quantized to multiples of 2 Hz the bandwidth of each formant was 1 ERB. The labels ‘‘F1’’ and ‘‘F2’’ indicate that a particular formant was the first formant (F1) or second formant (F2) of each vowel.

Center frequency (Hz)	Lower cutoff (Hz)	Upper cutoff (Hz)	Vowel			
			u	i	a	ε
250	224	276	F1	F1		
650	604	698			F1	F1
950	888	1016	F2		F2	
1850	1742	1964		F2		F2

interaural correlation ( $f$ )

$$= \max \left( \frac{\sum_{t=0}^{t=250 \text{ ms}} l(t, f) r(t + \Delta t, f)}{\sqrt{\sum_{t=0}^{t=250 \text{ ms}} l(t, f)^2 \sum_{t=0}^{t=250 \text{ ms}} r(t, f)^2}} \right) \quad (1)$$

The interaural decorrelation in each channel was calculated by subtracting the interaural correlation in that channel from one

interaural decorrelation ( $f$ )

$$= 1 - \text{interaural correlation } (f). \quad (2)$$

Interaural decorrelation was then plotted as a function of the center frequency of the channel to generate the binaural excitation pattern. A logarithmic y axis was chosen to emphasize the correspondence between the binaural and monaural excitation patterns and to reflect, approximately, the fact that the just-noticeable difference in interaural decorrelation increases as interaural decorrelation increases (e.g., Pollack and Trittipoe, 1959).

Interaural decorrelation measures the degree to which the inputs at the left and right ears differ, where ‘‘differ’’ means that the inputs cannot be equated using any combination of internal time delay or intensity change.<sup>4</sup> Interaural decorrelation has been proposed as one of the detection cues underlying the binaural-masking level difference (e.g., Durlach *et al.*, 1986). The focus of the present paper is on the use of suprathreshold levels of interaural decorrelation to recover speech from noise, as in Fig. 1. The across-frequency profile of interaural decorrelation reveals the spectral structure of the vowel because the value of interaural decorrelation in any frequency channel depends upon the relative level of the vowel and the noise in that frequency channel. Recall that the excitation patterns in Fig. 1 were generated by analyzing a vowel with an ITD of +400  $\mu$ s masked by a diotic noise. In those frequency channels where the levels of vowel and noise are approximately equal, the internal time delay which best compensates for the vowel (−400  $\mu$ s) cannot compensate for the noise. Equally, the internal time delay which best compensates for the noise (0  $\mu$ s) cannot compensate for the vowel. Therefore, the level of interaural decorrelation in these channels is relatively high. In other channels that are dominated by either the vowel or the noise, an internal delay of either −400 or 0  $\mu$ s equalizes the filter outputs, so the level of interaural decorrelation is lower. Listeners

have high sensitivity to interaural decorrelation. The just-noticeable difference in interaural decorrelation from a value of zero is in the range from 0.02 to 0.05 for steady-state sounds (e.g., Gabriel and Colburn, 1981; Akeroyd and Summerfield, 1999). By the 0.05 criterion, both F1 and F2 of the vowel in Fig. 1 could be detected in the binaural excitation pattern at −10 dB, and F1 could be detected at −20 dB.

Figure 1 demonstrates that the binaural excitation pattern has didactic value as a means of summarizing binaural information that may support speech identification in noise. Beyond that, however, its status is less well-established than that of the monaural excitation pattern, in at least two respects. First, it has no proven physiological underpinning in the form of a population of units whose activity corresponds to the profile of interaural decorrelation across frequency. Second, the capacity of listeners to distinguish different degrees of interaural decorrelation may not be as great as is suggested by the depth of modulation of excitation across frequency in Fig. 1. Nonetheless, the explanatory power of the pattern is shown by the demonstration that the arrangement of peaks in the pattern can account for the perceived frequencies of an important subset of the dichotic pitches (Culling *et al.*, 1998a, b). In those experiments, listeners performed as if they could access the binaural excitation pattern and interpret its structure in a similar fashion to a monaural excitation pattern. Thus, it is plausible that listeners might also extract evidence of formants from the binaural excitation pattern.

## B. Two-formant vowels

Although evidence of up to five formants can be found in natural vowels, adequate approximations to some vowels can be created by synthesizing a single formant, and the majority of vowels is adequately defined by synthesizing two formants (e.g., Delattre *et al.*, 1952). A synthetic sound consisting of just two narrow bands of noise centered on those frequencies also creates a vowel-like percept. The percept is weak, but the identity of the vowel can be reported reliably. The present experiments used such a set of synthetic two-formant vowel-like stimuli. They were similar to stimuli devised by Culling and Summerfield (1995). F1 was set to either 250 or 650 Hz and F2 was set to either 950 or 1850 Hz (Table I). The four possible combinations give the British-English monophthongal vowels /u/ (found in ‘‘who’d’’), /i/ (‘‘heed’’), /a/ (‘‘hard’’), and /ε/ (‘‘haired’’). The experi-

ments capitalize on the fact that both formants must be detected and their frequencies determined for each vowel in this set to be identified uniquely.

The formants were constructed by increasing either the level or the interaural decorrelation of narrow bands within a diotic white noise. Each band had a width of 1 equivalent-rectangular bandwidth (“ERB”; Glasberg and Moore, 1990). “Monaural formants” were created by increasing the level of a band at either 250, 650, 950, or 1850 Hz. This manipulation created peaks in the monaural excitation pattern at each ear. For this reason, the peaks are referred to as monaural formants although the stimulus was diotic and identical peaks were present at each ear. “Binaural formants” were created by increasing the interaural decorrelation of a band at either 250, 650, 950, or 1850 Hz. This manipulation creates peaks in the binaural excitation pattern.

In summary, formants are described as either monaural or binaural because a monaural formant can be identified from the information present at either ear, whereas a binaural formant can be identified only by a comparison of information from both ears. Likewise, “monaural vowels” had two monaural formants and “binaural vowels” had two binaural formants. Stimuli with two monaural formants or two binaural formants are collectively termed “same-mode” vowels; stimuli with one monaural and one binaural formant are termed “mixed-mode” vowels.

## I. EXPERIMENT 1: INTEGRATION OF INFORMATION ACROSS EXCITATION PATTERN

Experiment 1 investigated whether listeners can combine evidence of formants from the monaural and binaural excitation patterns. The experiment measured the accuracy of identification of synthetic vowels (Fig. 2, left column). Four classes of vowel were used. The same-mode vowels had either two monaural formants (termed “MM” using a terminology of *first formant*, *second formant*) or two binaural formants (“BB”). The mixed-mode vowels had either a monaural first formant combined with a binaural second formant (“MB”) or a binaural first formant combined with a monaural second formant (“BM”). If listeners can combine formant information across excitation patterns, they would achieve the same accuracy in identifying mixed-mode vowels as same-mode vowels.

If instead listeners cannot combine evidence in this way, their responses to the mixed-mode vowels must be determined by attention to each excitation pattern independently. Thus, listeners would report the vowel heard when the monaural formant was presented in isolation or the vowel heard when the binaural formant was presented in isolation. This “single-formant” hypothesis was tested by including four control conditions in which the stimuli contained just one formant. The conditions are termed “M-,” “-M,” “B-,” and “-B” (Fig. 2, right column). Listeners were required to respond /u/, /i/, /a/, or /ε/ according to which vowel they heard—or the closest vowel they heard—when each single-formant stimulus was presented. Four versions of the single-formant hypothesis were evaluated as quantitative models for predicting the accuracy of identification in the two-formant

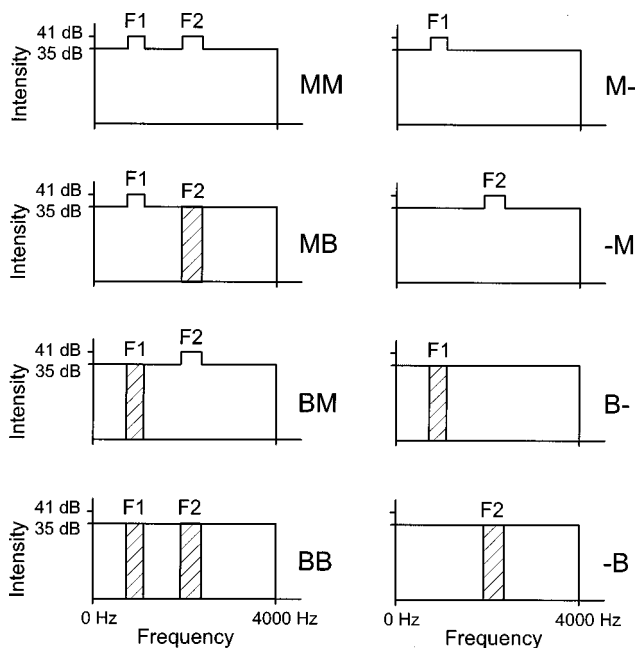


FIG. 2. Schematic illustrations of the stimuli for experiment 1. The left column contains two-formant stimuli. The right column contains one-formant stimuli. Monaural formants (“M”) were created by increasing the level of 1-ERB-wide bands in a diotic noise by 6 dB. Binaural formants (“B”) were created by setting the interaural correlation of 1-ERB-wide bands to zero. Binaural formants are illustrated as hatched bars. Note that binaural formants do not entail an increase in spectral level. The formant frequencies are specified in Table I.

conditions from accuracy observed in the single-formant conditions.

### A. Method

#### 1. Stimuli

The stimuli were created by modifying a digital, flat-spectrum, diotic noise constructed by summing together a sine and cosine frequency response at each frequency from 2 to 4000 Hz (inclusive), so giving the left  $L(t)$  and right  $R(t)$  waveforms

$$L(t) = \sum_{f=2 \text{ Hz}}^{f=4000 \text{ Hz}} (A_{L,f} \sin(2\pi ft) + B_{L,f} \cos(2\pi ft)), \quad (3)$$

$$R(t) = \sum_{f=2 \text{ Hz}}^{f=4000 \text{ Hz}} (A_{R,f} \sin(2\pi ft) + B_{R,f} \cos(2\pi ft)), \quad (4)$$

where  $A_{L,f}$ ,  $A_{R,f}$ ,  $B_{L,f}$  and  $B_{R,f}$  were random variables chosen from a Gaussian distribution. For all frequencies  $f$  other than those defining the binaural formants,  $A_{L,f}$  was equal to  $A_{R,f}$  and  $B_{L,f}$  was equal to  $B_{R,f}$ . The monaural formants were created by setting the intensity of a 1-ERB-wide band to be 6 dB higher than the flat-spectrum noise. The binaural formants were created by setting, within a 1-ERB-wide band, the random variables  $A_{L,f}$  and  $B_{L,f}$  to be statistically independent of  $A_{R,f}$  and  $B_{R,f}$ . This method creates an interaurally decorrelated band with a nominal interaural correlation of zero. The formant frequencies are specified in Table I.



Twenty independent tokens of each stimulus were synthesized at a sampling rate of 20 kHz with 16-bit amplitude quantization. They were converted to analog by a Loughborough Sound Images digital-to-analog converter (AM/D16DS). Each channel was passed separately through a custom-built attenuator, a custom-built headphone amplifier, and a further attenuator (Marconi Instruments type TF2612). They were presented to listeners over both channels of a Sennheiser HD-414 headset. Listeners sat in a double-walled sound booth. The presentation of stimuli and the collection of responses were controlled by a Dell 486 personal computer.

The noise spectrum level of the diotic noise was 35 dB (*re*: 20  $\mu$ Pa). All the stimuli had a duration of 500 ms, including 20-ms onset and offset raised-cosine ramps applied after the summation of components.

## 2. Procedure

A single-interval, four-alternative procedure was used to measure the accuracy of vowel identification. On each trial, one stimulus was presented and listeners were required to identify it using one of four labeled buttons. No feedback about the accuracy of responses was given. Stimuli from the eight conditions (MM, BB, MB, BM, M-, -M, B-, and -B) were randomized together. In each block each combination of vowel and condition occurred three times. There were four sessions of four blocks each, giving a total for the two-formant conditions of 192 trials per condition per listener (=4 vowels $\times$ 3 trials per block $\times$ 16 blocks) and a total for the one-formant conditions of 96 trials per condition per listener (=2 formants $\times$ 3 trials per block $\times$ 16 blocks). At the beginning of each session listeners completed a practice block of 24 trials in which feedback as to the identity of each stimulus was given. These blocks consisted of three examples of each MM vowel<sup>5</sup> and three examples of the same vowels but with the amplitudes of components outside the 1-ERB-wide formant bands set to zero (i.e., without the flat-spectrum diotic noise surrounding the formant peaks). These blocks served to remind listeners of the identity of each vowel in the stimulus set.

## 3. Listeners

Four listeners participated. Their hearing levels were less than or equal to 20 dB HL at octave frequencies between 500 Hz and 4 kHz inclusive. Their ages were 27 years (Listener A), 48 (B), 21 (C), and 24 (D). Listeners A and B were the authors. The other two listeners were paid for their participation. Before data collection began, all listeners received extensive practice in identifying monaural and binaural vowels similar to those used in the experiment.

## 4. Statistical analyses

A  $p < 0.05$  confidence interval based on the binomial distribution tested whether accuracy of identification was above chance for each combination of listener and condition. A single-factor repeated-measures analysis of variance (ANOVA) and a *post hoc* Tukey HSD test tested whether accuracy in any two-formant condition differed from accu-

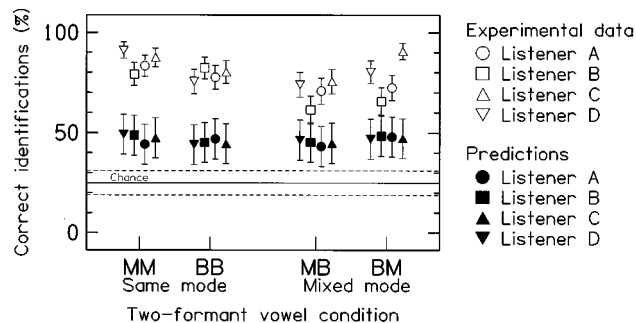


FIG. 3. Results of experiment 1. Mean percent-correct identifications in each two-formant condition, averaged across the four vowels, are shown as open symbols. Predictions from method 1 (described in the Appendix) of the levels of performance in each condition from patterns of responses in the one-formant conditions are shown as filled symbols. The horizontal dashed lines plot the 95%-confidence interval for chance performance. Error bars plot the 95%-confidence intervals estimated using the binomial distribution. Each symbol is based on 192 observations (experimental data) or 96 observations (predictions).

racy in any other two-formant condition. A two-tailed planned comparison tested whether overall accuracy in the two mixed-mode conditions differed from overall accuracy in the two same-mode conditions.

A two-factor (formant frequency $\times$ presentation mode) repeated-measures ANOVA tested whether accuracy in any one-formant condition differed from accuracy in any other one-formant condition. Note that, because *two* formants are required to define each vowel, no single response can be correct in the one-formant conditions. Instead, a response to either of the vowels partially defined by the formant was marked as correct; for example, when the formant was at 250 Hz both /u/ or /i/ were scored as correct responses (Table I).

## B. Results

### 1. Two-formant conditions

The results for the two-formant conditions are shown in Fig. 3. The open symbols plot the mean accuracy of identification in each condition, averaged across the four vowels. The solid symbols plot the predictions from one version of the single-formant hypothesis (see the Appendix). The horizontal dashed lines plot the 95%-confidence interval for chance performance. The error bars plot the 95%-confidence intervals of the observed levels of accuracy of identification.

In all conditions, accuracy of identification was above chance. The mean identification scores were 85% (MM), 79% (BB), 70% (MB), and 77% (BM). Accuracy of identification differed between the four conditions [ $F(3,9)=4.8$ ,  $p=0.03$ ]. The only significant comparison revealed by the *post hoc* test was that the MB vowels were identified less accurately than the MM vowels (Tukey HSD=12.3% at  $p=0.05$ ; difference in means=14.9%). The planned comparison showed that the mixed-mode vowels were identified less accurately than the same-mode vowels [ $t(12)=3.0$ ,  $p=0.01$ ]; the difference in accuracy was 8.3%.

### 2. One-formant conditions

The results for the one-formant conditions are shown in Table II. The majority of responses was to either of the two

TABLE II. Results of experiment 1. Proportion of responses to each vowel in the single-formant conditions, averaged across the four listeners. The responses marked in bold correspond to the two vowels of which each single formant is a component.

Formant frequency (Hz)	Condition	Mode	Responses (percent)			
			u	i	a	ε
250	M-	Monaural	<b>66</b>	<b>23</b>	5	6
650	M-	Monaural	5	2	<b>35</b>	<b>58</b>
950	-M	Monaural	<b>27</b>	1	<b>71</b>	1
1850	-M	Monaural	1	<b>61</b>	1	<b>37</b>
250	B-	Binaural	<b>55</b>	<b>38</b>	4	3
650	B-	Binaural	4	4	<b>36</b>	<b>55</b>
950	-B	Binaural	<b>19</b>	1	<b>76</b>	5
1850	-B	Binaural	11	<b>45</b>	7	<b>36</b>

“correct” vowels in which the formant could occur (cells marked in bold), although responses were not distributed evenly between the two. Only occasionally was a response made to either of the other vowels. Correct performance did not differ between the monaural and binaural conditions [ $F(1,3)=1.7, p=0.3$ ]. There was no main effect of formant frequency [ $F(3,9)=0.96, p=0.5$ ], although there was an interaction between frequency and mode [ $F(3,9)=7.3, p=0.009$ ], in that performance was worst in the 1850-Hz binaural-formant condition.

### 3. Relationship between one-formant and two-formant results

One account of the identification of the two-formant vowels is that listeners combined evidence of F1 and F2. The alternative, single-formant hypothesis, is that listeners based their identification responses on either F1, or F2, but not both. Superficially, this second hypothesis is plausible. Consider first that there was a tendency for listeners to identify each single formant predominantly as one of the four vowels, rather than dividing their responses equally between the two vowels to which the formant could belong (Table II). This outcome was expected. In general, a single formant, placed near the frequency of a concentration of energy in a natural vowel, may be heard as that vowel (e.g., Delattre *et al.*, 1952). Thus, a single formant at a low frequency may be heard as /u/, at a low/mid frequency as /ε/, at a high/mid frequency as /a/, and at a high frequency as /i/. This pattern can be seen in Table II.

The pattern suggests that listeners might have identified two-formant vowels by attending to a particular (“dominant”) single formant in each pair. For example, the 250-Hz monaural formant was identified as /u/ on 66% of presentations as a single-formant stimulus, while the pairing of that formant with the 950-Hz monaural formant yielded the same percentage of /u/ responses (66%). The primary requirement for this explanation to be viable is that there should be a systematic hierarchy of dominance among the formants. The hierarchy can be constructed by assessing each two-formant stimulus as follows, starting with the stimulus composed of a 250-Hz F1 and a 950-Hz F2. This stimulus was identified predominantly as /u/, while, in isolation, the 250-Hz formant was identified predominantly as /u/ and the 950-Hz formant was identified predominantly as /a/. Thus, the 250-Hz for-

mant dominated the 950-Hz formant. The same logic applied to the other two-formant stimuli implies that the 1850-Hz formant dominated the 250-Hz formant, and that the 650-Hz formant dominated the 1850-Hz formant. To complete the hierarchy, it is necessary that the 650-Hz formant should dominate the 950-Hz formant, so that the combination of these two formants would be heard as /ε/. In fact, the combination was heard strongly as /a/, so that 950-Hz formant dominated the 650-Hz formant. Thus, the pattern of dominance is circular, rather than forming a systematic hierarchy.

This outcome is strong evidence against the idea that listeners identified the two-formant vowels by attending selectively to one of the constituent formants because, in the absence of a systematic hierarchy of dominance, listeners would not know which formant in a two-formant stimulus should receive their attention. The conclusion is consolidated in the Appendix, which evaluates four quantitative methods for predicting the two-formant results from the one-formant results. One of these methods assumes, implausibly, that listeners knew which single formant should receive their attention to maximize their accuracy. Both this method, and the other three which did not include the assumption, predict a lower level of accuracy in identifying two-formant vowels than the listeners achieved. Taken together, the arguments in this section and the analyses in the Appendix imply that listeners identified the two-formant vowels from the evidence of both formants.

### C. Discussion

There are four results of interest. First, listeners achieved accuracy that was significantly above chance when identifying two-formant vowels composed of one monaural formant and one binaural formant. Second, accuracy of identification of these vowels was nevertheless poorer than the accuracy achieved with vowels composed of two monaural formants or two binaural formants. The penalty from the requirement to combine information across excitation patterns was small; however, it amounted to a reduction in accuracy of less than 10%. Third, accuracy of identification in the mixed-mode conditions did not depend on which formant was monaural and which formant was binaural. This result was not expected, given that the ease with which interaural decorrelation can be detected declines with increasing fre-

quency (e.g., Gabriel *et al.*, 1992; Akeroyd and Summerfield, 1999). The expected result was shown in other ways, however. In the single-formant conditions, the binaural formant at 1850 Hz was identified less consistently than the binaural formants at the three lower frequencies. In addition, compared with performance when both formants were monaural, performance was significantly lower when the higher formant was binaural but not when the lower formant was binaural. The fourth result was that in all conditions accuracy was higher than predicted by any of the versions of the single-formant hypothesis that are evaluated in the Appendix. We conclude that listeners based their identification responses to the two-formant vowels on evidence of both formants together, rather than by attending to one mode on some trials and the other mode on other trials.

Each binaural formant was created by interaurally decorrelating a narrow region of an originally diotic noise. This manipulation creates a dichotic pitch (Akeroyd and Summerfield, 2000) perceptually similar to the Huggins pitch (Cramer and Huggins, 1958). Culling *et al.* (1998a) showed that pairs of formants created using a different dichotic pitch—the binaural edge pitch (Klein and Hartmann, 1981)—can also be identified as vowels. They used a similar design to the BB condition of the present experiment, but with F1 set to either 225 or 625 Hz and F2 set to either 975 or 1925 Hz. They observed a mean accuracy of identification of 59%. Although above chance, this value is smaller than was observed here in the BB condition. The difference probably arises because the binaural edge pitch can be difficult to perceive at the highest frequency (1925 Hz) at which Culling *et al.* placed a formant. In contrast, the results from the one-formant conditions of the present experiment show that the highest formant (1850 Hz) was detected reliably in binaural conditions (Table II), though with lower accuracy than the formants at the other three frequencies.

In summary, the results of experiment 1 indicate that listeners can combine formant information from the monaural excitation pattern with formant information from the binaural excitation pattern. The authors' experience supports this conclusion. In the mixed-mode conditions we perceived a single noisy vowel, although with some stimuli we were also aware of the tonal quality of the binaural formant.

## II. EXPERIMENT 2: SELECTIVE ATTENTION TO ONE EXCITATION PATTERN

Experiment 2 sought to determine whether the integration of monaural and binaural information observed in experiment 1 is mandatory or optional. The design of experiment 1 deliberately encouraged integration, insofar as listeners were instructed to make a single response on each trial, and integration resulted in percepts that permitted more consistent patterns of response than did attention to either mode alone. Accordingly, in experiment 2 a strong *disincentive* for integration was introduced. Pairs of two-formant vowels were presented simultaneously. One vowel was defined monaurally (in the sense described in Sec. B of the Introduction) and the other was defined binaurally. Binaural and monaural vowels can be distinguished because they have different timbres and lateral positions. A binaural

vowel presented in isolation has a tonal timbre and is lateralized to one or the other side of the head. A monaural vowel presented in isolation has a noisy timbre and is lateralized at the center of the head. Listeners could therefore be instructed to identify one or the other vowel. The design of the stimuli ensured that any combination of formants across excitation patterns would generally result in an incorrect identification.

The spectral contrast of each monaural formant was varied between 0 and 12 dB. If listeners can selectively attend to the binaural excitation pattern, then the accuracy of identification of the binaural vowel will be above chance and will be independent of the spectral contrast of the monaural vowel. Furthermore, if listeners can selectively attend to the monaural excitation pattern, then the accuracy of identification of the monaural vowel will not be reduced by the presence of the binaural vowel.

### A. Method

There were two sets of stimuli. In one set, each stimulus was composed of a binaural vowel paired with different monaural vowel ("binaural-present" stimuli; Fig. 4). Only the 12 nonidentical pairs that can be formed from four vowels were included. The second set of stimuli consisted of monaural vowels alone ("binaural-absent" stimuli). Listeners performed two tasks in separate sessions. In one task, they heard only the binaural-present stimuli and were required to identify the binaural vowel. In the other task, they heard a randomized mixture of binaural-present and binaural-absent stimuli and were required to identify the monaural vowel. Listeners undertook the sessions with the monaural task before those with the binaural task.

The stimuli were constructed, calibrated, and presented to listeners using the same techniques and apparatus as in experiment 1. A single value of interaural correlation—nominally zero—defined the binaural formants. Five different values of spectral contrast—0, 3, 6, 9, and 12 dB—defined the monaural formants (when the spectral contrast was 0 dB, the stimulus had a flat spectrum and no monaural vowel was specified). The two monaural formants in a stimulus always had the same spectral contrast. Twenty independent tokens of each stimulus were created. Pilot testing showed that monaural vowels composed of formants defined by 3 dB of spectral contrast were less clear than the binaural vowels, while monaural vowels composed of formants defined by 12 dB of spectral contrast were clearer than the binaural vowels. Thus, the experiment would explore the ability of listeners to segregate binaural formants with one degree of clarity when in competition with monaural formants covering a range of clarity.

A single-interval, four-alternative forced-choice procedure similar to experiment 1 was used. Across multiple sessions, listeners undertook a series of blocks containing a total of 240 trials for each spectral contrast for each of the binaural task with the binaural-present stimuli, the monaural task with the binaural-present stimuli, and the monaural task with the binaural-absent stimuli (= 12 vowel pairs  $\times$  20 blocks  $\times$  1 trial per block in each case). Each session was preceded by a 24-trial practice block. For the binaural tasks this block contained single binaural vowels (BB). For the monaural

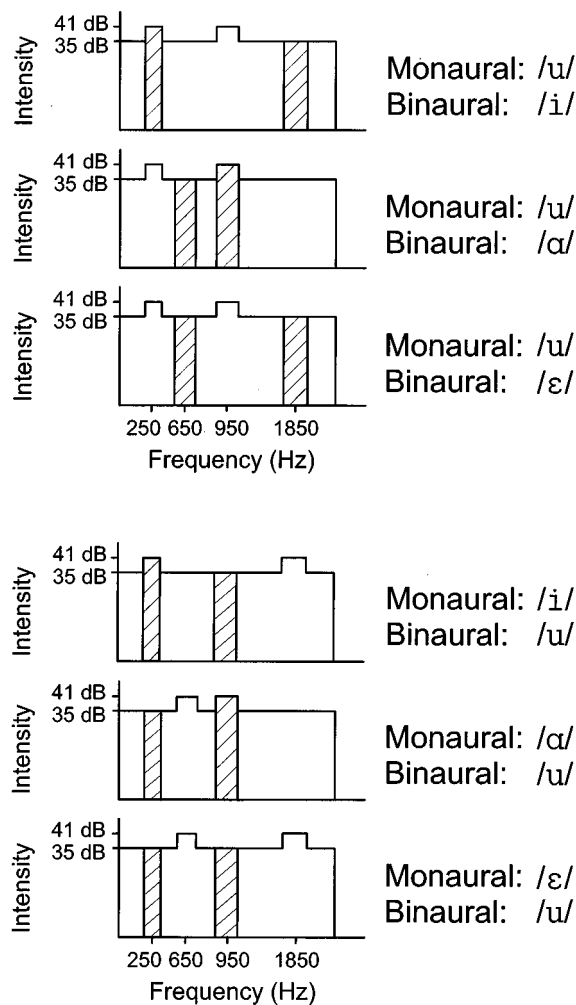


FIG. 4. Schematic illustrations of a subset of the binaural-present stimuli for experiment 2 (cf Fig. 2). All combinations for a monaural /u/ (upper three panels) and for a binaural /u/ (lower three panels) are shown. Some combinations require formants to be presented at three frequencies; other combinations require formants at four frequencies. Note that in the three-formant combinations, a binaural formant coincides with a monaural formant and so is associated with an increase in spectral level.

tasks this block contained single monaural vowels (MM). Listeners were thus reminded of the timbre and lateral position of the to-be-identified vowels. The same four listeners as before participated.

A  $p < 0.05$  confidence interval based on the binomial distribution tested whether accuracy of identification by individual listeners in individual conditions was above chance. A single-factor (spectral contrast) repeated-measures ANOVA tested whether accuracy of identification in the binaural task depended on the spectral contrast of the monaural vowels. A two-factor (binaural presence/absence  $\times$  spectral contrast) repeated-measures ANOVA tested whether accuracy of identification in the monaural task depended on the presence or absence of the binaural vowel. In the 0-dB spectral contrast condition, no monaural vowel was defined. Accordingly, the majority of the ANOVAs considered the range of spectral contrast from 3 to 12 dB. The exception was the analysis of the number of correct binaural identifications in the binaural task. The binaural vowel did exist in the 0-dB

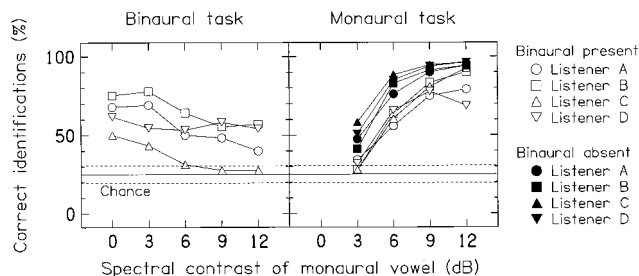


FIG. 5. Results of experiment 2. The left panel shows mean percent-correct identifications of the binaural vowel in the binaural task. The right panel shows mean percent-correct identifications of the monaural vowel, in the monaural task for the binaural-present stimuli (open symbols), and for the binaural-absent stimuli (solid symbols). The results are averaged across the 12 vowel pairs. The horizontal dashed lines plot the 95%-confidence interval for chance performance. Each symbol is based on 240 observations.

condition, and so the ANOVA considered the full range of spectral contrasts from 0 to 12 dB.

## B. Results

### 1. Binaural task

The results for the binaural task are shown in the left panel of Fig. 5. The mean accuracy of identification of the binaural vowel deteriorated as the spectral contrast of the monaural vowel increased; accuracy decreased from 64% at 0-dB spectral contrast to 45% at 12-dB spectral contrast [ $F(4,12) = 9.4$ ,  $p = 0.001$ ]. Nonetheless, accuracy of identification was above chance in all but two of the 20 combinations of listener and spectral contrast.

### 2. Monaural task

The results for the monaural task are shown in the right panel of Fig. 5. The mean accuracy of identification of the monaural vowel in the binaural-absent stimuli (filled symbols) was higher than in the binaural-present stimuli (open symbols); the mean difference was 17% [ $F(1,3) = 67$ ,  $p = 0.004$ ]. With the binaural-present stimuli accuracy improved as the spectral contrast increased: from 31% at 3-dB spectral contrast to 83% at 12-dB spectral contrast [ $F(3,9) = 52$ ,  $p < 0.0001$ ]. Accuracy of identification was above chance in all but two of the 16 combinations of listener and spectral contrast. With the binaural-absent stimuli accuracy also improved: from 50% at 3 dB to 95% at 12 dB [ $F(3,9) = 159$ ,  $p < 0.0001$ ]. Accuracy of identification was above chance in all 16 combinations of listener and spectral contrast.

## C. Discussion

In the majority of conditions listeners achieved levels of performance that were above chance when identifying a monaural vowel in the presence of a competing binaural vowel and vice versa. Evidence that listeners can selectively attend to each excitation pattern individually comes from the conditions where the spectral contrast of the monaural vowel was 6 dB. In this condition, all four listeners performed significantly above chance both when identifying the binaural vowel and when identifying the monaural vowel.<sup>6</sup> The same

pattern was also shown by listeners A, B, and D at spectral contrasts of 9 and 12 dB. The second relevant result is that accuracy in identifying a binaural vowel declined as the spectral contrast of the competing monaural vowel was increased. Also, the accuracy of identifying the monaural vowel was reduced when a competing binaural vowel was present. Therefore, selective attention is not perfect, in that listeners cannot completely discount the other excitation pattern.

The stimuli in the binaural condition with 0-dB spectral contrast were the same as those in the BB condition of experiment 1. Performance was, however, poorer in experiment 2 (64%) than experiment 1 (79%). The difference may reflect a decline in the listeners' motivation to perform well in a difficult condition. However, differences in the requirements of the two experiments may also have contributed. In experiment 1 only a single vowel was presented on each trial and listeners were required to report whatever vowel they heard, without regard for the modality from which the evidence of the vowel originated. In the binaural conditions of experiment 2, listeners were required to attend to the binaural excitation pattern and to actively disregard evidence from the monaural excitation pattern. On trials where the spectral contrast of the monaural vowel was 6, 9, or 12 dB, the stimuli provided a strong percept of a monaural vowel. To perform accurately, therefore, it was necessary to ignore the strong percept and report the weaker competing percept of the binaural vowel. These trials were randomized together with trials where the spectral contrast of the monaural vowels was 0 or 3 dB. Adopting the strategy of ignoring the strong percept on these trials would, however, have resulted in listeners rejecting the binaural vowel. Instead, on some trials, they might have mistaken the phonetic impression created by the noise as evidence of the binaural vowel, thereby lowering their accuracy. In summary, differences between the experiments in the context in which stimuli were presented could have contributed to the relatively poor performance of listeners in the binaural task when the spectral contrast of the monaural vowel was 0 or 3 dB in experiment 2.

### III. EXPERIMENT 3: EFFECT OF ONSET ASYNCHRONY ON SELECTIVE ATTENTION

The foregoing argument implies that a manipulation which draws attention to the binaural formants should improve the ability of listeners to hear out the binaural vowel in a stimulus that also contains a monaural vowel. Two co-occurring stimuli are more distinct perceptually and the mutual contamination of their perceptual properties is reduced, if they start at different times than if they start at the same time (e.g., Darwin and Carlyon, 1995). Accordingly, in experiment 3 an onset asynchrony was introduced between the monaural and binaural formants.

#### A. Method

The stimuli were constructed, calibrated, and presented to listeners using the same techniques and apparatus as in experiments 1 and 2. An onset asynchrony was introduced into stimuli similar to those used in experiment 2. In these conditions one vowel started 480 ms before the other vowel

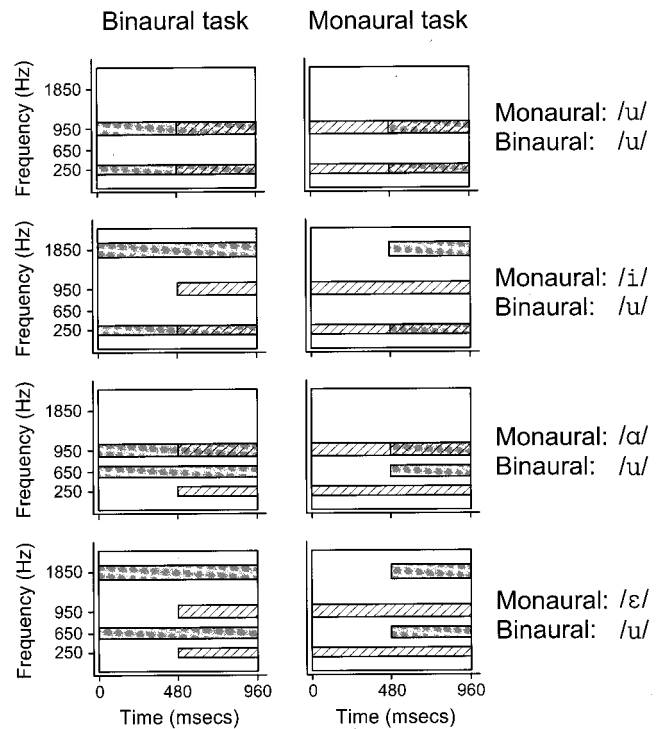


FIG. 6. Schematic spectrograms of a subset of vowel pairs used in experiment 3. All the combinations for a binaural /u/ are illustrated. A monaural formant is represented as a light-gray bar. A binaural formant is represented as a hatched bar. In the binaural task (left column) listeners were required to identify the binaural vowel. In the monaural task (right column) they were required to identify the monaural.

(Fig. 6). The two vowels ended together after a further 480 ms. In the binaural “onset-asynchrony” conditions the binaural vowel started *second*. This order was reversed in the monaural onset-asynchrony conditions. In the “simultaneous” conditions both vowels were 500 ms in duration and began and ended together, as in experiment 2.

Each binaural onset-asynchrony stimulus was created by digitally editing two of the stimuli used in experiments 1 and 2. The 20-ms offset ramp of an MM vowel from experiment 1 and the 20-ms onset ramp of a monaural–binaural vowel pair from experiment 2 were digitally removed and then the two new waveforms were abutted. The click resulting from spectral splatter at the join of the two waveforms was removed by applying a steep 4-kHz low-pass digital filter, thus limiting the bandwidth of the new stimulus to the original bandwidth of its constituents. The identity of the MM vowel was the same as the monaural vowel in the monaural–binaural vowel pair. Each new stimulus was 960 ms in duration, including 20-ms raised-cosine ramps at its offset and onset. A similar procedure was used to create each monaural onset-asynchronous stimulus, except that a BB vowel was used instead of an MM vowel. Its identity was the same as the binaural vowel in the monaural–binaural vowel pair. Finally, the simultaneous stimuli were constructed in the same way as the monaural–binaural vowel-pair stimuli used in experiment 2. Ten independent tokens of each stimulus were created.

In contrast to experiment 2, all 16 pairings of the vowels

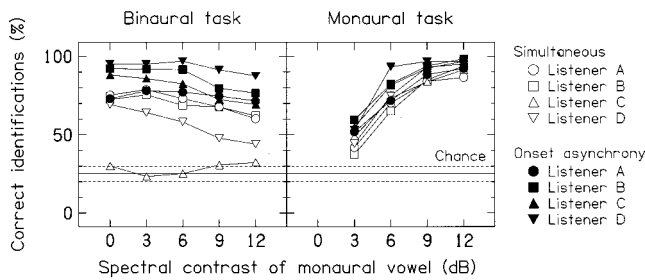


FIG. 7. Results of experiment 3. The left panel shows mean percent-correct identifications of the binaural vowel in the binaural task, in the onset-asynchrony conditions (filled symbols) and the simultaneous conditions (open symbols). The right panel shows mean percent-correct identifications of the monaural vowel in the monaural task, in the onset-asynchrony conditions (solid symbols) and the simultaneous conditions (open symbols). The results are averaged across the 16 vowel pairs. The horizontal dashed lines plot the 95%-confidence interval for chance performance. Each symbol is based on 320 observations.

were generated, including the identical pairings as well as the nonidentical pairings. This arrangement ensured that the identity of the vowel that started first in an onset-asynchrony condition could not restrict the possible choices for the vowel that started second. As a result, meaningful comparisons can be made between levels of performance in the onset-asynchrony and simultaneous conditions.

A single-interval, four-alternative forced-choice procedure similar to that of experiment 2 was used. Each combination of task (binaural and monaural) and synchrony (onset-asynchrony and simultaneous) was undertaken in a separate session. In each block, each combination of vowel pair and spectral contrast occurred once in a random order. There were four sessions of five blocks each, giving a total of 320 trials per spectral contrast per combination of task and synchrony per listener (= 16 vowel pairs  $\times$  20 blocks  $\times$  1 trial per block). Each session was preceded by a 24-trial practice block. For the binaural tasks this block contained single binaural vowels (BB). For the monaural tasks it contained single monaural vowels (MM). The same four listeners as before participated.

A  $p < 0.05$  confidence interval based on the binomial distribution tested whether accuracy of identification by individual listeners in individual conditions was above chance. A two-factor (onset-asynchrony/simultaneous  $\times$  spectral contrast) repeated-measures ANOVA and a nonparametric sign test based on the binomial distribution were performed for the data of each task separately to establish whether performance depended upon onset asynchrony or the spectral contrast of the monaural vowels.

## B. Results

### 1. Binaural task

The results for the binaural task are shown in the left panel of Fig. 7. The results are averaged across the 16 vowel pairs. Listener C performed at chance in the simultaneous conditions. The other three listeners performed above chance in all conditions. Although the difference in accuracy between the onset-asynchrony conditions (solid symbols) and simultaneous conditions (open symbols) was 27%, it was not

statistically significant in the ANOVA [ $F(1,3) = 6.7$ ,  $p = 0.08$ ]. The lack of significance was due to the large individual differences; Listener A showed a particularly small effect. Nevertheless, of the 16 individual comparisons between onset-asynchrony and simultaneous conditions (= 4 listeners  $\times$  4 nonzero levels of spectral contrast), 15 show numerically higher accuracy in the onset-asynchrony condition. A nonparametric sign test shows that the probability of 15 out of 16 outcomes having the same sign, when the probability of an outcome having either sign is 0.5, is 0.0002. In both conditions accuracy deteriorated as the spectral contrast of the monaural vowel was increased [ $F(4,12) = 17.0$ ,  $p < 0.001$ ]. For the onset-asynchrony condition accuracy decreased from 87% at 0-dB spectral contrast to 83% at 12-dB spectral contrast. For the simultaneous condition accuracy decreased from 62% at 0-dB spectral contrast to 50% at 12-dB spectral contrast.

### 2. Monaural task

The results for the monaural task are shown in the right panel of Fig. 7. The four listeners performed above chance in all conditions. The accuracy of identification of the monaural vowel was significantly higher in the onset-asynchrony conditions than in the simultaneous conditions; the difference in mean identification score was 8.3% [ $F(1,3) = 29$ ,  $p = 0.01$ ]. Fifteen of the 16 comparisons showed higher accuracy in the onset-asynchrony conditions than in the simultaneous conditions (sign test;  $p = 0.0002$ ).

## C. Discussion

Making some allowance for individual differences in performance, the introduction of an onset asynchrony between a binaural vowel and monaural vowel improved the accuracy of identification of the vowel which started second. Numerically, the average advantage was larger when listeners attended to the binaural vowel (27%) than the monaural vowel (8%). In the binaural task, there was a wider spread of performance among listeners in the simultaneous condition than the onset-asynchrony condition. The introduction of the onset asynchrony reduced the spread and raised performance to over 70% correct for each listener, independent of the spectral contrast of the competing monaural vowel. In other words, listeners could identify the binaural vowel relatively accurately, provided that their attention was drawn to it. The authors' experience supports this explanation. The onset asynchrony increased the clarity of the differences in tonality and lateralization which distinguished the monaural and binaural vowels.

The poor performance of Listener C on the binaural task with the simultaneous stimuli is puzzling, given that he performed as well as the other listeners in the other three conditions. While the possibility cannot be ruled out that his motivation was simply low in this condition, the difficulties created by the context in which the stimuli were presented (discussed in Sec. II C) may also have contributed to his poor performance.

The monaural task produced a different pattern of results: the spread of performance among listeners was smaller

than in the binaural task and the onset asynchrony led to only a small increase in accuracy. Subjectively, the onset asynchrony did not cause the monaural vowel to stand out strongly as a separate object from the binaural vowel and the background noise.

#### IV. CONCLUSIONS

The main result of these experiments is the demonstration that evidence of vowel formants can be combined across the monaural and binaural excitation patterns (experiment 1), despite listeners having some ability to attend selectively to either excitation pattern (experiment 2). The vowel in noise illustrated in Fig. 1 shows why the strategy of integrating evidence can be advantageous. The formants create peaks in both the monaural and binaural excitation patterns. If the cues that might cause listeners to segregate the peaks into separate auditory objects are weak—as the results of the three experiments show they are—then the peaks will be integrated into a single representation. This outcome accords with the physical reality that a single source of sound has generated both sets of peaks. Integration is therefore appropriate. In contrast to this situation, the binaural vowels were heard out consistently when they started after the monaural vowel (experiment 3). In this situation, spectral peaks in one excitation pattern start before those in the other excitation pattern. The peaks cannot, therefore, have originated from the same source. Logically, they should be segregated.

Evidence of spectral peaks in the binaural excitation pattern can be difficult to hear out when there is competing structure in the monaural excitation pattern. Listeners performed relatively poorly when required to report binaural vowels in experiments 2 and 3, even when those vowels were not in competition with monaural vowels. Poor performance occurred despite listeners having received training in identifying binaural vowels, and having performed adequately in experiment 1.

These findings set limits on the circumstances in which listeners are likely to benefit from the strategy of integrating monaural and binaural evidence of formants. There are two primary requirements for successful application of the strategy to recover speech from competing sounds. First, the competing sounds must be interaurally highly correlated so that they can be selectively decorrelated at specific frequencies by the formants of the speech. Second, the competing sounds must have a uniform spectral structure so that all of the peaks in the two excitation patterns derive from the speech. These conditions are *not* met when there is reverberation, which itself decorrelates the competing sounds, or when the competing sounds consist of voices, which themselves possess spectral structure. Instead, these conditions generally are met when speech is masked by interaurally coherent white or pink noise. Although such noises are encountered outside the laboratory, they are unusual. Instead, the major application of the strategy of integration is likely to be in demonstrations of the BILD.

#### V. SUMMARY

- (1) It was hypothesized that the binaural identification of speech in noise would be facilitated over a range of signal-to-noise ratios by the ability to combine information found only in the monaural excitation pattern (i.e., spectral peaks in intensity) with information found only in the binaural excitation pattern (i.e., spectral peaks in interaural decorrelation).
- (2) The results of experiment 1 confirmed that such integration can occur. Listeners attained levels of performance that were significantly above chance when identifying two-formant vowel-like sounds in which one formant was defined by a peak in the monaural excitation pattern while the other formant was defined by a peak in the binaural excitation pattern.
- (3) Experiments 2 and 3 demonstrated that integration is not mandatory, although the intrinsic cues for segregating evidence of formants in the monaural and binaural excitation patterns are relatively weak. Experiment 2 showed that the accuracy of identification of a monaural vowel is impaired by the presence of a simultaneous binaural vowel, and vice versa. Experiment 3 showed that segregation, particularly of the binaural vowel, can be increased by introducing an onset asynchrony between the competing vowels.
- (4) Overall, the results show that listeners can integrate evidence of formants from the binaural and monaural excitation patterns. That strategy should be advantageous when identifying speech binaurally in noise because sounds from a single source, the talker, create peaks simultaneously in both excitation patterns. The strategy is likely to contribute to demonstrations of the binaural intelligibility level difference.

#### ACKNOWLEDGMENTS

We thank John Foster for technical assistance in the early stages of this project. Hedwig Gockel, two anonymous reviewers, and associate editor Wes Grantham made helpful comments on a previous version of the paper. Part of this work was presented at the meetings of the Acoustical Society of America held in Spring 1998 [M. A. Akeroyd, A. Q. Summerfield, and J. R. Foster, "Integrating monaural and binaural spectral information," *J. Acoust. Soc. Am.* **103**, 2976(A) (1998) and Proceedings of the 16th ICA and 135th Meeting-ASA, Vol. III, pp. 1975–1976], and Spring 1999 [A. Q. Summerfield, J. F. Culling, and M. A. Akeroyd, *J. Acoust. Soc. Am.* **105**, 1158(A) (1999)].

#### APPENDIX: ANALYSIS OF SINGLE-FORMANT RESPONSES IN EXPERIMENT 1

Four methods were evaluated for generating numerical predictions of the response patterns in the two-formant conditions from the patterns in the single-formant conditions. Table AI lists the observed proportions of correct responses made to each of two-formant stimuli in each of the four conditions, along with the proportions predicted by the four methods. The following terminology is used in describing the methods. The proportions of responses to an isolated F1

TABLE AI. Top panel: Observed percentages of correct responses to individual two-formant stimuli and the mean across stimuli in four conditions (MM, BB, MB, BM). Bottom panel: percentages predicted by the four methods described in the Appendix. In cases marked by underlining, the predicted percentage equals or exceeds the observed percentage.

		Formant frequencies (Hz)				
	Condition	250–950	250–1850	650–950	650–1850	Mean
Observed responses						
	MM	66	88	96	92	<b>85</b>
	BB	77	71	95	72	<b>79</b>
	MB	64	58	79	80	<b>70</b>
	BM	64	86	89	70	<b>77</b>
		Formant frequencies (Hz)				
Method	Condition	250–950	250–1850	650–950	650–1850	Mean
Predicted responses						
1 and 4	MM	46	42	53	48	<b>47</b>
	BB	37	41	56	46	<b>45</b>
	MB	42	34	55	47	<b>45</b>
	BM	41	49	54	46	<b>48</b>
2	MM	27	61	71	37	<b>49</b>
	BB	55	38	36	55	<b>46</b>
	MB	<u>66</u>	23	35	58	<b>46</b>
	BM	<u>27</u>	61	71	37	<b>49</b>
3	MM	66	61	71	58	<b>64</b>
	BB	<u>55</u>	45	76	55	<b>58</b>
	MB	<u>66</u>	45	76	58	<b>61</b>
	BM	<u>55</u>	61	71	55	<b>61</b>

are  $u_1$ ,  $i_1$ ,  $a_1$ , and  $\varepsilon_1$  (where  $u_1$  is the proportion of /u/ responses,  $i_1$  of /i/ responses,  $a_1$  of /a/ responses, and  $\varepsilon_1$  of /ε/ responses, respectively). The proportions of responses to an isolated F2 are  $u_2$ ,  $i_2$ ,  $a_2$ , and  $\varepsilon_2$ .

Method 1 was based on the assumptions that listeners (a) attend to only one of the formants on each trial, but (b) attend equally often to the lower and higher formant on different trials, and (c) make the response which the attended formant would receive if it were presented in isolation. In effect, therefore, the method predicts that the response pattern to a two-formant vowel is the average of the response patterns made to the constituent single formants. The proportions of the four responses were predicted to be:  $(u_1 \times 0.5 + u_2 \times 0.5)$ ,  $(i_1 \times 0.5 + i_2 \times 0.5)$ ,  $(a_1 \times 0.5 + a_2 \times 0.5)$ , and  $(\varepsilon_1 \times 0.5 + \varepsilon_2 \times 0.5)$ . These percentages of correct responses predicted by this method are plotted as the solid symbols in Fig. 3. Table AI shows that they are lower than the observed percentages for all 16 stimuli.

The weighting of formants adopted in method 1 can be denoted “(0.5, 0.5),” where the first number in brackets is the proportion of trials on which evidence is taken from F1 and the second number is the proportion of trials on which evidence is taken from F2. Method 2 was a variant of method 1. It explored the consequences of attending to one formant to the exclusion of the other on all trials in a condition. In other words, within a condition, the weighting was either (1, 0) or (0, 1). The effect of adopting these extreme settings of the weights, rather than (0.5, 0.5), was small. In the MM condition, for example, it caused the predicted overall accuracy of identifications to range from 46% (1, 0) to 49% (0, 1). Table AI lists the highest accuracy that could be

achieved with the weights fixed across the formant pairings within a condition, but varying between conditions. The (0, 1) weighting gave the higher mean score in the MM and BM conditions. The (1, 0) weighting gave the higher mean score in the BB and MB conditions. For only 1 of the 16 stimuli did the method predict an accuracy equal to, or higher than, that observed in the two-formant conditions.

Method 3 was a further variant of method 1. It was based on the assumption that, given a two-formant vowel, listeners attend to the formant which, individually, is most likely to yield the correct response. It is not clear how listeners could learn this *post hoc* strategy. Nonetheless, it was evaluated to test whether performance in the two-formant conditions exceeds that predicted by the optimal, though implausible, one-formant strategy. It is illustrated by considering the pairing of a 250-Hz monaural F1 and a 950-Hz monaural F2. The correct response to this pairing is /u/. As a single formant, the 250-Hz monaural F1 is heard as /u/ on 66% of presentations, while the 950-Hz monaural F2 is heard as /a/ on 71% of presentations. Thus, the correct response to the two-formant stimulus is obtained by attending to the 250-Hz F1. If listeners could adopt this strategy, they would identify the two-formant stimulus as /u/ on 66% of trials. Table AI shows that this method predicts an accuracy of identification equal to, or higher than, the observed accuracy for only 2 of the 16 two-formant stimuli.

Method 4 was based on the idea that listeners attend to the two formants individually, and assess the probability that each provides evidence of each of the four vowels. This analysis yields four response probabilities for each formant. Two responses are selected randomly in accord-



ance with these probabilities. They are combined to generate a single response according to the following rules: (i) If both responses are the same, that response is chosen; (ii) if the responses differ, a random choice is made between them with an equal probability of either being chosen. Thus, for example, the probability of the response /u/ being chosen is given by the sum of seven products:  $(u_1 \times u_2) + (u_1 \times i_2)/2 + (u_1 \times a_2)/2 + (u_1 \times \varepsilon_2)/2 + (i_1 \times u_2)/2 + (a_1 u_2)/2 + (\varepsilon_1 \times u_2)/2$ . The accuracy of correct responses predicted by this method is the same as that predicted by method 1 and hence, like that method, underestimates the accuracy observed in the two-formant conditions.

Unlike methods 2 and 3, methods 1 and 4 do not assume that listeners are biased to attend predominantly to one or other of the formants in a pair. Such a lack of bias is compatible with the circular pattern of dominance among the single formants (Sec. IB 3). For this reason, the predictions selected for inclusion in Fig. 3 are those provided by method 1 (and identically by method 4).

<sup>1</sup>The frequencies of the higher formants were 2547 Hz ( $F_3$ ), 3272 Hz ( $F_4$ ), and 4500 Hz ( $F_5$ ).

<sup>2</sup>The binaural excitation patterns defined in this paper are similar to the “residual activation” spectra (occasionally referred to as “minimal activation spectra”) generated by Culling and Summerfield’s (1995) modified equalization–cancellation model.

<sup>3</sup>The range of values of  $\Delta t$  was constant across frequency. It was chosen to accommodate the internal delay required to cancel an interaurally inverted noise in order to yield a masking-level difference at 100 Hz. Each summation is performed over the 250-ms duration of the stimulus.

<sup>4</sup>Grantham (1995, pp. 302–303) provided a succinct definition of interaural correlation: “Interaural correlation is loosely defined as the point-by-point correlation coefficient computed for a stimulus segment after an appropriate delay is imposed on one of the inputs to maximize the correlation.”

<sup>5</sup>In these practice blocks the MM stimuli had a spectral contrast of 12 dB, not 6 dB.

<sup>6</sup>With 6 dB of spectral contrast in the binaural task, Listener C scored 31.3% compared with a chance level of 30.5%.

Akeroyd, M. A., and Summerfield, A. Q. (1999). “A binaural analog of gap detection,” *J. Acoust. Soc. Am.* **105**, 2807–2820.

Akeroyd, M. A., and Summerfield, A. Q. (2000). “The lateralization of simple dichotic pitches,” *J. Acoust. Soc. Am.* (to be published).

Blauert, J. (1983). *Spatial Hearing* (MIT, Cambridge).

Cramer, E. M., and Huggins, W. H. (1958). “Creation of pitch through binaural interaction,” *J. Acoust. Soc. Am.* **30**, 413–417.

Culling, J. F., and Summerfield, Q. (1995). “Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay,” *J. Acoust. Soc. Am.* **98**, 785–797.

Culling, J. F., Summerfield, Q., and Marshall, D. H. (1994). “Effects of simulated reverberation on the use of binaural cues and fundamental-frequency differences for separating concurrent vowels,” *Speech Commun.* **14**, 71–95.

Culling, J. F., Summerfield, A. Q., and Marshall, D. H. (1998a). “Dichotic pitches as illusions of binaural unmasking. I. Huggins’ pitch and the ‘binaural edge pitch,’” *J. Acoust. Soc. Am.* **103**, 3509–3526.

Culling, J. F., Marshall, D. H., and Summerfield, A. Q. (1998b). “Dichotic pitches as illusions of binaural unmasking. II. The Fourcin pitch and the dichotic repetition pitch,” *J. Acoust. Soc. Am.* **103**, 3527–3539.

Darwin, C. J., and Carlyon, R. P. (1995). “Auditory grouping,” in *Hearing*, edited by B. C. J. Moore (Academic, London).

Delattre, P., Liberman, A. M., Cooper, F. S., and Gerstman, L. J. (1952). “An experimental study of the acoustic determinants of vowel color: Observations on one- and two-formant vowels synthesized from spectrographic patterns,” *Word* **8**, 195–210.

Durlach, N. I., Gabriel, K. J., Colburn, H. S., and Trahiotis, C. (1986). “Interaural correlation discrimination. II. Relation to binaural unmasking,” *J. Acoust. Soc. Am.* **79**, 1548–1557.

Gabriel, K. J., and Colburn, H. S. (1981). “Interaural correlation discrimination. I. Bandwidth and level dependence,” *J. Acoust. Soc. Am.* **69**, 1394–1401.

Gabriel, K. J., Koehnke, J., and Colburn, H. S. (1992). “Frequency dependence of binaural performance in listeners with impaired binaural hearing,” *J. Acoust. Soc. Am.* **91**, 336–347.

Glasberg, B. R., and Moore, B. C. J. (1990). “Derivation of auditory filter shapes from notched-noise data,” *Hear. Res.* **47**, 103–138.

Grantham, D. W. (1995). “Spatial hearing and related phenomena,” in *Hearing*, edited by B. C. J. Moore (Academic, London).

Klatt, D. H. (1980). “Software for a cascade/parallel formant synthesizer,” *J. Acoust. Soc. Am.* **67**, 971–995.

Klein, M. A., and Hartmann, W. M. (1981). “Binaural edge pitch,” *J. Acoust. Soc. Am.* **70**, 51–61.

Levitt, H., and Rabiner, L. R. (1967). “Binaural release from masking for speech and gain in intelligibility,” *J. Acoust. Soc. Am.* **42**, 601–608.

Licklider, J. C. R. (1948). “The influence of interaural phase transitions upon the masking of speech by white noise,” *J. Acoust. Soc. Am.* **20**, 150–159.

Patterson, R. D., Allerhand, M. H., and Giguère, C. (1995). “Time-domain modeling of peripheral auditory processing: A model architecture and a software platform,” *J. Acoust. Soc. Am.* **98**, 1890–1894.

Pollack, I., and Trittipoe, W. J. (1959). “Binaural listening and interaural noise cross correlation,” *J. Acoust. Soc. Am.* **31**, 1250–1252.

# Cross-correlation procedures for measuring noise and distortion in AGC hearing aids

James M. Kates<sup>a)</sup>

*AudioLogic, 4870 Sterling Drive, Boulder, Colorado 80301*

(Received 16 November 1999; revised 31 January 2000; accepted 13 March 2000)

The magnitude-squared coherence function (MSC) has been used to measure noise and distortion in linear and compression hearing aids. However, the MSC will overestimate the distortion in a linear time-varying system such as a compression amplifier. The reduction in coherence caused by varying the gain in an otherwise linear system can be substantial, and can lead to large errors in estimating the distortion present in a compression hearing aid. The effects of gain changes in a linear system can be reduced by measuring the normalized system input–output cross correlation, which emphasizes the variance in the system phase response and deemphasizes the system gain fluctuations. Estimates of the total noise and distortion produced using the MSC, phase variance, and notched-noise measurement techniques are compared for additive noise, clipping distortion, and compression amplification. The MSC is found to give the most accurate results for estimating the noise and distortion in a linear time-invariant system, and the notched noise measurements are the most accurate for a compression system. The phase variance is found to give reasonable measurements for a time-varying gain as long as the system variations are slow relative to the length of the analysis data segments. © 2000 Acoustical Society of America. [S0001-4966(00)04706-8]

PACS numbers: 43.66.Ts, 43.58.Ry [RVS]

## INTRODUCTION

Cross-correlation procedures compare the system output with the system input to determine the system behavior. The coherence function (Carter *et al.*, 1973), for example, is a normalized cross-correlation measurement that determines the degree to which the system output is linearly related to the system input. The signal-to-distortion ratio (SDR), where the distortion term includes all nonlinear effects and noise in the system, can be computed from the coherence. The coherence can be computed for an arbitrary input signal, and is independent of the linear system response once the bias in the estimation procedure has been removed (Kates, 1992).

Coherence measurements are becoming important in evaluating hearing aids. Coherence measurement procedures using a broadband noise input signal have been developed for hearing aids (Preves *et al.*, 1989; Dyrlund, 1989; Preves, 1990; Kates, 1992; Fortune and Preves, 1992; Schneider and Jamieson, 1995); and a standard has evolved from this work (ANSI, 1992). The use of a broadband stimulus is essential for determining the response of an automatic gain control (AGC) hearing aid where the system gain as a function of frequency depends on both the signal amplitude and spectral distribution. The coherence is attractive for determining the nonlinear behavior of such a system because it measures all forms of distortion and not just the harmonic distortion traditionally measured (ANSI, 1996).

But there is a problem with using coherence to measure the properties of a linear time-varying system such as a compression amplifier. A linear time-varying system is one that varies over time, but in which no nonlinear distortion mechanism such as clipping or amplifier saturation is present. Olof-

sson (1992) has shown that the cross correlation of the output with the input of a linear time-varying system must always be less than the result for a system having constant gain. Thus an AGC hearing aid will have lower measured coherence than a linear time-invariant (non-AGC) instrument by virtue of the dynamic gain changes present even if no additional distortion, such as amplifier saturation, occurs. Studies using the magnitude-squared coherence (MSC) to indicate distortion levels in compression hearing aids (Dyrlund, 1989; Kates, 1993; Schneider and Jamieson, 1995) are therefore suspect because the dynamics of the compressor will reduce the measured coherence.

Coherence measurements have also been used to quantify the perceived sound quality of devices such as hearing aids. Kates and Kozma-Spytek (1994) were able to model sound quality judgments made by normal-hearing subjects listening to peak-clipped speech by using frequency-weighted coherence measurements relating the output to the input of the distortion system. An extension of that work (Kozma-Spytek *et al.*, 1996) to speech quality judgments by hearing-impaired subjects was inconclusive, however, because the frequency weights used to model the quality judgments by different subjects with similar degrees of hearing loss showed large intersubject variation.

Coherence measures the envelope and phase fluctuations of the output signal in comparison with the input. The envelope and phase fluctuations are assumed to be caused exclusively by noise and distortion. However, the envelope fluctuations may also be caused by the gain changes due to the compression amplifier in an AGC hearing aid. Thus a procedure that predominantly measures the phase fluctuations will be much less sensitive to the effects of the compression amplifier as long as the dynamic-range compression affects the gain and not the phase response of the hearing aid. An ex-

<sup>a)</sup>Electronic mail: jkates@audiologic.cirrus.com

ample of such a system would be a linear-phase digital compression hearing aid. If both the phase and gain of the hearing aid are varying in response to the incoming signal, then a noise and distortion estimation procedure using a notched-noise test signal may be the best solution. Such a procedure positions one or more notches in the excitation spectrum and then estimates the noise and distortion from the signal power that falls in each notch in comparison to the power in the surrounding peaks (Burnett, 1967; Kates, 1990).

This paper begins with a review of the procedure used to calculate the magnitude-squared coherence. The effect on the MSC of varying the system gain is then documented. The phase variance, a new distortion measure that reduces the effects of gain changes, is then introduced, and the notched-noise measurement technique is also described. The performance of the phase variance is then compared to that of the MSC and notched-noise technique for additive noise, clipping distortion, and compression amplification. The SDR computed from the phase variance is shown to be much less affected by AGC operation in a dynamic-range compressor than the SDR computed from the coherence, but neither cross-correlation procedure is as robust as the notched-noise technique in the presence of system gain fluctuations.

## I. MAGNITUDE-SQUARED COHERENCE

### A. Definition

The coherence function (Carter *et al.*, 1973) is the normalized cross-spectral density given by

$$\gamma(\omega) = S_{xy}(\omega) / [S_{xx}(\omega)S_{yy}(\omega)]^{1/2}, \quad (1)$$

where  $S_{xy}(\omega)$  is the cross-spectral density between the two stationary random processes  $x(n)$  and  $y(n)$  having autospectral densities  $S_{xx}(\omega)$  and  $S_{yy}(\omega)$ , respectively. The magnitude-squared coherence function (MSC) is then given by

$$|\gamma(\omega)|^2 = |S_{xy}(\omega)|^2 / S_{xx}(\omega)S_{yy}(\omega), \quad (2)$$

with  $0 \leq |\gamma(\omega)|^2 \leq 1$ .

When  $x(n)$  is the input to a system and  $y(n)$  is the output, the MSC represents the fraction of the output signal power that is linearly dependent on the input at each analysis frequency. The complementary fraction  $1 - |\gamma(\omega)|^2$  gives the output power at each frequency that is unrelated to the input at that frequency and thus represents the nonlinear distortion and noise. The SDR is therefore given by (Carter *et al.*, 1973)

$$\text{SDR}(\omega) = |\gamma(\omega)|^2 / [1 - |\gamma(\omega)|^2]. \quad (3)$$

When  $|\gamma(\omega)|^2 = 1$ , all of the output is linearly related to the input and the SDR approaches infinity. When  $|\gamma(\omega)|^2 = 0$ , there is no relationship between the output and the input and the SDR is zero as well.

The MSC is typically estimated using fast Fourier transform (FFT) techniques (Carter *et al.*, 1973). The data sequences  $x(n)$  and  $y(n)$  are divided into a number  $M$  of overlapping windowed data segments. The cross-spectrum and autospectra are computed for each segment using the FFT

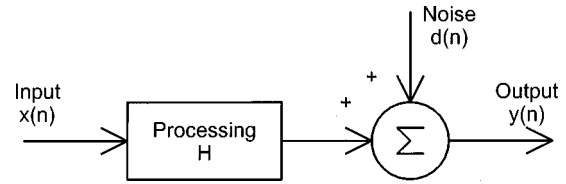


FIG. 1. Block diagram of a linear processing system with additive noise.

and are then averaged across segments, and the MSC is computed from the averages. For  $M$  data segments, the estimated MSC is given by

$$|\hat{\gamma}(\omega)|^2 = \frac{|\sum_{m=0}^{M-1} X_m(\omega) Y_m^*(\omega)|^2}{\sum_{m=0}^{M-1} |X_m(\omega)|^2 \sum_{m=0}^{M-1} |Y_m(\omega)|^2}, \quad (4)$$

where the asterisk denotes the complex conjugate, and  $X_m(\omega)$  and  $Y_m(\omega)$  are the spectra of the  $m$ th windowed data segments of  $x(n)$  and  $y(n)$ , respectively, computed using the FFT algorithm. Typically, a Hanning (von Hann) or Hamming data window and 50% overlap are used for the data segments. The estimated MSC is subject to bias effects due to the finite segment length and finite number of segments, and unbiasing techniques have been developed by Kates (1992).

### B. Time-varying system

The MSC definition assumes that the input and output signals are stationary, that is, the signal statistics do not change over time. If the system is time varying, as occurs in a compression amplifier, then the underlying stationarity assumptions are violated and the MSC is no longer valid. A simple signal processing system is shown in Fig. 1. If the transfer function  $H(\omega)$  is time varying, even for a stationary input  $x(n)$ , the estimated MSC will be misleading.

As an example, let the system transfer function switch between two gain values  $a$  and  $b$  with a duty cycle of 50%. The input signal is white Gaussian noise and there is no additive noise. The auto- and cross-spectra are then given by

$$\begin{aligned} S_{xx}(\omega) &= E[|X(\omega)|^2], \\ S_{yy}(\omega) &= \frac{1}{2}(a^2 + b^2)E[|X(\omega)|^2], \\ S_{xy}(\omega) &= \frac{1}{2}(a + b)E[|X(\omega)|^2]. \end{aligned} \quad (5)$$

The MSC for this system is then given by

$$|\gamma(\omega)|^2 = \frac{1}{2} \frac{(a + b)^2}{(a^2 + b^2)}. \quad (6)$$

If the gain is constant, as given by  $b = a$ , then the MSC = 1. However, any fluctuations in the gain reduce the MSC. The SDR, computed from the MSC using Eq. (3) and converted to dB, is plotted in Fig. 2 as a function of the ratio  $b/a$ . Reducing the ratio  $b/a$  reduces the SDR despite the fact that there is no distortion mechanism such as clipping or amplifier saturation nor any additive noise in this example. When  $b = 0$  the MSC from Eq. (6) becomes 0.5, which gives a SDR of 0 dB. Thus gain variations, as occur in a compression amplifier, will reduce the MSC and associated SDR estimated for a hearing aid. It is important to note that the rate

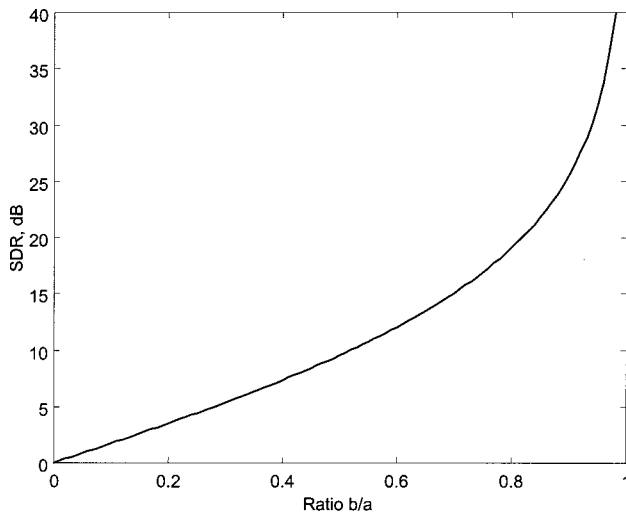


FIG. 2. Signal-to-(distortion plus noise) ratio (SDR) in dB for a linear time-varying processing system switching between gains of  $a$  and  $b$  with a 50% duty cycle.

of the gain fluctuations is immaterial in this example; slow or rapid changes in gain will give the same reduction in coherence as long as the ratio  $b/a$  is the same. In a compression hearing aid, however, the amount of distortion will depend on the interaction of the compressor attack and release times and the rate of the signal envelope fluctuations.

## II. PHASE VARIANCE

For a linear system, the phase of the system transfer function can be extracted from the input–output cross correlation, with the phase given by

$$\varphi(\omega) = \tan^{-1} \{ \text{Im} [S_{xy}(\omega)] / \text{Re} [S_{xy}(\omega)] \}. \quad (7)$$

The phase of the system transfer function can be estimated using the short-time FFT much like the procedure used to estimate the coherence. If the cross-correlation FFTs are averaged across the entire data record before the phase is computed, then the same stationarity assumptions apply for the estimated phase as applied for the estimated coherence. Normalizing the cross correlation on a segment-by-segment basis, however, gives the short-time estimates of  $\cos[\varphi_m(\omega)] + j \sin[\varphi_m(\omega)]$  for data block  $m$ . The average of these real and imaginary components will be much less sensitive to compression amplifier gain fluctuations because of the segment-by-segment normalization, and the variance of the normalized signal can be used to indicate the undesired phase fluctuations caused by noise and distortion.

Define the normalized cross correlation for data segment  $m$ :

$$P_m(\omega) = \frac{X_m(\omega)Y_m^*(\omega)}{|X_m(\omega)Y_m^*(\omega)|}. \quad (8)$$

The estimate of the normalized cross correlation then becomes

$$\rho(\omega) = \frac{1}{M} \sum_{m=0}^{M-1} P_m(\omega). \quad (9)$$

The SDR of a linear time-varying system can be estimated using the normalized cross correlation as the starting point. The phase variance describes the fluctuations in  $P_m(\omega)$  and is given by

$$\sigma^2(\omega) = \left( \frac{1}{M} \sum_{m=0}^{M-1} \{ \text{Re}^2 [P_m(\omega)] + \text{Im}^2 [P_m(\omega)] \} \right) - (\text{Re}^2 [\rho(\omega)] + \text{Im}^2 [\rho(\omega)]). \quad (10)$$

As shown in Appendix A, the SDR is approximately

$$\text{SDR}(\omega) = 2/\sigma^2(\omega). \quad (11)$$

The SDR computed using Eqs. (10) and (11) assumes that the linear time-varying system has a constant phase response even though the gain may be changing, and that the gain changes occur slowly relative to the analysis block size. These assumptions would apply, for example, to a linear-phase digital compression system having time constants long relative to the analysis window. However, compressor gain changes occurring on the same time scale as the analysis window, as discussed in Appendix B, will reduce the SDR. The SDR computed using the phase variance is then interpreted in the same way as the SDR computed using the coherence, that is, it assumes that all of the measured phase fluctuations are due to noise and distortion.

## III. NOTCHED-NOISE MEASUREMENTS

The notched-noise distortion measurement uses the procedure developed by Kates (1990). The noise test signal is speech-shaped Gaussian noise which is then convolved with a comb filter to create a series of interleaved peaks and valleys. White Gaussian noise is first passed through a three-pole Butterworth high-pass filter with a cutoff frequency of 200 Hz and a three-pole Butterworth low-pass filter with a cutoff frequency of 5000 Hz. The band-limited noise signal is then passed through a one-pole low-pass filter having a cutoff frequency of 900 Hz to approximate the long-term spectrum of speech. The comb filter has 0-dB pass-band gain and 62-dB stop-band attenuation, with valleys at 625 Hz and multiples of 625 Hz and pass-band peaks occurring between each adjacent pair of valleys. The assumed system sampling rate is 20 kHz.

The distortion measurement procedure is based on determining how much energy from the peaks of the comb-filtered noise signal spills over into the valleys. The power spectrum of the output signal is computed using the method of modified periodograms (Welch, 1967), with the data processed in blocks of 2048 samples using a Hamming window and 50% overlap. The distortion energy is computed as the mean-squared energy in the center of the valley, and the signal energy is computed as the mean-squared energy in the peaks to either side of the valley with the bias equal to the distortion energy subtracted out. The transition regions between the peaks and the valley are not used.

A modification to the notched-noise technique is to replace the speech-shaped noise with a segment of speech. The speech segment is passed through the comb filter to generate a notched speech excitation signal, and the distortion is then measured using the same procedure as for the notched-noise

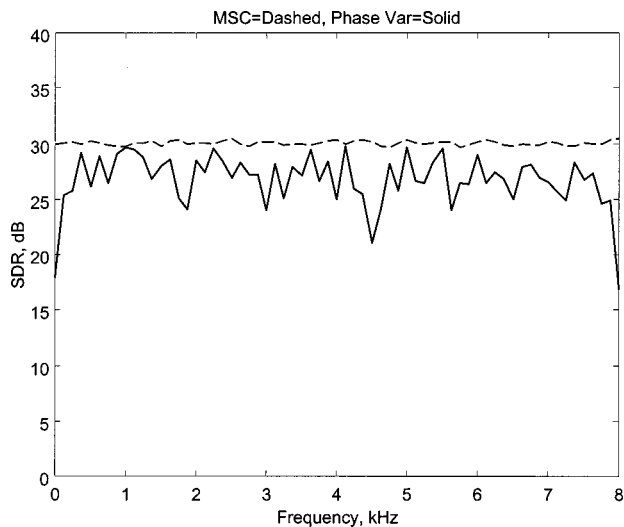


FIG. 3. The SDR in dB as a function of frequency computed from the MSC and phase variance. The input signal is white Gaussian noise, and the interference is independent white Gaussian noise at a SNR of 30 dB.

test signal. The notched speech signal preserves the envelope fluctuations of the original speech, and is thus a useful signal for testing a compression hearing aid.

To compute the AI-weighted SDR, the peak-to-valley ratios are first converted to dB. The SDR is then given by a weighted sum of the dB values corresponding to the valleys at 625–5000 Hz. The weights are derived from the Articulation-Index weights of Kryter (1962), and are tabulated by Kates (1990).

#### IV. SIMULATION RESULTS

The MSC, phase variance, and notched-noise test procedures were implemented to estimate the SDR of a corrupted signal. The simulation sampling rate was 16 kHz for the MSC and phase variance, and 20 kHz for the notched noise. The analysis procedures for the MSC and phase variance used a block size of 128 samples (8 ms) weighted with a Hamming window and a 128-point FFT for the frequency analysis unless otherwise stated. The segments had an overlap of 64 samples. The notched-noise measurements used a FFT block size of 2048 samples (102.4 ms) with an overlap of 1024 samples. Five seconds of data were processed for each measurement condition.

An example of the SDR computed from the MSC and phase variance as a function of frequency is shown in Fig. 3. The input signal is white Gaussian noise with additive uncorrelated white Gaussian noise interference at a SNR of 30 dB. The system had a flat unity gain. The SDR derived from the MSC is very well behaved under these measurement conditions, giving an accurate result at all frequencies. The SDR derived from the phase variance shows much more variance across frequency and underestimates the SNR. The greater variance in the SDR computed from the phase variance is probably due to the fact that the phase fluctuations are the only part of the signal variation being used, whereas the SDR computed from the MSC is based on both the phase and amplitude fluctuations and thus has more information available about the signal.

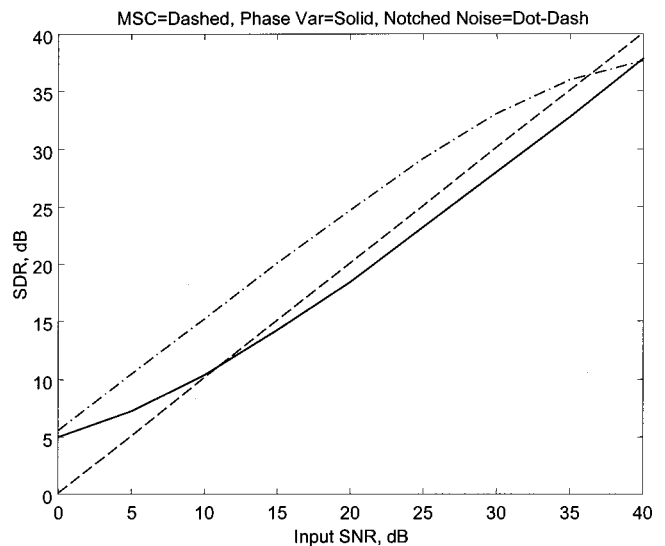


FIG. 4. The AI-weighted SDR in dB computed from the MSC, phase variance, and notched-noise techniques. The input signal for the MSC and phase variance is white Gaussian noise, and the input for the notched-noise test is comb-filtered speech-shaped Gaussian noise. The interference is independent white Gaussian noise at the indicated SNR.

The SDRs derived from the three measurement procedures are shown in Fig. 4 for additive noise. The system input for the MSC and phase variance was white Gaussian noise, while the comb-filtered speech-shaped noise was used for the notched-noise measurements. Uncorrelated white Gaussian noise interference was added at the indicated SNR. The system had flat unity gain. The logarithmic SDR values for the MSC and phase variance were averaged across frequency using the Kryter Articulation Index (AI) weights (Kryter, 1962; Kates and Kozma-Spytek, 1994) to give the AI-weighted SDR values shown in the figure. The MSC yields an essentially perfect mapping of the input SNR into the estimated SDR. The phase variance overestimates the SDR at input SNR values of less than 10 dB, and then underestimates the higher SNR values by an average of about 2 dB. The deviations of the SDR computed from the phase variance are primarily the result of the approximation used to compute the SDR from the phase variance. The notched-noise measurements give SDR estimates approximately 6 dB greater than those computed using the MSC at the lower SNR values, and then start to approach the asymptotic level determined by the comb-filter and windowed FFT sidelobes as the input SNR increases.

The SDRs computed from the three measurement procedures are shown in Fig. 5 for symmetric peak clipping distortion. The system input for the MSC and phase variance was white Gaussian noise with unit variance, and the input for the notched-noise measurements was the comb-filtered speech-shaped noise. The clipping threshold is expressed in units of the signal standard deviation; a clipping threshold of 1 corresponds to clipping the signal at its rms level. The clipping threshold was then varied over the indicated range, giving infinite clipping near the threshold of zero and essentially no clipping at the threshold of 3.5. The system had flat unity gain, and the logarithmic SDR values for the MSC and phase variance were again averaged across frequency using

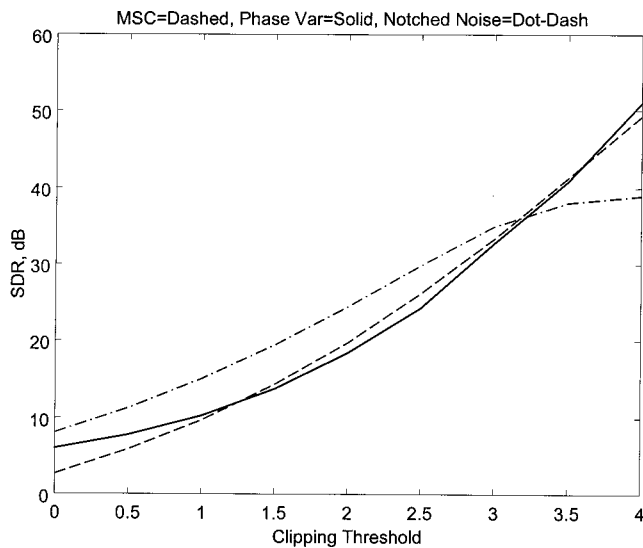


FIG. 5. The AI-weighted SDR in dB computed from the MSC, phase variance, and notched-noise techniques. The input signal for the MSC and phase variance is white Gaussian noise having a variance of 1, and the input for the notched-noise test is comb-filtered speech-shaped Gaussian noise having a variance of 1. The signal is symmetrically peak clipped at the indicated clipping threshold  $re$ : the signal rms level.

the Kryter Articulation Index (AI) weights to give the AI-weighted SDR values shown in the figure. The SDR computed from the MSC is monotonically increasing with clipping threshold, and is assumed to be an accurate representation of the distortion level. The SDR computed from the phase variance overestimates the SDR for the higher amounts of distortion generated by the lower clipping thresholds, but the SDR values above 10 dB match those of the MSC curve quite closely. The SDR computed from the notched-noise measurements again lies about 6 dB above that estimated from the MSC at low input SNR values, and then approaches the asymptotic value at high input SNR values. Note that the SDR computed from any of the three procedures is monotonic with clipping level, so rankings of the amount of distortion based on any of the SDR estimates will be accurate.

Both the additive noise and the clipping distortion experiments are measurement conditions where the MSC is expected to be accurate because the underlying signal processing system has constant gain. Results for a simulated broadband compression hearing aid for several different input signals are presented in Table I. The broadband compression

TABLE I. The AI-weighted averages of the SDR calculated from the magnitude-squared coherence (MSC) and phase variance. The simulated broadband compression hearing aid had a compression ratio of 1.5:1 with an ANSI attack time of 5 ms and an ANSI release time of 50 ms. The compression threshold was set to 40 dB below the rms signal level. The analysis block length was 128 samples.

Test signal	MSC SDR (dB)	Phase variance SDR (dB)
White Gaussian noise	35.3	40.7
Speech-shaped noise	31.2	36.7
Amplitude-modulated speech-shaped noise	18.9	35.3
Speech segment	16.3	28.2

TABLE II. The AI-weighted averages of the SDR in dB calculated using the notched-noise technique. The simulated broadband compression hearing aid had a compression ratio of 1.5:1 with an ANSI attack time of 5 ms and an ANSI release time of 50 ms. The compression threshold was set to 40 dB below the rms signal level. The analysis block length was 2048 samples.

Test signal	SDR dB	
	No compression	Compression
Notched noise	38.6	37.3
Notched speech	38.8	36.0

amplifier had a compression ratio of 1.5:1 with an ANSI attack time of 5 ms and an ANSI release time of 50 ms (ANSI, 1996). The compression threshold was set to 40 dB below the rms signal level. There was no additive noise or clipping distortion in the compression system. The amplitude modulation applied to the speech-shaped noise had a modulation index of 0.6, which gave a modulation envelope peak-to-valley ratio of 12 dB, and the modulation frequency was 6 Hz.

The tabulated results indicate that as the envelope fluctuations of the input signal increase, causing greater gain variations in the compression amplifier, the SDR computed from the MSC decreases. This illustrates the problem of using the MSC to compute the performance statistics of a linear time-varying system. Even with no external source of noise or clipping distortion, the gain fluctuations of the compression system cause a substantial reduction in the computed MSC and SDR. This would lead one to conclude that the compression system had large amounts of distortion when in fact what is present is only artifacts of the measurement procedure. The SDR computed from the phase variance, on the other hand, shows much less sensitivity to the gain fluctuations, although some sensitivity is still present, and would be a more accurate indicator of the residual distortion resulting from the compression amplification.

Compression results for the SDR computed using the notched-noise technique are presented in Table II. The simulated compression hearing aid is the same as used for the results of Table I. The two test signals are the notched-noise test signal defined by Kates (1990) and the speech segment passed through the comb filter to give a notched speech signal. The SDR computed from the notched-noise technique is much less affected by the compressor than that computed from the MSC or the phase variance. The SDR estimated using the notched speech signal is 36.0 dB, as compared to 28.2 dB for the phase variance and 16.3 dB for the MSC using the speech signal. The SDR computed from the notched-noise technique appears to be much more robust in the presence of gain fluctuations than that computed from the correlation-based procedures.

The SDR computed from the phase variance is sensitive to the analysis block size. The phase variance computation assumes that the system gain is relatively constant over the block duration. If the gain changes more rapidly than the block rate, the SDR computed from the phase variance will be reduced. This effect is illustrated in Fig. 6 for the speech segment as the MSC and phase variance analysis block size is increased from 32 samples (2 ms) to 4096 samples (256

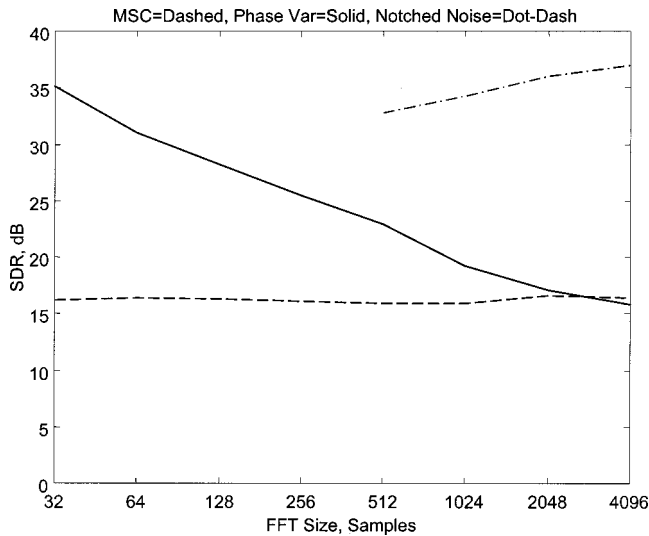


FIG. 6. The AI-weighted SDR in dB computed from the MSC, phase variance, and the notched-noise techniques for a simulated broadband compression hearing aid having a compression ratio of 1.5:1 with an ANSI attack time of 5 ms and an ANSI release time of 50 ms. The compression threshold was 40 dB below the rms signal level. The input signal for the MSC and phase variance is the speech segment, and the input for the notched-noise test is the speech segment passed through the notched-noise comb filter.

ms) at the 16-kHz sampling rate. The attack time of the compressor is 5 ms, which is comparable to the analysis block size used in the SDR calculations. The SDR computed from the MSC remains constant at about 16.2 dB independent of the block size. However, the SDR computed from the phase variance is inversely proportional to the block size, decreasing approximately 3 dB for each doubling of the block length. Thus one should use the shortest block size that is compatible with the required frequency resolution when computing the SDR from the phase variance.

The SDR computed using the notched speech approach, however, is essentially immune to the analysis block size. A minimum block size of 512 samples is needed to give adequate frequency resolution for estimating the power in the signal valleys. As the block size increases, the SDR computed from the notched speech signal increases due to reduced interference between the FFT sidelobes and the residual power level estimated in the valleys.

## V. CONCLUSIONS

There are many techniques that can be used to estimate the noise and distortion in a hearing aid. The unbiased MSC has the advantages of providing an accurate noise and distortion estimate for an arbitrary test signal, but the results are invalid for a compression instrument because of the system gain fluctuations. On the other hand, notched noise measurements can give a valid noise and distortion estimate for any system, but require the use of a special test signal or notch-filtered input. The phase variance thus occupies a useful middle position, giving reasonable measurements for a time-varying gain and an arbitrary test signal as long as the system variations are slow relative to the length of the analysis data segments. For rapid gain variations, however, the phase variance will lead to erroneous results.

For those situations where the input and output signals are known to be stationary, the MSC is the preferred estimation procedure. The SDR derived from the MSC accurately measures noise and distortion levels in linear time-invariant systems. However, there are many situations, such as a compression amplifier in a hearing aid, where the system gain is varying with time. For these situations the phase variance is preferable to the MSC because it is much less sensitive to the effects of gain fluctuations. Even though the phase variance is somewhat less accurate in estimating the noise and distortion, the net error will be much smaller than for the MSC in a compression system. However, the phase variance will also be sensitive to the changes in a linear time-varying system if the phase varies in addition to the gain or if substantial gain variation occurs within the analysis data segments.

The notched-noise technique has the best immunity to the effects of compressor gain fluctuations. The SDR estimated using the notched-noise technique shows an approximate 6-dB bias in comparison with that estimated using the MSC for stationary test conditions. However, the notched-noise measurements were essentially unaffected by the compressor gain changes induced by the speech signal. One could therefore calibrate the notched-noise measurements under stationary conditions using the MSC as a reference, and then proceed to use the notched-noise approach for systems suspected of having gain fluctuations.

The detection of noise and distortion for an arbitrary test signal is based on measuring amplitude and phase perturbations in the system output that were not present in the input. The MSC measurement is based on the assumption that all perturbations in the output are due to noise and distortion, while the phase variance is based on the assumption that the phase perturbations are due to noise and distortion and ignores slow amplitude fluctuations. The notched-noise measurement works on a different principle; it is based on the assumption that the power that appears in the spectral notches is due to noise and distortion, and it ignores both amplitude and phase perturbations in the bands carrying the input signal power. Thus one needs to choose the measurement procedure based on knowledge of the underlying system behavior to ensure that the measurements are meaningful rather than corrupted by artifacts.

## ACKNOWLEDGMENTS

The work presented in this article was supported in part by GN ReSound. The author would like to thank Dave Preves for comments on a draft version of this paper.

## APPENDIX A: COMPUTING THE SDR FROM THE PHASE VARIANCE

Consider the linear system of Fig. 1 with additive noise. The normalized cross correlation estimated for this system using Eq. (9) is

$$\rho(\omega) = \frac{1}{M} \sum_{m=0}^{M-1} \frac{H^*(\omega)|X(\omega)|^2 + D^*(\omega)X(\omega)}{|H^*(\omega)|X(\omega)|^2 + D^*(\omega)X(\omega)}, \quad (\text{A1})$$

where  $D(\omega)$  is the Fourier transform of the noise signal shown in Fig. 1. Because the normalization is performed for

ach data segment, Eq. (A1) can be rewritten as

$$\rho(\omega) = \frac{1}{M} \sum_{m=0}^{M-1} \{\cos[\varphi_m(\omega)] + j \times \sin[\varphi_m(\omega)]\}. \quad (\text{A2})$$

If the phase response of the system transfer function remains constant, as would occur in a linear-phase digital system, then the system phase for each data segment can be represented as

$$\varphi_m(\omega) = \varphi(\omega) + \delta_m(\omega), \quad (\text{A3})$$

where  $\varphi(\omega)$  is the constant linear system phase shift and  $\delta_m(\omega)$  is the phase fluctuation caused by the additive noise. The normalized cross correlation of Eq. (A2) then becomes

$$\rho(\omega) = \frac{1}{M} \sum_{m=0}^{M-1} \{\cos[\varphi(\omega) + \delta_m(\omega)] + j \times \sin[\varphi(\omega) + \delta_m(\omega)]\}. \quad (\text{A4})$$

The variance of the normalized cross correlation given in Eq. (10) then becomes

$$\cos[\delta_m(\omega)] = \frac{\text{Re}[1 + D^*(\omega)/X^*(\omega)]}{\{\text{Re}^2[1 + D^*(\omega)/X^*(\omega)] + \text{Im}^2[1 + D^*(\omega)/X^*(\omega)]\}^{1/2}}. \quad (\text{A7})$$

Substituting Eq. (A7) into Eq. (A6) and approximating  $(1+x)^{-1/2}$  as  $1-x/2$  leads to

$$\sigma^2(\omega) = \frac{1}{M} \sum_{m=0}^{M-1} 2 \left\{ \text{Re}^2 \left[ \frac{D^*(\omega)}{X^*(\omega)} \right] + \frac{1}{2} \left| \frac{D^*(\omega)}{X^*(\omega)} \right|^2 \right\}. \quad (\text{A8})$$

Assuming that half of the noise power is in the real part and half of the power in the imaginary part then yields

$$\text{SDR}(\omega) = 2/\sigma^2(\omega). \quad (\text{A9})$$

Equation (A9) gives the approximate solution for the SDR computed from the phase variance.

## APPENDIX B: EFFECTS OF GAIN CHANGES ON THE PHASE VARIANCE

The phase variance can be affected by changes in the system gain. The nature of these effects can be quite complicated due to the interaction of the compressor attack and release time constants and the rate of the signal envelope modulations. A general explanation of the effect of gain fluctuations on the phase variance is given in this appendix.

Assume that the phase variance is being measured using data segments of length  $L$ . Further assume that the compressor gain goes from 1 to  $1+g$  and back to 1 over the duration of the segment, with the short-term gain function given by

$$g(n) = 1 + \frac{g}{2} [1 - \cos(2\pi n/L)]. \quad (\text{B1})$$

The  $L$ -point FFT of the gain,  $G(\omega)$ , consists of a peak at  $\omega=0$  having magnitude  $1+g/2$  flanked by sidelobes having

$$\sigma^2(\omega) = \frac{1}{M} \sum_{m=0}^{M-1} [\{\cos[\varphi(\omega) + \delta_m(\omega)] - \cos[\varphi(\omega)]\}^2 + \{\sin[\varphi(\omega) + \delta_m(\omega)] - \sin[\varphi(\omega)]\}^2]. \quad (\text{A5})$$

Evaluating the summation and simplifying the trigonometric terms leads to

$$\sigma^2(\omega) = \frac{1}{M} \sum_{m=0}^{M-1} 2\{1 - \cos[\delta_m(\omega)]\}. \quad (\text{A6})$$

The variance of the normalized cross correlation thus depends only on the phase shift from segment to segment and is independent of the linear time-invariant system transfer function or of the gain fluctuations in a linear time-varying system having a constant phase response.

The variance of the normalized cross-correlation depends on the amplitudes of the signal and additive noise. Since the linear system transfer function is immaterial as long as it has constant phase with respect to time, it can be replaced by 1. This substitution yields

magnitude  $g/2$ . The system output spectrum is then given by the frequency-domain convolution of the input spectrum with the gain spectrum:

$$Y(\omega) = X(\omega) * G(\omega). \quad (\text{B2})$$

The effect of the variation in the gain function is thus a smoothing in the spectral domain. Changes in the amplitude of the gain function will change the amount of spectral smoothing, and changes in the rate of the gain increase or decrease will change the width of the spectral smoothing function.

The spectral smoothing caused by the short-term gain changes will change both the amplitude and phase of the output signal. The phase variance calculation of Eq. (9) compensates for the amplitude changes but not for the phase changes. Thus the system phase shift will change from analysis block to analysis block as the spectral smoothing adjusts to the changes in the short-term compressor gain. The net result is an increase in the phase variance due to short-term gain changes.

ANSI S3.42-1992. (1992). "American National Standard Testing Hearing Aids with a Broad-Band Noise Signal" (American National Standards Institute, New York).

ANSI S3.22-1996. (1996). "American National Standard Specification of Hearing Aid Characteristics" (American National Standards Institute, New York).

Burnett, E. D. (1967). "A new method for the measurement of nonlinear distortion using a random noise test signal," Bull. Prosthet. Res. **10**, 76-92.

Carter, G. C., Knapp, C. H., and Nuttall, A. H. (1973). "Estimation of the



- magnitude-squared coherence function via overlapped fast Fourier transform processing," *IEEE Trans. Audio Electroacoust.* **21**, 337–344.
- Dyrlund, O. (1989). "Characterization of non-linear distortion in hearing aids using coherence: A pilot study," *Scand. Audiol.* **18**, 143–148.
- Fortune, T. W., and Preves, D. A. (1992). "Hearing aid saturation and aided loudness discomfort," *J. Speech Hear. Res.* **35**, 175–185.
- Kates, J. M. (1990). "A test suite for hearing aid evaluation," *J. Rehabil. Res. Dev.* **27**, 255–278.
- Kates, J. M. (1992). "On using coherence to measure distortion in hearing aids," *J. Acoust. Soc. Am.* **91**, 2236–2244.
- Kates, J. M. (1993). "Hearing aid design criteria," *J. Speech-Lang. Path. and Audiology Monograph Supplement* **1**, 15–23.
- Kates, J. M., and Kozma-Spytek, L. (1994). "Quality ratings for frequency-shaped peak-clipped speech," *J. Acoust. Soc. Am.* **95**, 3586–3594.
- Kozma-Spytek, L., Kates, J. M., and Revoile, S. (1996). "Quality ratings for frequency-shaped peak-clipped speech: Results for listeners with hearing loss," *J. Speech Hear. Res.* **39**, 1115–1123.
- Kryter, K. D. (1962). "Methods for the calculation and use of the articulation index," *J. Acoust. Soc. Am.* **34**, 1689–1697.
- Olofsson, A. (1992). "Measurements on AGC hearing instruments by use of broad-band test signals and a psychoacoustical model," *Karolinska Institutet, Teknisk Audiologi, Report TA126*.
- Preves, D. A. (1990). "Expressing hearing aid noise and distortion with coherence measurements," *ASHA* **32**, 56–59.
- Preves, D. A., Beck, L. B., Burnett, E. D., and Teder, H. (1989). "Input stimuli for obtaining frequency responses of automatic gain control hearing aids," *J. Speech Hear. Res.* **32**, 189–194.
- Schneider, T., and Jamieson, D. G. (1995). "Using maximum length sequence coherence for broadband distortion measurements in hearing aids," *J. Acoust. Soc. Am.* **97**, 2282–2292.
- Welch, P. D. (1967). "The use of the fast Fourier transform for the estimation of power spectra: A method based on short, modified periodograms," *IEEE Trans. Audio Electroacoust.* **AU-15**, 70–73.

# Lexical boundary error analysis in hypokinetic and ataxic dysarthria

Julie M. Liss<sup>a)</sup> and Stephanie M. Spitzer

*Motor Speech Disorders Laboratory, Arizona State University, Box 871908, Tempe, Arizona 85281*

John N. Caviness, Charles Adler, and Brian W. Edwards<sup>b)</sup>

*Department of Neurology, Mayo Clinic—Scottsdale, 13400 Shea Boulevard, Scottsdale, Arizona 85259*

(Received 21 April 1999; revised 15 October 1999; accepted 1 March 2000)

This investigation is the second in a series to examine a potential source of reduced intelligibility in dysarthric speech, namely the mismatch between listeners' perceptual strategies and the acoustic information available in the dysarthric speech signal. Lexical boundary error (LBE) analysis was conducted on listener transcripts from phrases produced by speakers with hypokinetic dysarthria, ataxic dysarthria, and normal controls. By design, the hypokinetic and ataxic dysarthric tapes elicited similar intelligibility (words-correct) scores. However, they elicited different numbers and patterns of lexical boundary errors. The nature of the error pattern differences can be traced to the listeners' use of available syllabic strength information to segment the acoustic stream. Specifically, although both dysarthric speech samples elicited numerous lexical boundary errors, those for the hypokinetic speech generally conformed to predictions offered from studies of degraded normal speech. Those for the ataxic speech did not conform strongly to such predictions. It appears that the prosodic deficits of the ataxic speech (tendency toward syllabic isochrony, excessive loudness variation, and reduced vowel working space consequent to reductions in vowel strength) posed more of a problem for listeners than did the prosodic deficits of the hypokinetic speech (rapid rate, monotony, reduced vowel working space). © 2000 Acoustical Society of America. [S0001-4966(00)02006-3]

PACS numbers: 43.70.Dn, 43.70.Fq, 43.71.Gv, 43.71.Bp [DOS]

## INTRODUCTION

Though the construct of speech intelligibility is fundamental to both clinical and research endeavors in dysarthria, surprisingly little is known about the perceptual processes that underlie it. For example, it is not known which perceptual strategies listeners invoke to decipher the degraded acoustic signal of dysarthria, or the extent to which speech signal characteristics might influence the choice or the effectiveness of these perceptual strategies. Answers to these questions will bear on both theories of speech perception and the clinical use of speech intelligibility measures.

This paper is the second in a series that examines the interface between the dysarthric speech signal characteristics and the listener's response to that signal in an effort to identify a source for reductions in intelligibility. To explore this interface, it was necessary to adopt a paradigm that accommodated both listener and speaker variables in a meaningful and interpretable way (see Lindblom, 1990; see also Connolly, 1986; Weismer and Martin, 1992). Toward this end, we chose to examine the perceptual task of segmenting the continuous acoustic stream into its component words. Lexical segmentation was selected for several reasons, including the primacy of the task for lexical access, the availability of a working model for perceptual strategies to accomplish the task, and the ability of the working model to accommodate dysarthric speech. Each of these is addressed in turn.

By most theoretical accounts, identification of word boundaries in connected speech triggers lexical search (e.g., Cole and Jakimik, 1980; Forster, 1989; Goldinger *et al.*, 1989; Gow and Gordon, 1995; Marslen-Wilson, 1989). If speech intelligibility is defined as a listener's success in deciphering the words spoken, word-boundary identification is an obvious and critical step. A large number of investigations have examined acoustic cues that may be involved in the prelexical processing of speech input (Cutler and Butterfield, 1990, 1991, 1992; Cutler and Norris, 1988; Gaygen and Luce, 1998; Grosjean and Gee, 1987; Gow and Gordon, 1995; Klatt, 1980; Lehiste, 1972; Mattys and Samuel, 1997; Nakatani and Schaffer, 1978; Quené, 1992; Vitevitch and Luce, 1998). This research has examined both segmental and suprasegmental cues produced by normal speakers that may be involved in the signaling of word boundaries to listeners.

One contemporary view is that listeners rely on utterance prosody—in particular, the juxtaposition of strong and weak syllables—to guide lexical segmentation decisions [metrical segmentation strategy hypothesis (MSS); Cutler and Norris, 1998; Cutler and Butterfield, 1992]. According to this model, strong syllables are those that contain full vowels and that may or may not receive prosodic stress. Weak syllables contain reduced vowels and do not receive prosodic stress (Cutler and Butterfield, 1990, 1991; Cutler and Carter, 1987; Fear *et al.*, 1995; Smith *et al.*, 1989). Central to the MSS hypothesis is the assumption that, in English, segmentation of the speech signal is activated by the occurrence of a strong syllable. This hypothesis is supported by research that has examined patterns of naturally occurring “slips of the

<sup>a)</sup>Electronic mail: julie.liss@asu.edu

<sup>b)</sup>Deceased.

ear,” and by lexical boundary errors elicited by the presentation of speech at very low listening levels, just above threshold (Cutler and Butterfield, 1992). Specifically, when listeners recognize strong syllables as word onsets, they will be more likely to erroneously insert lexical boundaries before strong than before weak syllables. They also will be more likely to erroneously delete lexical boundaries before weak than before strong syllables. Although this model does not claim to completely explain the cognitive processes involved in lexical segmentation, it does provide a theoretical framework upon which to base hypotheses and interpret results.

Because the MSS hypothesis emphasizes prosodic contrastivity, it is an attractive model for the investigation of intelligibility in dysarthria. All varieties of dysarthria affect the prosody of connected speech in one way or another (Darley *et al.*, 1969; Duffy, 1995). If prosodic cues play a crucial role in signaling word boundary location in normal speech perception, their disturbance in dysarthric speech likely has a negative impact on lexical segmentation. By examining the precise nature of the prosodic disturbances relative to the patterns of errors they elicit, we obtain information about the source of intelligibility decrement. This information is then dysarthria- (or, more accurately, speech pattern-) specific.

In our previous study, we examined perceptual strategies and hypokinetic dysarthric speech within the context of the MSS hypothesis (Liss *et al.*, 1998). Hypokinetic dysarthria was chosen as an entry point for this line of research because it was believed that the cardinal speech features, by their very nature, serve to diminish syllabic strength contrastivity. Although the precise acoustic correlates of syllabic strength have yet to be determined, there is evidence that they lie in the domains of relative syllable duration, intersyllabic pitch and loudness differences, and vowel quality or vowel strength (Cutler and Butterfield, 1990, 1991, 1992; Fear *et al.*, 1995; Halle and Keyser, 1971; Klatt, 1980; Lehiste, 1972; Nakatani and Schaffer, 1978; Quené and Koster, 1998; van Ooijen *et al.*, 1997). Thus, the hypokinetic dysarthric speech features of rapid speaking rate, a tendency toward monotony and monoloudness, and phoneme imprecision (Ackermann and Ziegler, 1991; Adams, 1991; Darley *et al.*, 1969; Duffy, 1995; Forrest *et al.*, 1989; Logemann and Fisher, 1981; Ludlow and Bassich, 1984; Ramig, 1992; Weismer, 1984, 1991) should reduce syllabic strength contrastivity.

We expected to see that, if listeners rely on syllabic strength contrasts to identify word boundaries in hypokinetic speech, the LBE patterns should reveal reduced effectiveness of this strategy. Instead, we found that the general patterns of LBEs were identical to those predicted from studies of degraded normal speech (Cutler, 1993; Cutler and Butterfield, 1992; Cutler and Norris, 1988). Even though the phrases produced by these speakers contained reduced syllabic strength contrasts according to both perceptual and acoustic evaluation, the information was apparently sufficiently robust to be utilized by the listeners. However, two important points were of note. First, the strength of adherence to the predicted pattern of LBEs across all speech samples was less than that reported in the literature for normal degraded

speech. Second, speech samples with the most severely degraded strong–weak contrasts (according to perceptual and acoustic indices) elicited the lowest strength of adherence values. Thus, this investigation provided evidence that listeners rely on acoustic information about syllabic strength to locate word boundaries in hypokinetic dysarthric speech, but that listeners have difficulty applying the strategy as syllabic strength information is increasingly degraded.

Having established that the severity of syllabic strength degradation is a source of reduced intelligibility in hypokinetic dysarthria, the present study posed the following questions: Does the *form* of the dysarthria have a definable effect on listeners’ abilities to apply the Metrical Segmentation Strategy? For comparison, we selected ataxic dysarthria because it differs antithetically from hypokinetic dysarthria in its pattern of prosodic deficit. Ataxic dysarthria is characterized by a slow speaking rate with excess and equal stress, and excessive loudness variation, with irregular articulatory breakdown and articulatory impression (Ackermann and Hertrich, 1994; Darley *et al.*, 1969; Duffy, 1995; Kent *et al.*, 1979). Our goal was to select two groups of speakers, hypokinetic and ataxic dysarthrics, whose speech samples did not differ in terms of intelligibility scores, but whose speech samples exhibited very specific and different forms of prosodic deficit. In this way, when listeners mis-segment the acoustic stream, the perceptual errors could be interpreted relative to specific production characteristics rather than to fundamental differences in intelligibility.

Thus, both dysarthric groups exhibited reductions in syllabic strength contrastivity but for different reasons. The question was whether these two speech patterns would show evidence of posing different challenges to the listeners’ task of perceptual segmentation. If this is found to be the case, we will have identified different sources of intelligibility reduction for these two speech pattern types.

Based on our previous study (Liss *et al.*, 1998), we expected that listeners would have difficulty identifying lexical boundaries in the hypokinetic speech, but that they would make use of the available syllabic strength information to venture guesses. This would be apparent in two ways: the presence of a large number of LBEs, and a LBE pattern that generally conformed to that predicted by the MSS hypothesis. We also expected that because syllabic strength contrastivity is reduced in these phrases, the strength of adherence to the predicted pattern would be less than that reported in the literature for degraded normal speech. If these results were found, they would replicate those of our previous investigation and strengthen the validity of those original findings.

With regard to the ataxic speech, we expected that listeners would have difficulty identifying lexical boundaries in the phrases—again, because of the presumed decrement in syllabic strength contrasts. This would manifest as a large number of LBEs. However, it was not known the extent to which the nature of the prosodic deficit would affect the systematic application of the metrical segmentation strategy, nor whether listeners would show evidence of using syllabic strength contrasts to parse the phrases. Within our sample of phrases, the opportunity to produce lexical boundary dele-

tions and insertions before strong and weak syllables was exactly opposite the MSS prediction pattern (see Liss *et al.*, 1998; Table III). Thus, if the LBE pattern conformed to the predicted pattern and did so strongly, it could be surmised that the syllabic strength information was sufficient for the effective use of this perceptual strategy. If the LBE pattern conformed weakly, it would signify a mismatch between the acoustic information and the perceptual strategy. In other words, it would indicate that listeners are relying on inferior or insufficient acoustic information to conduct the task of lexical segmentation. If the LBE pattern matched the opportunity for certain errors to occur (i.e., more deletion than insertion errors; more insertions before weak than strong syllables; more deletions before strong than weak syllables), it could be interpreted either as the failure of the perceptual strategy and/or use of an alternate perceptual strategy that did not capitalize on syllabic strength information.

## I. METHOD

### A. Listeners

The 60 listeners were 30 men and 30 women who ranged in age from 18–44 years. Most were undergraduate students at Arizona State University who were compensated for their participation in this study. All listeners self-reported normal hearing, were native speakers of Standard American English, and reported having little or no experience listening to dysarthric speech. The listeners were quasi-randomly assigned to one of the listening groups (Control, Hypokinetic,<sup>1</sup> and Ataxic) such that each group contained ten men and ten women.

### B. Speech stimuli

Speech stimuli consisted of three audiotapes of 60 phrases produced by three groups of speakers: six speakers with hypokinetic dysarthria, six with ataxic dysarthria, and six neurologically normal control speakers. It was critical to the success of this investigation that two conditions were met in terms of the two dysarthria tapes. First, LBEs must be interpretable relative to syllabic strength. This required careful construction of the stimulus phrases themselves. Second, the two dysarthria tapes must be of equivalent intelligibility (as measured by words-correct scores), and all of the phrases on each tape must be representative of the operational definitions of the respective dysarthria. This allowed differences in the dependent variables to be interpreted relative to differences in speech production characteristics as they relate to syllabic strength contrasts. Toward this end, speakers and phrase tokens were selected through a series of steps.

The 60 phrases, modeled after Cutler and Butterfield (1992), were designed to permit the interpretation of LBEs. The phrases themselves were of low interword predictability to reduce the contribution of semantic information to word perception. They consisted of six syllables that alternated in phrasal stress patterns. Half of the phrases alternated strong–weak (SWSWSW), and the other half alternated weak–strong (WSWSWS). The majority of the strong and weak syllables contained full and reduced vowels, respectively. The phrases ranged in length from three to five words and no

word contained more than two syllables. None of the words in the phrases was repeated except articles and auxiliary verbs; all English phonemes except /zh/ were represented.<sup>2</sup>

To meet the requirements of similar within-group speech characteristics and equivalent between-group phrase intelligibility, more speakers were recruited and more phrases were recorded than ultimately were used. Forty-seven hypokinetic and 48 ataxic dysarthric speakers were identified by their neurologists and speech pathologists as potential subjects for this investigation. Following an initial telephone conversation with the first author, 13 hypokinetic and 12 ataxic speakers were invited to participate based on their moderate to severe levels of intelligibility deficit and predominant speech characteristics. Our operational definition of hypokinetic dysarthria was as follows: perceptually rapid speaking rate with monopitch and monoloudness; little use of variation in pitch or loudness to achieve differential syllabic stress; imprecise articulation that gives the impression of a blurring of phonemes and syllables; and a breathy and perhaps hoarse/harsh voice. Ataxic dysarthria was defined as perceptually slow speaking rate with a tendency toward equal and even syllable duration (scanning speech); excessive loudness variation; and irregular articulatory breakdown. Of the 25 speakers who provided samples, five hypokinetic and six ataxic speakers ultimately were not used because of additional speech characteristics either not present or not noted during the initial telephone screening. These characteristics included the presence of a distinctive regional accent, a pervasive vocal tremor and oral dyskinesia, insufficient impairment of intelligibility, and the absence of one or more components of our operational definitions of dysarthria subtype. Thus, speech samples from eight hypokinetic dysarthrics, six ataxic dysarthrics, and eight neurologically normal controls were regarded as potential candidates for construction of the listening tapes.

Speech samples were collected during a single hour-long session with each speaker. The details of the speech protocol are reported in Liss *et al.* (1998). Briefly, each speaker produced words and sentences for the *Assessment of Intelligibility of Dysarthric Speech (AIDS)*: Yorkston and Beukelman, 1981), a reading passage, vowel productions, several minutes of spontaneous speech, and the set of stimulus phrases. Each speaker typically produced three or four iterations of each stimulus phrase during the course of speech sample collection. The first token, which contained no word omissions, substitutions, dysfluences, or interword pauses, and which most closely represented our operational definitions, was selected as the experimental token. The 60 phrases per speaker were low-pass filtered at 10 kHz, digitized at a 22-kHz sampling rate, and stored in a computer file using CSpeech Laboratory Automation System (Milenkovic and Read, 1992). Prior to downloading onto DAT audiocassettes for use in the perceptual experiment, the files were subjected to a customized gain program (MATLAB) to equilibrate mean rms energy across the phrases.

The next step was to identify the 18 speakers whose phrases would comprise each of the three listening tapes, and to select the 60 phrase productions (ten per speaker) for each tape. The first and second authors used information provided

TABLE I. Duration data for speaker groups.

Speaker group	Median phrase duration <sup>a</sup>	Median strong syllable duration <sup>b</sup>	Median weak syllable duration <sup>c</sup>	Median strong-weak vowel duration ratio <sup>d</sup>
Control	1681.5	339.0	171.5	1.88
Hypokinetic	1192.5	231.5	121.0	1.86
Ataxic	2253.5	462.0	272.0	1.48

<sup>a</sup> $H(2) = 135.9$ ,  $P < 0.0001$ , all comparisons significant  $P < 0.05$ .

<sup>b</sup> $H(2) = 193.4$ ,  $P < 0.0001$ , all comparisons significant  $P < 0.05$ .

<sup>c</sup> $H(2) = 141.2$ ,  $P < 0.0001$ , all comparisons significant  $P < 0.05$ .

<sup>d</sup> $H(2) = 7.98$ ,  $P = 0.0185$ , control and hypokinetic versus ataxic significant  $P < 0.05$ . The second and third are based on syllable duration measures. The ratios presented in the final column are based on intraphrase strong-weak vowel duration contrasts, not on values in the previous data columns.

by eight certified speech language pathologists,<sup>3</sup> and the results of the previous study on hypokinetic dysarthria (Liss *et al.*, 1998) to accomplish this task. First, the six hypokinetic speakers whose intelligibility levels most closely matched the six ataxic speech samples were identified. Second, ten phrases from each speaker were selected such that all 60 phrases were represented for each dysarthria tape. Although the goal of phrase selection was only to create hypokinetic and ataxic tapes with similar overall intelligibility distributions, it was possible to select phrase pairs with similar intelligibility levels for the vast majority of phrases. Thus, two dysarthria tapes were constructed—and there was every indication from the preliminary steps that the two would elicit similar words-correct scores from the listeners in the present study.<sup>4</sup>

Finally, three men and three women whose age range was most similar to that of the 12 dysarthric speakers were selected from among the eight control speakers. Because the intelligibility of their phrases ranged from 98%–100%, phrase selection for the listening tape was straightforward. Ten phrases were selected from each of the six speakers in the quasi-random fashion, such that no speaker provided consecutive phrases.

The three audiotapes each contained one production of the 60 phrases, ten phrases per each of the six speakers in each group. Phrases were preceded by the phrase number (1–60) spoken by a neurologically normal female, and followed with a 12-s interstimulus silent interval. The subjective recording quality was judged to be high and signal intensity consistent across speakers and phrases.

### 1. Perceptual-acoustic characterization of phrases

As noted earlier, the precise acoustic correlates of syllabic strength contrastivity have yet to be determined but are thought to lie in the domains of vowel quality, relative syllable duration, and pitch and loudness contours. Perceptually, all of the phrases produced by the dysarthric speakers were characterized by reductions in syllabic contrastivity, as compared with those produced by the control group. However, the source of this reduced contrastivity differed between the two dysarthria subtypes, as per the operational definitions herein. Syllabic contrastivity reductions in the hypokinetic phrases were thought to arise from the perceptually rapid rate, blurring of syllabic boundaries, and monotony.

The ataxic phrases were characterized by a tendency toward equal and even syllable duration, and a labored, slow rate of speech. The acoustic correlates of these perceived reductions were examined for all of the stimulus phrases. Intraphrase measures included duration (phrase, syllable, and vowel), and fundamental frequency and amplitude variation. Intra-group measures of strong vowel formant frequencies ( $F1$ – $F2$ ) were plotted to determine the “outer limits” of the vowel working space for each set of phrases. By inference, smaller working spaces may reflect reduced capacity for strong-weak distinctions because the strong vowels are relatively reduced.<sup>5</sup>

### 2. Duration

Phrase duration was obtained during the initial editing of the phrases by placing cursors on the first and last acoustic evidence of phonemes on the spectrographic display. This included the first or last glottal pulse in the case of initial or final voiced phonemes, respectively; the beginning or end of noise energy in the case of initial or final fricatives; and the beginning or end of the burst release in the case of initial or final stop consonants. One-hundred milliseconds of silence was then appended to the beginning and end of each phrase to reduce onset–offset effects and the entire screen was saved as a digital file for all subsequent acoustic analysis.

Syllable and vowel durations were measured using the cursor function of CSpeech on a spectrographic display. To promote uniform measurement criteria and high measurement reliability, a hardcopy of each phrase as produced by the control speaker was segmented according to traditional phoneme landmarks (Crystal and House, 1998a, b; Weismer, 1984), and operational definitions of each segment were created by the first author. The hardcopy and definitions were used as a guide for the on-screen segmentation of the other control and dysarthria phrases. The first author made all measurements, and approximately 20% of the dysarthric phrases were resampled in a quasi-random fashion by the first author and one other judge for reliability purposes. Reliability measures for the control phrases were not undertaken because of the use of the hardcopy reference. Data for reliability of vowel and syllable measures were collapsed. The mean interjudge difference was 44 ms (range, 0–123 ms), and intrajudge differences ranged from 0–112 ms (mean, 31 ms).

TABLE II. Fundamental frequency and amplitude variation for the speaker groups.

Speaker group	Intraphrase $F_0$ coefficient of variation <sup>a</sup>	$F_0$ range across phrases (Hz)	Intraphrase amplitude coefficient of variation <sup>b</sup>
Control	0.1365	125	0.136
Hypokinetic	0.0497	75	0.134
Ataxic	0.1245	131	0.183

<sup>a</sup> $H(2)=65.1$ ;  $P<0.0001$ ; hypokinetic significantly different from ataxic and control at  $P<0.05$ .

<sup>b</sup> $H(2)=32.6$ ;  $P<0.001$ ; ataxic significantly different from hypokinetic and control at  $P<0.05$ .

The first three data columns of Table I contain the median phrase, strong syllable, and weak syllable durations for each of the speaker groups. In all cases, the duration measures of the hypokinetic phrases were significantly shorter and the ataxic phrases significantly longer than those of the other two groups.

An index of local durational contrastivity between vowels in strong and weak syllables was computed. For each phrase, the duration of each strong vowel was divided by the duration of its adjacent weak vowel. To avoid the artifact of final syllable lengthening (Oller, 1973; Klatt, 1975), vowels located in final syllables were not included. Therefore, each phrase yielded four S/W ratios and a mean. The median values for each group are presented in the final data column of Table I. A Kruskal–Wallis one way ANOVA on ranks and *post hoc* Student–Newman–Keuls revealed that the ratio of the ataxic phrases was significantly lower than that of the other two groups. Although the ratio was slightly lower for the hypokinetic than control phrases, this did not reach significance.

To summarize, the phrase, syllable, and vowel duration measures support the perception of rapid hypokinetic and slow ataxic speaking rate relative to the controls. The perception of equal and even syllable duration in the ataxic phrases as compared to those of the control and hypokinetic phrases was supported by the strong-to-weak vowel duration ratios.

### 3. Vowel quality

First and second formant frequencies were measured at the temporal midpoints of seven occurrences of the vowels /i/, /ae/, /a/, and /u/ embedded in strong syllables (see Liss *et al.*, 1998 for word contexts). Measurements were made using both broadband spectrograms and LPC displays. Hard-copy visuals and operational definitions also were created on the control phrases to promote reliable measurement. A total of 84 vowels were measured (4 vowels $\times$ 7 instances $\times$ 3 groups). Mean formant values were used to create  $F_1$ – $F_2$  plots to define the vowel quadrilateral for each group. Geometric area values were calculated by summing the areas of the two triangles created by bisecting each quadrilateral (see Fourakis, 1991; Liss *et al.*, 1998; Tsao and Weismer, 1999; Turner *et al.*, 1995). Approximately 30% of the dysarthric vowels were remeasured and inter- and intrajudge reliability were found to be high and acceptable: 87% of the differences between the original and second measures of  $F_1$  were less than 40 Hz; 80% of the differences between the original and

second measures of  $F_2$  were less than 60 Hz. Two instances that exceeded the 40- and 60-Hz values were reassessed and modified as appropriate.

The perceptual impression of reduced vowel strength contrasts in the dysarthric phrases as compared to the control is supported by the indirect measure of vowel working space, the geometric area occupied by the vowel quadrilateral derived from vowels in strong syllables. Specifically, the area generated by the control vowels was approximately 50% larger than either of the two dysarthric quadrilaterals.

### 4. $F_0$ and amplitude variation

Fundamental frequency ( $F_0$ ) and its variation within each digitized phrase was computed automatically using the CSpeech short-term autocorrelation function with center clipping. All pitch traces were inspected visually to identify and edit tracking errors, which occurred occasionally in the dysarthric phrases. The rms amplitude envelope of each phrase was converted automatically to mean decibels and variation around the mean was calculated. To compare mean intraphrase  $F_0$  and amplitude variations across phrases and speaker groups, coefficients of variation were calculated by dividing each standard deviation by its mean.<sup>6</sup> These results, along with  $F_0$  ranges, are provided in Table II.

In summary, the values support the perceptual impression that the hypokinetic phrases tended toward monotonicity. Ataxic and control phrases had higher coefficients of variation for  $F_0$  than the hypokinetic group [ $H(2)=65.1$ ,  $P<0.0001$ ; comparisons significant as  $P<0.05$ ]. The values also support the perceptual impression that the ataxic phrases contained excess loudness variation. With regard to loudness variation, ataxic dB coefficients of variation were significantly greater than those of the control and hypokinetic groups [ $H(2)=32.6$ ,  $P<0.0001$ ; comparisons significant at  $P<0.05$ ].

### C. Procedures

The three groups of listeners were instructed to listen to each phrase and to write down exactly what they heard. They were told that all phrases consisted of real words in the English language produced by several different male and female speakers. They were told that some of the phrases may be difficult to understand, but that they should guess if they did not know what the speaker was saying. They were told that if they could not venture a guess, they were to use a slash to indicate that part of the phrase they could not understand.

TABLE III. Examples of coding lexical boundary errors from the listeners' transcriptions.

Target phrase	Listener response	Error type(s)
to rest and not appear	to wish <i>another beer</i>	DW, IS
alive and eager smile	<i>a liar</i> when <i>he goes</i> well	IS, IW
their jury is below	there's <i>serious</i> below	DW
always lobby water	always <i>love your</i> water	IW
will travel after court	will travel <i>if you're</i> caught	IW
a term arranged inside	I turn <i>her engine</i> fine	IS, DW, IS
confuse the very back	<i>confusing with</i> the bank	DW, IW
amend the slower page	<i>I meant to sort</i> the page	IS, IW
the rally found some light	the <i>raw impounds</i> some light	IW, DS
answer dying temper	<i>entertaining</i> temper	DS

Note: IS refers to insertion of a lexical boundary before a strong syllable; IW refers to insertion before a weak syllable. DS and DW refer to deletions of lexical boundaries before strong and weak syllables, respectively. The first five examples are from transcripts from the ataxic tape and the second five are from hypokinetic tape transcripts.

The listeners were seated in individual cubicles. The audio tapes were presented via the Tandberg Educational sound system in the ASU Language Laboratory over high-quality Tandberg supra-aural headphones. Equivalent sound pressure levels across headphones were verified with a headphone coupler sound level meter (Quest 215 Sound Level Meter). Listeners were instructed to adjust the volume to a comfortable listening level (in 4-dB increments up or down) during the preliminary instructions. They were directed not to alter the volume once the stimulus phrases had begun. The listeners transcribed three practice phrases, which were read by a neurologically normal female speaker. Listeners who made more than one word-transcription error in the practice phrases were not eligible for the study. No listeners were excluded by this criterion.

#### D. Analysis

The total data set consisted of 60 transcripts of the 60 phrases. Three trained judges independently coded the listener transcripts for the presence and type of LBEs. Lexical boundary violations were defined as erroneous insertions or deletions of lexical boundaries. These insertions or deletions were coded as occurring either before strong or before weak syllables (as determined by the target phrasal stress pattern of the phrase, SWSWSW or WSWSWS). Thus, four error types were possible and each phrase had the possibility of containing more than one LBE. Examples from the actual transcripts are provided in Table III. Other tabulations for each transcript included words-correct scores for intelligibility calculations, the number of word substitutions (defined as incorrect words that did not violate lexical boundaries, and that may or may not bear phonetic similarities to the target), and the number of words for which no response was attempted.

The 60 phrases consisted of 360 syllables, 60 of which were phrase-initial syllables and were therefore not subject to LBEs. Of those 300 non-phrase-initial syllables, 102 were word-initial strong syllables; 80 were word-initial weak syllables; 48 were non-word-initial strong syllables; and 70 were non-word-initial weak syllables. The occurrence of each word-initial strong syllable in the target corresponded to the opportunity for the deletion of a lexical boundary be-

fore a strong syllable. Similarly, word-initial weak syllables corresponded to the opportunity for the deletion of a lexical boundary before a weak syllable; non-word-initial strong syllables to the opportunity for an insertion of a lexical boundary before a strong syllable; and non-word-initial weak syllables to the opportunity for an insertion of a lexical boundary before a weak syllable. Thus, the opportunities for producing the different types of LBEs were not equal, but are representative of the opportunities generally available in the English language (Cutler and Carter, 1987).

The codes generated by each of the three judges were merged into one composite data set that reflected instances in which there was 100% agreement among the judges. Twenty potential LBEs were discarded due to interjudge disagreement. The number, type (insertion or deletion), and location (before strong or before weak syllables) were then tallied for each listener group. Words-correct scores were calculated for each tape. Chi-squares, *t*-tests, and analyses of variance were conducted to identify significant differences between category and group means. When the data violated the assumption of a normal distribution, the Kruskal-Wallis analysis of variance by ranks was performed to test for differences within coding categories and between listener groups. *Post hoc* comparisons were evaluated by the Student-Newman-Keuls procedure.

## II. RESULTS

### A. Transcript tabulations

Table IV contains the results of the coding and tabulation procedures for each of the three listening groups. Intelligibility values are listed in the first column of this table. Recall that every effort was made during phrase selection and tape construction to ensure that the mean intelligibility scores elicited by the dysarthric tapes would not differ. The

TABLE IV. Transcript tabulations for the three listening groups.

	% words correct	No. word-substitutions	No. no response	No. LBE
Control	98.3	53	0	22
Hypokinetic	41.8	1059	669	820
Ataxic	43.2	925	1035	610

TABLE V. Lexical boundary error distributions for the dysarthric groups.

Speaker group	LB insertions		LB deletions		IS–IW ratio	DW–DS ratio
	IS	IW	DS	DW		
Hypokinetic	386	225	74	135	1.7	1.8
Ataxic	247	210	75	78	1.2	1.0

Note: The values in the first four data columns are raw scores.

mean words-correct scores of 43.2% for the ataxic phrases and 41.8% for the hypokinetic phrases were found not to be significantly different, so LBE analysis proceeded without modification of the phrase composition.

Of the words that were not correctly transcribed, some consisted of word substitutions, some were not attempted by the listeners (no response), and some consisted of words that violated lexical boundaries. These are listed in the next data columns of Table IV.

The transcripts of the hypokinetic tape contained more instances of word substitution errors than did the transcripts of the ataxic tape (1059 vs 925), although the magnitude of this difference is not great. Actual examples of word substitution errors include, “excess” for “expect,” “quacked” for “stopped,” “little” for “legal,” and “carcass” for “caucus.” In contrast, the ataxic tape transcripts contained many more instances in which no attempt was made for given words than the hypokinetic tape transcripts (1035 vs 669 words). Thus, of the responses that did not obviously violate lexical boundaries, the hypokinetic and ataxic tapes elicited approximately the same number of words-correct and a similar number of word substitution errors, but the transcripts of the ataxic tape contained more instances of unattempted word transcriptions.

Both dysarthric samples elicited many LBEs from the listeners. In all, 1430 LBEs were identified unanimously by the three independent judges: 610 for the ataxic phrases and 820 for the hypokinetic phrases. The control group elicited only a negligible number of LBEs (22) and their data will not be discussed further. No predictions were made regarding the absolute differences in the incidence of LBEs elicited by the two dysarthric tapes. However, the hypokinetic tape elicited significantly more LBEs than the ataxic tape [ $t(38) = -3.227$ ;  $p = 0.003$ ].

## B. Lexical boundary error pattern

As seen in Table V, LBE insertions outnumbered deletions by approximately three times. This was the case for both the hypokinetic and ataxic phrases (2.92 and 2.98 times, respectively). The first four data columns of this table contain the numbers of insertion and deletion errors that occurred before strong and weak syllables for each group. Both groups elicited more LBE insertions before strong than before weak syllables, and more LBE deletions before weak than before strong syllables. However, this difference for error type and location for the ataxic data was very small. A chi-square analysis indicated a significant interaction between the variables of insert/delete and strong/weak for the data generated by the hypokinetic phrases [ $X^2(1, N=4) = 47.637$ ,  $p < 0.001$ ], but not for the data generated by the ataxic phrases.

In our previous paper (Liss *et al.*, 1998), we used two ratio calculations to express strength of adherence to the expected patterns of errors relative to syllabic strength. The first was the number of insertions before strong syllables relative to those before weak syllables (IS–IW). The second was the number of deletions before weak syllables relative to those before strong syllables (DW–DS). Ratio values of “1” indicate that insertions and deletions occur gradually as often before strong and weak syllables; the greater the positive distance from “1,” the greater the strength of adherence to the predicted pattern. For the present hypokinetic data, insertion errors occurred 1.7 times more often before strong than before weak syllables, and deletion errors occurred 1.8 times more often before weak than before strong syllables. The values for the ataxic data signify weak or lack of adherence to the expected pattern: IS–IW was 1.2 and DW–DS was 1.0.

## C. Liss *et al.* (1998) replication

Our previous investigation (Liss *et al.*, 1998) provided a large-scale detailed examination of LBEs elicited by hypokinetic speech. Seventy listeners transcribed 420 phrases produced by seven different speakers with hypokinetic dysarthria. The present investigation differed in several important ways. First, only six speakers provided ten phrases each to produce one listening tape of 60 phrases, but each of the 20 listeners heard ten phrases from all six speakers in their transcription task. This is in contrast to the 20-phrase blocks of three voices heard by the listeners in the first study. Second, although there was some overlap in speakers between the first and second studies, the range of dysarthria severity in the present investigation was more narrow. Also, the phrases selected for the listening tapes were selected on the basis of intelligibility and representativeness of our definition of hypokinetic dysarthria. This served to further increase the homogeneity of the phrases on the hypokinetic listening tape.

Despite these important differences, the results of the present investigation mirrored those of our previous report. In both studies, the hypokinetic speech samples elicited large numbers of LBEs; insertion errors exceeded deletion errors by nearly three times; insertion errors before strong syllables outnumbered those before weak syllables; and the opposite pattern occurred for deletion errors. Strength of adherence patterns were virtually identical in both studies. The IS–IW ratio of the present and previous studies was 1.7; the DW–DS ratio was 1.8 in the present and 1.5 in the previous study. This consistency of findings across the two studies suggests the results are robust and valid for hypokinetic speech.



### III. DISCUSSION

The present study provides compelling evidence that, as with the perception of normal speech, listeners attend to syllabic strength contrasts to segment the acoustic stream in the cases of hypokinetic and ataxic dysarthria. Further, the patterns of LBEs reveal that form of the prosodic deficit in the two dysarthria subtypes differentially influences the listeners' ability to apply the metrical segmentation strategy. This finding is critical for our understanding of speech intelligibility in dysarthria because it demonstrates the reciprocal relationship between listeners' perceptual strategies and the integrity of specific aspects of the acoustic signal. We therefore conclude that reductions in intelligibility for these two speech samples may be thought to arise from different perceptual challenges.

Several authors have suggested that listeners modify their perceptual strategies to decipher degraded speech (Forster, 1989; Marslen-Wilson, 1989; McQueen, 1991; Pisoni and Luce, 1986). This may take the form of hypothesis testing and postperceptual processing. In our previous paper, we hypothesized that when articulatory integrity is most impaired, syllabic contrastivity "matters more." Listeners may rely then on the relatively robust cues associated with syllabic strength contrasts in the face of phonetic uncertainty (Cutler and Butterfield, 1992).

The speakers, phrases, and phrase tokens used in this investigation were carefully selected to ensure that any differences in LBE patterns elicited by the two dysarthric tapes would be interpretable relative to syllabic strength contrasts. Like the hypokinetic speech, the ataxic tape elicited a large number of LBEs in which insertion errors outnumbered deletion errors by nearly three times. Although this finding alone is not interpretable relative to syllabic strength contrasts, it does speak to the listeners' use of syllable onset information to parse the acoustic stream. Recall that our corpus of phrases provided many times more opportunities for lexical boundary deletions than insertions. The high insertion-deletion ratio for both dysarthric tapes provides compelling evidence that the listeners exploited the statistical probabilities of the English language that favors the occurrence of single syllable words (Cutler and Carter, 1987).

In addition to syllable onset information, it is clear that listeners of both dysarthric tapes utilized available syllabic strength information to hypothesize about word-onsets. If they had not, the proportion of insertion and deletion errors that occurred before strong and weak syllables would have mirrored the opportunities for such errors to occur. As stated above, the pattern of LBEs for the hypokinetic tape adhered to the pattern predicted by the MSS, though less strongly than that reported for normal speech. The ataxic tape elicited a pattern of even weaker adherence to the predicted MSS pattern: LBE insertions and deletions occurred with equal frequency before strong and weak syllables.

Given these findings, it can be surmised that the prosodic deficits contributed differentially to the systematic application of the metrical segmentation strategy. Specifically, the characteristics of the ataxic dysarthric speech appeared to pose more of an impediment to the strategy than did the hypokinetic speech. It is not possible from this investigation

to know precisely which acoustic correlates of syllabic strength account for our pattern of results. However, we speculate that the relative syllabic isochrony of the ataxic phrases (see Table II) may have contributed to the decreased effectiveness of the perceptual strategy of attending to syllabic strength contrasts.

Hertrich and Ackermann (1998) published a preliminary report that examined the effects of synthetically altered segment durations on intelligibility and naturalness of ataxic and normal speech. Their goal was to determine whether "normalizing" segment durations in ataxic dysarthria would result in improved speech, and whether distorting normal temporal profiles to match ataxic templates would result in ataxic-sounding speech. The synthetic alterations of the ataxic speech were judged more favorably than the original utterances on a number of perceptual dimensions. However, intelligibility judgements were not affected by the temporal alterations.

The relevance of the Hertrich and Ackermann study is not so much its overall findings. Their speech sample (five sentences presented in seven resynthesized variants each) and listener pool (six speech-language pathologists) were quite limited. The perceptual measures were global ratings that were not sufficiently sensitive to quantify subtle changes. Instead, the relevance to the present report lies in the notion that temporal parameters in ataxic dysarthria may indeed affect listener performance in quantifiable ways. Temporal manipulations of the ataxic phrases produced for our present investigation would provide a perfect test for our hypothesis about syllabic isochrony and perceptual segmentation.

Although all of our hypotheses were based on LBE patterns and proportions, it is of note that the ataxic tape elicited overall fewer LBEs than the hypokinetic tape (see Table IV). If ataxic speech truly poses more of a challenge to listeners' application of the metrical segmentation strategy than hypokinetic speech, why did it appear to elicit fewer lexical boundary errors overall? We speculate that the difference in raw numbers is a byproduct of the transcription compositions. Recall that we assembled phrases for both the hypokinetic and ataxic tapes from larger pools of phrases. Our goal was to select phrases such that the tapes would be of equivalent intelligibility. For this study, the transcribers of the ataxic tape obtained slightly higher words-correct scores than did the transcribers of the hypokinetic tape. This difference was not statistically significant so we met our goal of equivalent intelligibility. However, it does mean that, functionally, there were fewer opportunities for LBEs in the transcripts of the ataxic phrases. If we had modified the ataxic tape by replacing more intelligible phrases with less intelligible ones (which was our plan had equivalent intelligibility not been attained), our LBE count may have increased. In addition, the ataxic phrases elicited a greater number of "no response" from the listeners. It is not immediately obvious why the transcribers of the ataxic tape chose not to venture guesses on many phrases. However, this also serves to reduce the possibility for the identification of LBEs. Taken together, the slight edge in intelligibility and the greater "no responses" may account for the difference in raw numbers.

The data could have been corrected to express the absolute number of LBEs as a function of opportunities available to determine the actual significance of the observed difference. However, none of our predictions about LBE patterns called for such a comparison. If we find significant differences in LBE raw numbers in our subsequent studies when all other things are equal (i.e., intelligibility and number of “no responses”), an alternative explanation will be warranted.

Finally, the findings of our study must be viewed within the limitations of the methodology and the design. First, despite equivalent intelligibility of the dysarthric tapes, and the measures taken to ensure the presence of cardinal dysarthria features for all phrases included on the tapes, individual differences did exist among speakers. Because speakers only provided ten phrases each, it was not possible to assess the full impact of individual variation on listener performance. Second, the results were interpreted relative to one presumed cognitive strategy of perceptual segmentation. Although the results strongly suggest that listeners indeed used this strategy, other strategies certainly were operative as well. The LBE analysis is not designed to reveal alternative strategies. A qualitative assessment of listener errors would be required to speculate about such alternative strategies. On a related issue, our study did not explicitly examine the influence of segmental integrity or goodness on lexical segmentation decisions. Listeners surely used hypotheses about phoneme identity to guide lexical searches. There is no doubt that this interacted with their hypotheses about which syllables constituted word-onsets. Again, a qualitative assessment of listener responses may shed light on the nature of this interaction and yield testable hypotheses that serve further to elucidate the nature of intelligibility deficits in these two dysarthria groups.

Despite the limitations of this investigation, the results offer compelling evidence that a normal model of language processing may be used successfully to examine the perception of dysarthric speech. This constitutes a significant paradigm shift away from traditional approaches to intelligibility that focus exclusively or primarily on output characteristics. Our findings suggest that the classical notions of motor speech disorders be revisited and revised to take into account relevant listener variables.

#### IV. CONCLUSION

The examination of LBEs revealed that listeners used the perceptual strategy of attending to syllabic strength to segment the acoustic stream in both ataxic and hypokinetic dysarthria. The pattern of errors elicited by the hypokinetic speech more closely adhered to predictions offered by the MSS hypothesis than did the errors elicited by the ataxic speech. It is hypothesized that the tendency toward syllabic isochrony in ataxic speech is especially detrimental to the application of the metrical segmentation strategy.

#### ACKNOWLEDGMENTS

This research was supported by Research Grant No. 5 R29 DC 02672 from the National Institute of Deafness and Other Communication Disorders, National Institutes of

Health. Gratitude is extended to the patients of the Mayo Clinic—Scottsdale who participated in this investigation, and to Diana Acevedo, Shannon Beatty, Kellie Filter Sciortino, Michele Marshall, and Melissa Miller for their assistance on the project.

<sup>1</sup>No listeners in the present investigation participated in the Liss *et al.* (1998) study.

<sup>2</sup>A list of the phrases is available electronically from the first author.

<sup>3</sup>Eight speech-language pathologists with experience in dysarthria independently transcribed the 60 phrases produced by a quasi-random sampling of the 14 dysarthric speakers (8 hypokinetic and 6 ataxic).

<sup>4</sup>If words-correct scores would have been different for the two listener groups, a phrase analysis would have been conducted to identify and replace those phrases responsible for the discrepancy.

<sup>5</sup>Although vowel space calculations have been used to capture reductions in articulatory excursions in dysarthria (Turner *et al.*, 1995), vowel spaces are notoriously variable in even the normal population (Tsao and Weismer, 1999). Their use in the present study is an attempt to quantify the perceptual impression of vowel reduction or centralization among our dysarthric speakers (see also Bradlow *et al.*, 1996).

<sup>6</sup>It is recognized that the dB ratio values represent a nonlinear function. Our purpose for these data was simply to make ordinal observations about amplitude variation, and not to extrapolate about magnitude of differences.

Ackermann, H., and Hertrich, I. (1994). “Speech rate and rhythm in cerebellar dysarthria: An acoustic analysis of syllable timing,” *Folia Phoniatr.* **46**, 70–78.

Ackermann, H., and Ziegler, W. (1991). “Articulatory deficits in Parkinsonian dysarthria: An acoustic analysis,” *J. Neurol., Neurosurg. Psychiatry* **54**, 1093–1098.

Adams, S. G. (1991). “Accelerating speech in a case of hypokinetic dysarthria: Descriptions and treatment,” in *Motor Speech Disorders: Advances in Assessment and Treatment*, edited by J. Till, K. Yorkston, and D. Beukelman (Brookes, Baltimore), pp. 213–228.

Bradlow, A. R., Torretta, G. M., and Pisoni, D. B. (1996). “Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics,” *Speech Commun.* **20**, 255–272.

Cole, R. A., and Jakimik, J. (1980). “How are syllables used to recognize words?” *J. Acoust. Soc. Am.* **67**, 44–56.

Connolly, F. (1986). “Intelligibility: A linguistic view,” *Br. J. Disord. Commun.* **21**, 371–376.

Crystal, T. H., and House, A. S. (1988a). “Segmental durations in connected-speech signals: Current results,” *J. Acoust. Soc. Am.* **83**, 1553–1573.

Crystal, T. H., and House, A. S. (1988b). “Segmental duration in connected-speech signals: Syllabic stress,” *J. Acoust. Soc. Am.* **83**, 1574–1583.

Cutler, A. (1993). “Phonological cues to open- and closed-class words in the processing of spoken sentences,” *J. Psycholinguist. Res.* **22**, 109–131.

Cutler, A., and Butterfield, S. (1990). “Durational cues to word boundaries in clear speech,” *Speech Commun.* **10**, 485–495.

Cutler, A., and Butterfield, S. (1991). “Word boundary cues in clear speech: A supplementary report,” *Speech Commun.* **9**, 335–353.

Cutler, A., and Butterfield, S. (1992). “Rhythmic cues to speech segmentation: Evidence from juncture misperception,” *J. Mem. Lang.* **31**, 218–236.

Cutler, A., and Carter, D. M. (1987). “The predominance of strong syllables in the English vocabulary,” *Comput. Speech Lang.* **2**, 133–142.

Cutler, A., and Norris, D. (1988). “The role of strong syllables in segmentation for lexical access,” *J. Exp. Psychol.* **14**, 113–121.

Darley, F., Aronson, A., and Brown, J. (1969). “Differential diagnostic patterns of dysarthria,” *J. Speech Hear. Res.* **12**, 246–269.

Duffy, J. R. (1995). *Motor Speech Disorders* (Mosby, St. Louis).

Fear, B. D., Cutler, A., and Butterfield, S. (1995). “The strong/weak syllable distinction in English,” *J. Acoust. Soc. Am.* **97**, 1893–1904.

Forrest, K., Weismer, G., and Turner, G. (1989). “Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults,” *J. Acoust. Soc. Am.* **85**, 2608–2622.

Forster, K. I. (1989). “Basic issues in lexical processing,” in *Lexical Representation and Process*, edited by W. Marslen-Wilson (MIT, Cambridge), pp. 75–107.

- Fourakis, M. (1991). "Tempo, stress and vowel reduction in American English," *J. Acoust. Soc. Am.* **90**, 1816–1827.
- Gaygen, D. E., and Luce, P. A. (1998). "Effects of phonotactic probability on the recognition of words in continuous speech," *J. Acoust. Soc. Am.* **104**, 1758(A).
- Goldinger, S. D., Luce, P. A., and Pisoni, D. B. (1989). "Priming lexical neighbors of spoken words: Effects of competition and inhibition," *J. Mem. Lang.* **28**, 501–518.
- Gow, D. W., and Gordon, P. C. (1995). "Lexical and prelexical influences on word segmentation: Evidence from priming," *J. Exp. Psychol.* **21**(2), 344–359.
- Grosjean, F., and Gee, J. (1987). "Prosodic structure and spoken word recognition," *Cognition* **25**, 135–155.
- Halle, M., and Keyser, S. J. (1971). *English Stress: Its Form, Its Growth, and Its Role in Verse* (Harper and Row, New York).
- Hertrich, I., and Ackermann, H. (1998). "Auditory perceptual evaluation of rhythm-manipulated and resynthesized sentence utterances obtained from cerebellar patients and normal speakers: A preliminary report," *Clinical Linguistics & Phonetics* **12**, 427–437.
- Kent, R. D., Netsell, R., and Abbs, J. H. (1979). "Acoustic characteristics of dysarthria associated with cerebellar disease," *J. Speech Hear. Res.* **22**, 627–648.
- Klatt, D. H. (1975). "Vowel lengthening is syntactically determined in a connected discourse," *J. Phonetics* **3**, 129–140.
- Klatt, D. H. (1980). "Speech perception: A model of acoustic-phonetic analysis and lexical access," in *Perception and Production of Fluent Speech*, edited by R. A. Cole (Erlbaum, Hillsdale, NJ), pp. 243–288.
- Lehiste, I. (1972). "The timing of utterances and linguistic boundaries," *J. Acoust. Soc. Am.* **51**, 2018–2024.
- Lindblom, B. (1990). "Explaining phonetic variation: A sketch of the H and H theory," in *Speech Production and Speech Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic, Dordrecht, The Netherlands), pp. 403–439.
- Liss, J. M., Spitzer, S. M., Caviness, J. N., Adler, C., and Edwards, B. (1998). "Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech," *J. Acoust. Soc. Am.* **104**, 2457–2466.
- Logemann, J. A., and Fisher, H. B. (1981). "Vocal tract control in Parkinson disease: Phonetic feature analysis of misarticulations," *J. Speech Hear. Disord.* **46**, 348–352.
- Ludlow, C. L., and Bassich, C. J. (1984). "Relationships between perceptual ratings and acoustic measures of hypokinetic speech," in *The Dysarthrias*, edited by M. R. McNeil, J. C. Rosenbek, and A. Aronson (College Hill, San Diego), pp. 163–195.
- Marslen-Wilson, W. (1989). "Access and integration: Projecting sound onto meaning," in *Lexical Representation and Process*, edited by W. Marslen-Wilson (MIT, Cambridge), pp. 3–24.
- Mattys, S. L., and Samuel, A. G. (1997). "How lexical stress affects speech segmentation and interactivity: Evidence from the migration paradigm," *J. Mem. Lang.* **36**, 87–116.
- McQueen, J. (1991). "The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity," *J. Exp. Psychol.* **17**, 433–443.
- Milenkovic, P. H., and Read, C. (1992). *CSpeech (Version 4)* [Computer software], Madison, WI.
- Nakatani, L. H., and Schaffer, J. A. (1978). "Hearing words without words: Prosodic cues for word perception," *J. Acoust. Soc. Am.* **63**, 234–245.
- Oller, D. K. (1973). "The effect of position in utterance on speech segment duration in English," *J. Acoust. Soc. Am.* **54**, 1235–1246.
- Pisoni, D. B., and Luce, P. A. (1986). "Speech perception: Research, theory, and the principal issues," in *Pattern Recognition by Humans and Machines: Speech Perception*, edited by E. C. Schwab and H. C. Nusbaum (Academic, Orlando), Vol. 1, pp. 1–50.
- Quené, H. (1992). "Durational cues for word segmentation in Dutch," *J. Phonetics* **20**, 331–350.
- Quené, H., and Koster, M. (1998). "Metrical segmentation in Dutch: Vowel quality or stress?" *Lang Speech* **41**, 185–202.
- Ramig, L. O. (1992). "The role of phonation in speech intelligibility: A review and preliminary data from patients with Parkinson disease," in *Intelligibility in Speech Disorders*, edited by R. D. Kent (Benjamin, Amsterdam), pp. 120–155.
- Smith, M. R., Cutler, A., Butterfield, S., and Nimmo-Smith, I. (1989). "The perception of rhythm and word boundaries in noise-masked speech," *J. Speech Hear. Res.* **32**, 912–920.
- Tsao, Y.-C., and Weismer, G. (1999). "Does interspeaker variation in habitual speaking rate predict the size of the acoustic vowel space?" *J. Acoust. Soc. Am.* **105**, 1097(A).
- Turner, G. S., Tjaden, K., and Weismer, G. (1995). "The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis," *J. Speech Hear. Res.* **38**, 1001–1013.
- van Ooijen, B., Bertoncini, J., Sansavini, A., and Mehler, J. (1997). "Do weak syllables count in newborns?" *J. Acoust. Soc. Am.* **102**, 3735–3741.
- Vitevitch, M. S., and Luce, P. A. (1998). "When words compete: Levels of processing in perception of spoken words," *Psych. Sci.* **9**(4), 325–329.
- Weismer, G. (1984). "Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination," in *The Dysarthrias*, edited by M. R. McNeil, J. C. Rosenbek, and A. Aronson (College Press, San Diego), pp. 101–130.
- Weismer, G. (1991). "Assessment of articulatory timing," in *Assessment of Speech and Voice Production: Research and Clinical Applications*, NIDCD Monograph, Vol. 1, pp. 84–95.
- Weismer, G., and Martin, R. (1992). "Acoustic and perceptual approaches to the study of intelligibility," in *Intelligibility in Speech Disorders*, edited by R. D. Kent (Benjamin, Amsterdam), pp. 67–118.
- Yorkston, K. M., and Beukelman, D. R. (1981). *Assessment of the Intelligibility of Dysarthric Speech (Manual)* (CC Publications, Oregon).

# Effect of speaking rate and contrastive stress on formant dynamics and vowel perception

Michel Pitermann<sup>a)</sup>

Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada

(Received 4 October 1999; accepted for publication 1 March 2000)

Vowel formants play an important role in speech theories and applications; however, the same formant values measured for the steady-state part of a vowel can correspond to different vowel categories. Experimental evidence indicates that dynamic information can also contribute to vowel characterization. Hence, dynamically modeling formant transitions may lead to quantitatively testable predictions in vowel categorization. Because the articulatory strategy used to manage different speaking rates and contrastive stress may depend on speaker and situation, the parameter values of a dynamic formant model may vary with speaking rate and stress. In most experiments speaking rate is rarely controlled, only two or three rates are tested, and most corpora contain just a few repetitions of each item. As a consequence, the dependence of dynamic models on those factors is difficult to gauge. This article presents a study of 2300 [iai] or [iei] stimuli produced by two speakers at nine or ten speaking rates in a carrier sentence for two contrastive stress patterns. The corpus was perceptually evaluated by naive listeners. Formant frequencies were measured during the steady-state parts of the stimuli, and the formant transitions were dynamically and kinematically modeled. The results indicate that (1) the corpus was characterized by a contextual assimilation instead of a centralization effect; (2) dynamic or kinematic modeling was equivalent as far as the analysis of the model parameters was concerned; (3) the dependence of the model parameter estimates on speaking rate and stress suggests that the formant transitions were sharper for high speaking rate, but no consistent trend was found for contrastive stress; (4) the formant frequencies measured in the steady-state parts of the vowels were sufficient to explain the perceptual results while the dynamic parameters of the models were not. © 2000 Acoustical Society of America. [S0001-4966(00)02406-1]

PACS numbers: 43.70.Fq [AL]

## INTRODUCTION

Previous studies indicate that the first two formant frequencies measured in the steady-state part of a vowel play an important role in its characterization (Helmholtz, 1885; Joos, 1948; Delattre *et al.*, 1952). However, the formants of the same vowel uttered by different speakers, in different contexts, at different speaking rates and with different stress patterns, exhibit a lot of variability (Joos, 1948; Peterson and Barney, 1952; Stevens and House, 1963; Lindblom, 1963).

The formant dependence on speaking rate and stress is usually described either by a centralization process or a context assimilation. Some studies have reported that formant frequencies of unstressed vowels or vowels uttered at fast speaking rates shift to a more central position around (500, 1500) Hz in the ( $F_1, F_2$ ) plane (Joos, 1948; Koopmans-van Beinum, 1980). This is called the *centralization effect*. A traditional view is that the vocal tract tends to become more similar to its "schwa" configuration for unstressed vowels and for fast utterances. In contrast, other studies have reported that vowel formants shift toward the values of their context instead of a central position (Stevens and House, 1963; Moon and Lindblom, 1994). This is called *contextual assimilation*.

An important issue related to vowel formant variability

is that different vowels can be characterized by the same steady-state values of the first two formants (Joos, 1948; Peterson and Barney, 1952; Stevens and House, 1963). Because static formant values do not seem sufficient to identify vowels, dynamic information is also analyzed.

Lindblom and Studdert-Kennedy (1967) showed that vowels were categorized differently when they were presented in semivowel context instead of isolation. In context, listeners heard formant values extrapolated beyond the values actually reached at the end of the transitions. This *perceptual overshoot* was confirmed by Nearey (1989) for occlusive contexts. But, another perceptual experiment presented in Nearey's article showed that the classification of sustained vowels was influenced by the presence of other sustained vowels. Hence, an acoustical contrast between stationary segments may also contribute to a perceptual overshoot. Some results conversely suggest that listeners can perceive formant values averaged over a full syllable, i.e., a value closer to the vowel context than the value actually reached (Pols and van Son, 1993). van Son (1993) proposed that a listener would exhibit a perceptual overshoot only when the vowel context had been identified.

Strange (1989a, b) presented several experiments where the vowel steady-state part of natural consonant-vowel-consonant (CVC) stimuli were excised from the speech signals and replaced by silence. The perceptual scores of these silent-center stimuli were at least as good as the scores of the

<sup>a)</sup>Electronic mail: mpiter@psyc.queensu.ca

excised stationary segments presented in isolation. She proposed the theory of *dynamic specification of vowels* stating that vowels in continuous speech are better identified by coarticulation with their context than by information intrinsic to the vowel nuclei. Strange's results were replicated with adults and prelingual babies aged between 6 and 12 months for one German and one English vowel contrast (Polka and Bohn, 1996), and with adults for the 14 German monophthongs (Strange and Bohn, 1998). On the other hand, Nearey and Assmann (1986) showed that isolated Canadian English monophthongs were slightly diphthongized, i.e., the steady-state stimuli contained significant formant frequency change over time. The same level of diphthongization was found for the vowels produced in context. The results of their experiments suggested that listeners may rely more on this *vowel-inherent spectral change* (VISC) than dynamic specification to identify vowels. Hence, vowels in continuous speech would be better characterized by their nuclei than by coarticulation with context. Andruski and Nearey (1992) showed perceptual results contradicting dynamic specification of vowels, but compatible with VISC. Jenkins, Strange, and Trent (1999) presented an experiment based on the silent-center paradigm showing the opposite, i.e., results contradicting VISC theory but compatible with dynamic specification of vowels.

Automatic classification procedures also showed that dynamic information may help to categorize vowels. Cassidy and Harrington (1992) compared the performance of a neural network in vowel categorization for logatomes [CVd] using either static or dynamic information. The neural network used either formant transitions or stationary formant values duplicated several times in order to have the same segment durations in both cases. The network had higher identification scores when dynamic information was available. In Beautemps (1993) a neural network could categorize [iai] and [iei] speech signals when only one 26-ms segment of each  $F_1$  transition of [iV] stimuli was available. Conversely, the network failed a complete categorization when the first three formants values measured in the stationary part were used. In contrast with the two previous works, Watson and Harrington (1999) showed that a Gaussian classification algorithm could categorize vowels in an /h\_d/ or /h\_/ context only on the base of vowel duration plus formant values measured in the steady-state part of the vowel. But, /h\_d/ and /h\_/ contexts produce less coarticulation than other consonant contexts (Stevens and House, 1963). The dynamics involved in the transitions could therefore be insufficient to play a role.

As Nearey (1989) proposed, all the previous contradicting results suggest that listeners may use many strategies together, but the perceptual weights given to each of them would depend on the situation and task. Perceptual results and automatic classification procedures indicate that some of these strategies may rely on formant dynamics. Its characterization could thereby improve our understanding of vowel categorization.

There are no good theories predicting formant dynamics or what dynamic features could be used in vowel classification. It is not even clear whether the shape of a formant

transition plays a perceptual role. Carré *et al.* (1994) activated the commands of a vocal-tract model (Carré and Mrayati, 1992) with different mathematical functions (linear, parabola, sine, etc.) to synthesize [iai] stimuli. An analysis of the formant transitions in the ( $F_1, F_2$ ) plane showed that the shape of the transition varied with the mathematical form of the model commands. The perception of [a] was affected only when the canonical formant frequencies corresponding to a good [a] were far from being reached. Nábělek and Ovchinnikov (1997) perceptually compared linear and quadratic formant contours in a vowel identification task. The two formant contours led to the same categorization.

Other results suggest that the shape of formant transitions might change vowel categorization. Divenyi *et al.* (1995) showed that perceptual overshoot depends on transition slope. In Carré *et al.* (1996), [abi] stimuli were synthesized by means of a vocal-tract model (Carré and Mrayati, 1992) with different phasing between the consonant and vowel gestures. The shape of the formant transitions varied with this "articulatory" phasing. For some formant transitions, listeners heard an extra vowel. The perceptual results were [æbi], [abi], or [abei], depending on the articulatory phasing. As a consequence, even if a description of formant dynamics may not need to be highly accurate in all situations (Carré *et al.*, 1994; Nábělek and Ovchinnikov, 1997), its characterization may lead to testable predictions in vowel identification.

Some attempts have been made to describe formant dynamics. For example, Broad and Clermont (1987) offered several models of formant transition. The more elaborate one was an exponential fit and it had good variance explanation. In van Son and Pols (1992), a professional radio speaker read a Dutch text at normal, then fast, speaking rate. Vowel duration was normalized during formant extraction, i.e., all formant time-series segments corresponding to vowels had the same length. No differences were found between the formant contours when the two speaking rates were compared, but this representation helped to distinguish front from back vowels. The normalized time series were also fitted to Legendre polynomial expansions in order to characterize the shape of the formant transitions, and the same results were found, i.e., no fit differences for the two speaking rates and the same classification performance. In Watson and Harrington (1999), formant time series were fitted to discrete cosine expansions to characterize formant contours of vowels uttered in [h\_d] and [h\_] contexts. This representation helped to identify the diphthongs. However good, none of these formant contour characterizations was proven sufficient to identify all vowels in a general context. New attempts may therefore be needed. This work presents models of formant dynamics and kinematics and analyzes how their parameters and the formant frequencies depended on speaking rate and contrastive stress.

To model the dynamics of the first two formants, the system underlying their production was assumed stationary during an analyzed transition.  $F_1$  and  $F_2$  were modeled separately for two reasons. First, it can be seen in Pitermann (1996) that the first two formants of the corpus presented here ([iai] and [iei] stimuli) were not synchronized. Hence,

TABLE I. Equations of the dynamic and kinematic models. Symbol  $x_n$  represents the formant frequency at time coordinate  $n\Delta t$ ;  $a, a_i, b, c, c_1, c_2, d, p, q, t_0$ , and  $x^*$  are the model parameters;  $k$  is the order of linear model (1);  $\Delta t$  is the sampling step (i.e., 9 ms).

	Dynamic models		Kinematic models	
Linear and exp.	$x_n = a_0 + \sum_{i=1}^k a_i x_{n-i}$	(1)	$x_n = x^* + pe^{-qn\Delta t}$	(2)
Critical damping	$x_n = a_0 + a_1 x_{n-1} - \frac{a_1^2}{4} x_{n-2}$	(3)	$x_n = x^* + (c_1 + c_2 n \Delta t) e^{-qn\Delta t}$	(4)
Logistic	$x_n = c + \frac{a(x_{n-1} - c)}{ad + (1-d)(x_{n-1} - c)}$	(5)	$x_n = c + \frac{a}{1 + e^{-b(n\Delta t - t_0)}}$	(6)

two independent models appeared a better choice. Second, there are more mathematical tools available to guide one-dimensional than two-dimensional modeling.

It is often presumed that an articulatory movement can approximately be described by a second-order linear model (Sonoda, 1987; Laboissière *et al.* 1995). A corpus with simple articulatory movements was chosen (only a front-back-front tongue movement and a jaw opening-closing may be involved in the production of [iai] and [iei] stimuli). As a consequence, and because it is the simplest model, the linear model was selected first.

In a revision of Lindblom's original target undershoot model (Lindblom, 1963), Moon and Lindblom (1994) proposed to describe the relationship between vowel duration and steady-state formant frequency by a second-order linear model at critical damping. Although the mathematical forms of a formant time series and of the relationship between vowel duration and formant frequency may be different, these two concepts may be closely tied. This model was therefore chosen, too. One may object that the second-order linear model at critical damping is a special case of the linear model first selected; hence, a study of the latter should be sufficient to back up or contradict the former. This point will be discussed in the "Results" section, IID.

A small-order linear model cannot accurately describe an asymmetrical transition. The first- and second half of each formant transition were therefore separately modeled. To model a whole transition, a nonlinear dynamic model having the logistic curve among its solutions was implemented. It will be called the *logistic model* in this article.

Fitting the parameters of a dynamic model or of one of its solutions (a kinematic model, i.e., a curve fitting) may

lead to different results. The reason is that the kinematic and dynamic model equations are different, and the associated cost functions to be minimized are different, too. For example, an exponential is the solution of a first-order linear dynamic model. Fitting an exponential or a linear model may be different. A linear least-squares method can be used for the linear model, not for the exponential. Hence, the two fits may lead to different results.

For each dynamic model, one of its solutions was therefore tested: an exponential, a combination of an exponential and time, and a logistic curve, respectively, for the linear, second-order linear at critical damping, and logistic models. Only the exponential was chosen as the solution for the linear model because, as will be justified in Sec. IID, the first order was sufficient to describe the half-transitions of the corpus.

The model equations are summarized in Table I and their graphics can be seen in Fig. 1.

Three general criticisms can be made about most studies dealing with speaking rate and formants extracted from natural speech. First, many corpora include a lot of different phonemes in different contexts with few examples of each particular case, e.g., Kuwabara (1985); van Son and Pols (1992). Many sources of variability may thereby be responsible for the results, and it is difficult to assess the role of each of them. Second, only two or three speaking rates are usually used in those experiments, e.g., Imaizumi and Kiritani (1989); Loevenbruck and Perrier (1993). Third, the speaking rates are generally not controlled. The speakers are usually asked to speak slowly, naturally or fast, e.g., van Son and Pols (1992); Beautemps (1993).

In contrast, this article presents an analysis of approxi-

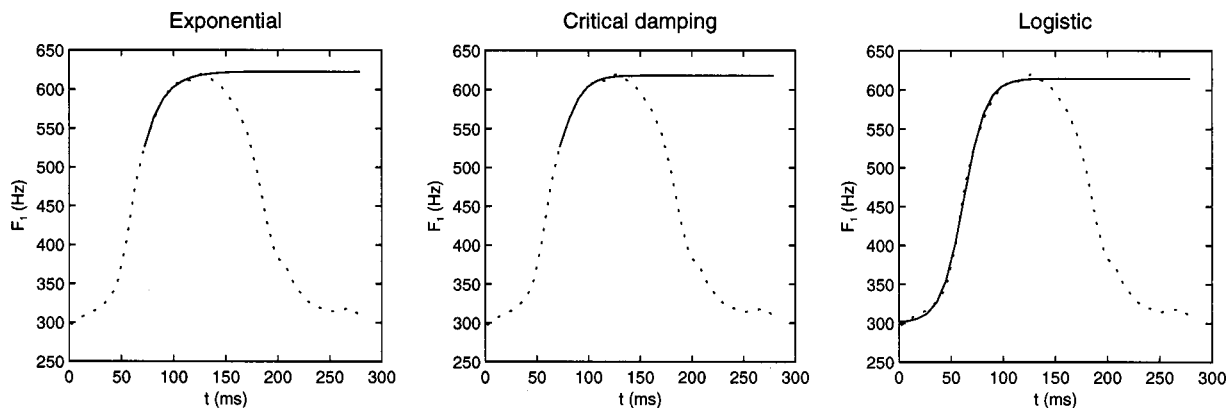


FIG. 1. Plots of each kinematic model (solid lines) fitted to a second-half or full transition of the first formant (dotted line) of the first part of an [iai] stimulus.

mately 2300 [iai] or [iɛi] stimuli uttered by two speakers in a carrier sentence at nine or ten speaking rates controlled by metronome for two contrastive stress patterns. The stimuli were categorized during a perceptual task. The first two formants were measured in the stationary parts of the vowels, and three dynamic and three kinematic models were fitted to the formant transitions. The contribution of static and dynamic information to the perceptual results was assessed, and the centralization versus context assimilation issue was discussed in light of the results. The dynamic and kinematic models gave a quantitative description of the formant contours to be related to the perceptual results. The high number of speaking rates controlled by metronome and the high number of repetitions of each corpus item provided more accuracy for the relationship between speaking rate and variables than what is usually presented.

## I. METHOD

### A. Speech material

The first two formants of [ia] and [iɛ] transitions taken from [iVi] segments were analyzed (V was [a] or [ɛ]). The segments were excised from two carrier sentence portions [iliViimedia] taken from two French sentences: “*Il y a immédiatement souscrit*” (“He immediately subscribed to it”) and “*Il y est immédiatement retourné*” (“He immediately returned to it”). The carrier sentence portions were produced in an anechoic room by two French-speaking male speakers, A and B, at ten and nine different speaking rates, respectively. All speaking rates were controlled by metronome. The speakers had to synchronize the beginning of the carrier segment with a metronome beat, and the plosion of [d] with the following one. The metronome rhythms were 48, 54, 60, 69, 80, 88, 96, 104, 112, and 120 beats/min. The lowest rate was very slow for the speakers; the highest was the limit bearable by each of them, i.e., the highest rate before they could not synchronize their production and the metronome beats. The speakers were asked to produce the whole corpus placing a contrastive stress on V, then a second time without any contrastive stress. I controlled the recording sessions, and each combination of speaking rate, contrastive stress, and vowel was repeated until 30 satisfying examples were produced. Around 4600 formant transitions were therefore analyzed (2 vowels×2 contrastive stresses×19 speaking rates×30 repetitions×2 formants).

The signals were recorded with a microphone (Neuman U 87 P 48) connected to a 8-kHz, low-pass filter and sampled at 32 kHz by a DAT recorder (Sony PCM-2300, 16-bit recording). The cutoff frequency of the low-pass filter was 8000 Hz, and its slope was 360 dB/octave. The corpus was then digitally transferred from the DAT to a computer. The sampling frequency was thereafter reduced to 16 kHz on the computer by taking away one sample out of two. The 8-kHz low-pass filter used during the recording sessions avoided any potential aliasing problems when the sampling frequency was reduced from 32 to 16 kHz.

## B. Perceptual analysis

The corpus was perceptually evaluated to spot improperly produced data, i.e., [iVi] speech signals for which the perceptual identity of V did not correspond to the instructions given to the speakers.

Eight French-speaking participants between 22 and 30 years of age and without phonetic training took part in the experiment. No listeners reported suffering from impaired hearing. The data of the last listener were lost due to an experimental mistake.

For each listener, one list per talker was made up of four different [iVi] segments for each combination of rate, stress, and vowel. For example, one list corresponding to talker B contained 144 items (4 occurrences×9 rates×2 stresses×2 vowels). Since the corpus was large enough, no speech signals were used twice. The stationary parts of [i] were not kept, thus the stimuli sounded more [jVj] than [iVi]. To eliminate onset and offset noise, the signal amplitude was multiplied by a linear 12.5-ms slope at the beginning and end of each stimulus.

Each subject listened to one list per speaker. The stimuli were randomly presented to each listener. Two successive items were separated by a 3-s silence during presentation. Four participants started with a speaker-A list, the other three with a speaker B. Before starting the perceptual test, the subjects heard 20 samples from another list to adjust the signal intensity at a comfortable level and become familiar with the assignment.

The task was to write down each vowel identified between the two [i]. When a listener hesitated, he or she was allowed to leave a blank or choose the closest vowel. No information was given about the carrier sentence, the vowel to be identified, or the phonetic transcription to use. It was therefore a test with open answers.

The signals were played back on Beyer Beyerdynamic dt325 headphones connected to a Sparc 20 SUN workstation containing a 16-bit linear digital-to-analog converter.

Owing to the loss of data, the results of only 28 out of 30 repetitions of each corpus item were available. However, this only amounted to a loss of less than 7% of the perceptual data.

## C. Formant measurements

Formant frequency estimation is a mathematically ill-formed problem, i.e., there are no theorems proving that a solution exists, that a solution is unique, and that a solution continuously depends on the initial conditions. Hence, several heuristics may be tried for each problem. In this study, the outcome of dynamically modeling a formant transition may be influenced by the quality of the formant extractor used. Several formant estimate procedures were therefore compared: a pure linear predictive coding (LPC) method (Atal and Hanauer, 1971), a method combining the results of one cepstral and two LPC algorithms (Soquet, 1995a), and a pitch-synchronous method (Soquet, 1995b). The pitch-synchronous method produced cleaner formant time series than the other two algorithms, but all techniques led to the

same conclusion as far as dynamic modeling was concerned. I thereby kept the LPC-based formant extractor to stick to the simplest tool.

The speech signals were sampled at 16 kHz, then modified by means of a  $(1-0.95z)$  preemphasis filter (in  $z$ -transform notation). Every 9 ms, the first two formant frequencies were estimated by means of an 18-order covariance LPC analysis in a 26-ms rectangular window. The lowest pole frequency was assigned to the first formant, the second pole frequency to the second formant.

All formant time series were visually examined to check for formant misclassifications. Formant insertions were manually erased. When a formant omission was found, the time series of the first two formants of the speech signal were reestimated by means of an LPC model with a different order. If no orders in the [10,20] interval could produce two time series without omissions, the speech signal was discarded. Less than 1% of the corpus was rejected this way.

Values of 0.70, 0.75, 0.80, 0.85, 0.90, and 0.95 were tested for the preemphasis filter. Since these values did not lead to different results as far as dynamic modeling of formant transitions was concerned, 0.95 was selected in order to be consistent with the literature (Makhoul 1975).

For each corpus item, formant frequencies of vowel [V] were defined as the  $F_1$  maximum and  $F_2$  minimum of their time series. This means that a formant frequency was measured at the time coordinate for which the time series took on a nearly horizontal slope inside segment [V].

## D. Formant modeling

### 1. Choice of models

Three dynamic models and one solution of each model were fitted to each formant transition. The three dynamic models were a linear, a second-order linear at critical damping, and a logistic model. The three kinematic models were an exponential, a combination of exponential and time, and a logistic curve. The model equations are summarized in Table I and their graphics can be seen in Fig. 1.

### 2. Parameter estimations

The coefficients of the six models were estimated by means of least-squares methods for each formant transition. Because there were approximately 4600 formant transitions in the corpus, around 4600 sets of parameters were estimated for each model. The merit function was the classical sum of the squared differences between model-generated and measured values. The coefficients of the linear model (1) that gave rise to the minimum merit function were determined by means of a conventional linear method. The coefficients of the nonlinear difference equation (5) and curves (2), (4), and (6) were estimated by means of the Levenberg–Marquardt algorithm. This is an optimizer that gradually switches from a steepest-descent method to a quadratic approximation when the merit function approaches a minimum (Press *et al.*, 1992).

The merit function of the second-order linear model at critical damping (3) was

$$E = \frac{1}{N-2} \sum_{i=3}^N \left( x_i - a_0 - a_1 x_{i-1} + \frac{a_1^2}{4} x_{i-2} \right)^2, \quad (7)$$

where  $E$  is the merit function,  $x_i$  is the  $i$ th sample of the time series,  $N$  is the length of the time series, and  $a_0$  and  $a_1$  are the parameters of the model. Setting to 0 the first derivative of the merit function with respect to  $a_0$  and  $a_1$  leads to a nonlinear system of two equations

$$\begin{aligned} a_0 &= \bar{x}_n - a_1 \bar{x}_{n-1} + \frac{a_1^2}{4} \bar{x}_{n-2}, \\ 0 &= \sigma_{x_{n-2}}^2 \left( \frac{a_1}{2} \right)^3 - 3 \sigma_{x_{n-1}x_{n-2}} \left( \frac{a_1}{2} \right)^2 \\ &\quad + (2\sigma_{x_{n-1}}^2 + \sigma_{x_n x_{n-2}}) \left( \frac{a_1}{2} \right) - \sigma_{x_n x_{n-1}}, \end{aligned} \quad (8)$$

where

$$\bar{x}_{n-k} = \frac{1}{N-2} \sum_{i=3}^N x_{i-k}, \quad (9)$$

$$\sigma_{x_{n-k}}^2 = \frac{1}{N-2} \sum_{i=3}^N (x_{i-k} - \bar{x}_{n-k})^2, \quad (10)$$

$$\sigma_{x_{n-k}x_{n-l}} = \frac{1}{N-2} \sum_{i=3}^N (x_{i-k} - \bar{x}_{n-k})(x_{i-l} - \bar{x}_{n-l}). \quad (11)$$

The procedure to estimate  $a_0$  and  $a_1$  for a time series  $x_i$  was straightforward. The third-degree polynomial (8) was solved to produce one or three real roots. For each  $a_1$  (one or three values), the corresponding  $a_0$  value was calculated by means of the first equation of (8). The couple  $(a_0, a_1)$  that gave rise to the minimum merit function (7) was selected.

The order of the linear difference equation (1) was determined by analyzing the random variable  $(\hat{a}_i - a_i)/\hat{\sigma}(\hat{a}_i)$ , where  $a_i$  is the  $i$ th unknown coefficient of the model for the full population of all possible formant transitions,  $\hat{a}_i$  is its estimate based on a sample of the population, and  $\hat{\sigma}(\hat{a}_i)$  is the estimate of its standard error. When the conditions of application of a linear regression are met, this variable is characterized by a  $t$  distribution (Spiegel, 1988, Chap. 14, p. 300). The null hypothesis “ $H_0: a_i=0$ ” could thereby be tested at the 0.05 level for any coefficient of the model. The order of the linear model (1) was iteratively increased by 1 as long as the null hypothesis for the highest-order coefficient  $a_k$  was rejected. The highest-order  $k$  for which “ $H_0: a_k=0$ ” was not rejected was selected.

The [iVi] segmentation procedures were the following. Complete formant transitions between preceding [i] and [V] were modeled by means of the logistic models (5) and (6). The first sample of a transition was manually chosen near the end of the horizontal part of the formant time series of the [i], the last sample was the next maximum for  $F_1$  or minimum for  $F_2$ .

In contrast to models (5) and (6), only the half-transitions before or after the points of inflection were modeled by means of the four other models. The position of each point of inflection was determined by means of parameter  $t_0$  of the logistic curve (6). To increase  $t_0$ -estimate accuracy, a speaker was assumed to use the same articulatory gesture for



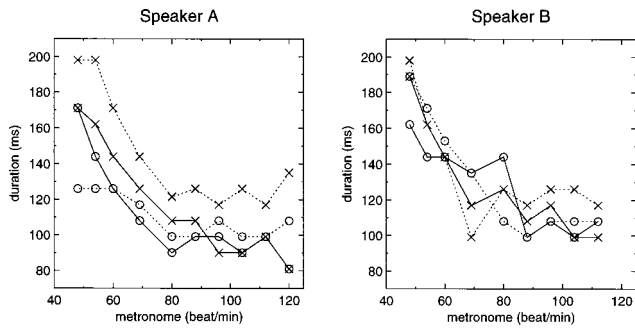


FIG. 2. Length medians of the [ia] or [iε]  $F_1$  transitions plotted as functions of speaking rate measured in metronome rhythm: “○” stressed vowels, “×” unstressed vowels, solid lines for [ia], and dotted lines for [iε]. Low speaking rates are on the left-hand side of each graph.

the 30 repetitions of a corpus item. Accordingly, the 30 transitions were presumed to be characterized by a single coefficient vector  $(a, b, c)$  while each transition assumed a different time position  $t_0$ . I used the Viterbi algorithm in order to estimate a unique  $(a, b, c)$  vector and 30 different  $t_0$  values simultaneously using the 30 transitions (Viterbi, 1967). The method consisted of alternatively estimating either one structural vector  $(a, b, c)$  or the 30 time positions  $t_0$  while keeping the other coefficient class fixed until convergence was achieved.

## II. RESULTS

### A. Transition length and speaking rate

Figure 2 shows how the duration of the  $F_1$  transitions varied with speaking rate. Each point of the plot is a median of 30 durations computed from 30 repetitions of a corpus item in one speaking rate and contrastive stress condition. The same pattern was observed for  $F_2$ .

The graphics show that formant transition duration decreased as speaking rate increased. A floor effect appeared after 80 beats/min, although the speakers continued to increase their speaking rate. Other parts of the segment sentence were therefore shortened for the highest speaking rates.

### B. Perceptual evaluation of the corpus

The purpose of the perceptual analysis was to spot the [iVi] tokens that were incorrectly produced.

Around 99.5% of vowels  $V$  were labeled /a/, /ε/, or /e/. Because [e]–[ε] distinction depended on regional accent, and because the listeners did not receive any references, labels /e/

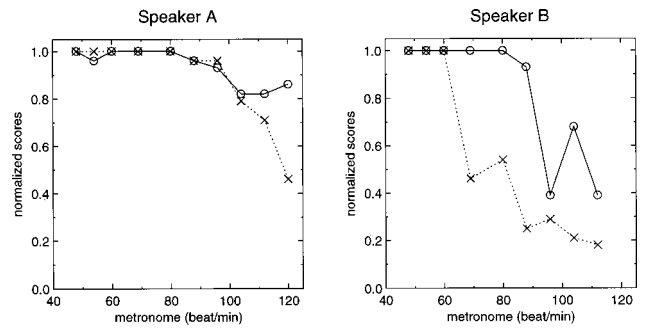


FIG. 3. Normalized identification scores averaged over the seven listeners for speakers A and B's vowel [a] as functions of speaking rate: “○” stressed [a], “×” unstressed [a]. Low speaking rates are on the left-hand side of each graph.

were replaced by /ε/. This way, 99.5% of the labels became /a/ or /ε/, and most mistakes were considered to be confusions between [a] and [ε].

More than 90% of vowels [ε] were correctly identified for 35 speaking rate and contrastive stress groups out of 38 [38=(9 rates×2 stresses)+(10 rates×2 stresses)]. The identification scores of the remaining three conditions were 86%.

Figure 3 shows the average identification scores normalized by the total number of answers of the seven listeners for vowel [a] as a function of speaking rate.

More than 90% of speaker A's [a] were correctly identified for the seven lowest speaking rates. The identification scores were below 85% for the other speaking rates.

The identification scores of speaker B's [a] were above 90% for only 9 categories out of 18 (the stressed [a] produced at the six lowest speaking rates and the unstressed [a] produced at the three lowest ones). For the other categories, the identification scores varied between 18% and 68%.

### C. Formant measurements

Processing about 4600 formant transitions by means of six dynamic or kinematic models led to around 27 600 fits. Due to noise, some of them were physically impossible and should be considered as aberrations. To lower their impact on the statistical description of the corpus, the median and a similar measure to semi-interquartile range were chosen instead of the mean and square deviation. For consistency, the formants were described in the same way.

Instead of the semi-interquartile range, the semi-68%-range leaving 16% instead of 25% of the elements in each tail of the distribution was used. The advantage of the semi-68% compared to the semi-interquartile range is that it has

TABLE II. Medians and semi-68% ranges of the first two formant frequencies (Hz) measured in the steady-state part of speakers' [i], [ε], or [a]. The medians were computed across all speaking rate and stress conditions.

	$F_1$				$F_2$			
	Speaker A		Speaker B		Speaker A		Speaker B	
	Median	Range	Median	Range	Median	Range	Median	Range
[i]	282	28	225	14	1965	36	1954	43
[ε]	373	20	353	32	1848	37	1890	52
[a]	530	73	513	144	1661	102	1732	168

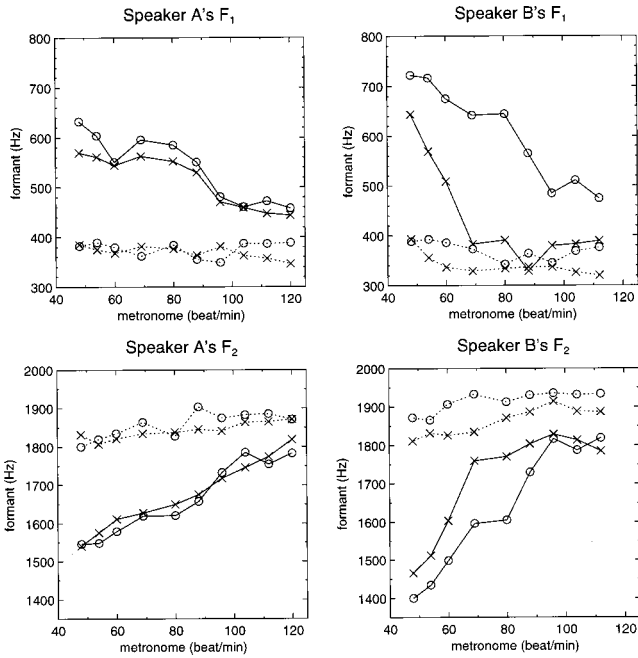


FIG. 4. Medians of the first two formant frequencies (Hz) measured in the steady-state part of speakers' [ε] and [a] for all combinations of speaking rate and contrastive stress. The data are plotted as functions of speaking rate measured in metronome rhythm: "○" stressed vowels, "×" unstressed vowels, solid lines for [a], and dotted lines for [ε]. Low speaking rates are on the left-hand side of each graph.

the same value as the standard deviation when the distribution is Gaussian. It is more convenient for comparisons when different statistical feature sets are used to describe the data.

Table II shows the medians and semi-68% ranges of the first two formants of [i], [ε], and [a] for both speakers. Those values were computed across all repetitions, speaking rates, and stresses. The variability of [i]'s formant frequencies was lower than [a]'s, but similar to [ε]'s.

Figure 4 shows the medians of the first two formants of each speaker's [ε] and [a] computed for each set of 30 repetitions of a corpus item for all combinations of speaking rate and contrastive stress. The data are plotted as functions of speaking rate measured in metronome rhythm. Table III

shows the results of two-factor (rate and stress) analyses of variance for all combinations of speaker, vowel, and formant. All analyses of variance were carried out with a linear additive model with replication. Interactions between factors were included in the model.

#### D. Model parameters

The first step to analyze the corpus using a linear model (1) was to estimate its order  $k$ . To validate the first order, the null hypothesis that  $a_1$  was equal to 0 was tested for each formant half-transition. The hypothesis was rejected at the 0.05 level in 97.7% of the cases. Hence,  $a_1$  should be considered as different from 0, so the first order was mandatory. For the second order, the null hypothesis was tested for  $a_2$ , and it was rejected at the 0.05 level in 4.5% of the cases. Thus,  $a_2$  could not be considered as different from 0, and the first order was sufficient to model the formant half-transitions. Accordingly, we used a first-order linear model and one of its solutions, an exponential.

Because a first-order linear model was sufficient to describe the data, the second-order model at critical damping proposed in Moon and Lindblom (1994) may appear inadequate. Being able to reject higher-order linear models, however, does not imply that a first-order linear model was underlying the data. It only means that, with the number of points available for each fit under the noise condition, no information could be found about higher-order parameters by a classical linear regression.

When a constraint between model parameters is added, the parameter space suitable for a fit is reduced. For example, when a linear second-order model at critical damping is responsible for some data, the noise can hide information about the second order. If the topology of the cost function is complex with many local minima, the fitting procedure may easily be caught in a local minimum far away from the global one. However, adding the critical constraint to the fitting process reduces the search in a 3D space [ $a_0$ ,  $a_1$ , and  $a_2$  from Eq. (1) of Table I] to a search in a 2D space [ $a_0$  and  $a_1$  from Eq. (3)], i.e., in a plane. Hence, the algorithm may

TABLE III. Results of analyses of variance of formant measurements for speaking rate and contrastive stress factors. One analysis was carried out for each combination of formant, speaker, and vowel. Labels "DF," "S:R", and "Resid" stand for degree of freedom, interaction between stress and rate, and residuals, respectively. When  $p < 1e-10$ , 0 was reported in the table.

	Speaker A						Speaker B					
	[ε]			[a]			[ε]			[a]		
	DF	F	p	DF	F	p	DF	F	p	DF	F	p
	$F_1$											
Stress	1	17.3	3.7e-5	1	79.9	0	1	296.3	0	1	4261.3	0
Rate	9	11.4	0	9	158.4	0	8	47.1	0	8	637.1	0
S:R	9	24.6	0	9	3.1	0.0011	8	11.8	0	8	83.7	0
Resid	604			620			569			589		
	$F_2$											
Stress	1	37.5	1.7e-9	1	34.4	7.4e-9	1	549.4	0	1	558.3	0
Rate	9	52.6	0	9	489.4	0	8	81.6	0	8	972.5	0
S:R	9	10.0	0	9	8.3	0	8	10.7	0	8	44.2	0
Resid	599			620			553			584		

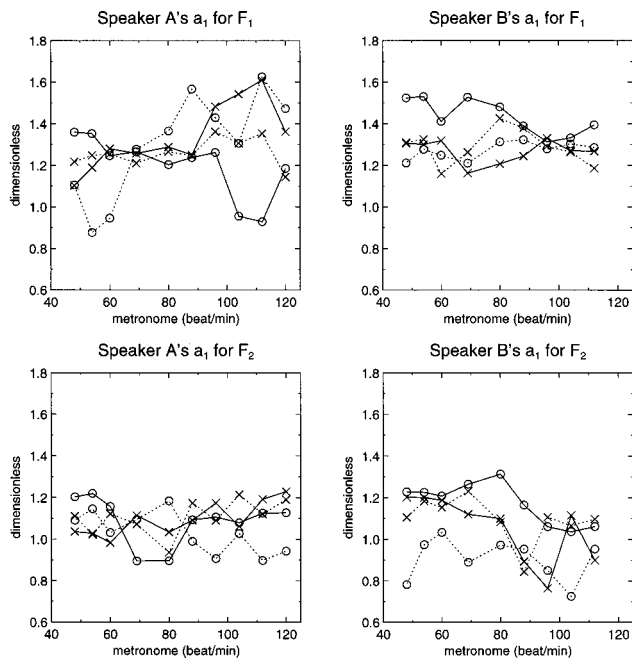


FIG. 5. Medians of parameter  $a_1$  of the linear model (1) fitted to the first-half formant transitions of speakers' [iε] or [ia] for each combination of speaking rate and contrastive stress. The data are plotted as functions of speaking rate measured in metronome rhythm: "o" stressed vowels, "x" unstressed vowels, solid lines for [a], and dotted lines for [ε]. Low speaking rates are on the left-hand side of each graph.

converge to a better local minimum or the global one because the search is better guided by the fitting procedure. For this reason, a second-order linear model at critical damping was also fitted to the transitions to follow Moon and Lindblom's (1994) suggestion.

In the two selected linear models (1) and (3),  $a_1$  indicated how fast a transition left an asymptote or came close to it and was dampened near it. The lower  $a_1$ , the faster a formant left or approached its asymptotic value and was dampened. It may therefore be expected that fast transitions were modeled with a low  $a_1$  value.

Figures 5 and 6 show the medians of the  $a_1$  estimates of the first-order linear model for each set of 30 repetitions of a corpus item for each combination of speaker, vowel, formant, contrastive stress, and speaking rate. The data are presented as functions of speaking rate measured in metronome rhythm. Tables IV and V show the results of two-factor (stress and rate) analyses of variance for all combinations of speaker, vowel, and formant. Figure 5 and Table IV refer to the first-half transitions, Fig. 6 and Table V to the second half.

The results produced by means of the two linear models (1) and (3) were slightly different in their details, but the conclusions of all analyses were the same. Therefore, only the first-order linear model will be discussed in the rest of the paper.

The exponential (2) and the kinematic model (4) were also fitted to the formant half-transitions. Since the parameters  $a_1$  of the dynamic linear models were linked to their kinematic equivalents by  $a_1 = \exp(-q\Delta t)$ , this value was computed for the kinematic models in order to compare the dynamic and kinematic results. Because the results were vir-

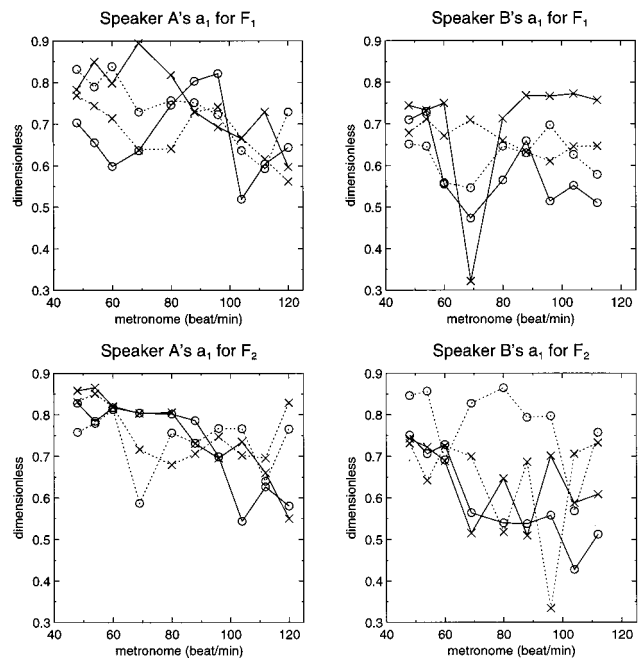


FIG. 6. Medians of parameter  $a_1$  of the linear model (1) fitted to the second-half formant transitions of speakers' [iε] or [ia] for each combination of speaking rate and contrastive stress. The data are plotted as functions of speaking rate measured in metronome rhythm: "o" stressed vowels, "x" unstressed vowels, solid lines for [a], and dotted lines for [ε]. Low speaking rates are on the left-hand side of each graph.

tually identical and the kinematic models did not add any new information, those kinematic models will not be considered further in this paper.

Up to this point, only half-transition modeling has been presented because low-order linear models with constant parameters could not describe a full transition. The logistic models (5) and (6) permitted complete transition modeling.

In the logistic models,  $d$  of (5) and  $b$  of (6) played a similar role to  $a_1$  and  $q$  of the linear (1) and exponential (2) models, and the relation  $d = \exp(-b\Delta t)$  was valid. In other words,  $d$  or  $b$  controlled the models when  $t \rightarrow \pm \infty$ , i.e., they described how a formant left its first asymptote and how it was dampened near the second one.

Figure 7 shows the medians of the  $d$  values of each set of 30 repetitions of a corpus item for each combination of speaker, vowel, formant, contrastive stress, and speaking rate. The data are plotted as functions of speaking rate measured in metronome rhythm. A smaller  $d$  value means that the formant left the initial value and was dampened near the final one more abruptly. Table VI shows the results of two-factor (rate and stress) analyses of variance for all combinations of speaker, vowel and formant.

Once again, dynamic or kinematic modeling produced approximately the same results. The kinematic ones were therefore not included in this paper.

Because a second-half transition may be more influenced by following gestures than the first half, the logistic models was also fitted to the first-half transitions. Modeling a first-half or a full transition led to similar results, except that there was more noise in the fits with the first half. However, this increase in noise was expected because fewer points were included in the fits. The tables are not presented, be-

TABLE IV. Results of analyses of variance of  $a_1$  [linear model (1)] estimates fitted to the first-half formant transitions for speaking rate and contrastive stress factors. One analysis was carried out for each combination of formant, speaker, and vowel. Labels “DF,” “S:R,” and “Resid” stand for degree of freedom, interaction between stress and rate, and residuals, respectively. When  $p < 1e-10$ , 0 was reported in the table.

	Speaker A						Speaker B					
	[ε]			[a]			[ε]			[a]		
	DF	<i>F</i>	<i>p</i>	DF	<i>F</i>	<i>p</i>	DF	<i>F</i>	<i>p</i>	DF	<i>F</i>	<i>p</i>
	$F_1$											
Stress	1	0.55	0.46	1	20.83	$6.3e-6$	1	0.67	0.41	1	68.86	0
Rate	9	0.76	0.65	9	1.98	0.040	8	5.51	$1.0e-6$	8	1.77	0.081
S:R	9	0.68	0.73	9	8.03	0	8	3.60	$4.4e-4$	8	3.41	$7.9e-4$
Resid	444			522			539			545		
	$F_2$											
Stress	1	0.51	0.47	1	0.57	0.45	1	9.23	0.0025	1	0.87	0.35
Rate	9	1.26	0.26	9	0.97	0.46	8	1.23	0.28	8	0.90	0.52
S:R	9	0.76	0.65	9	0.93	0.50	8	1.30	0.24	8	0.82	0.59
Resid	514			502			514			561		

cause they did not add any new information. The analysis simply indicated that the first- and second-half transitions were not different enough to lead to different conclusions.

Table VII shows standard linear Pearson correlations measured between the perceptual scores and the median of the formant measurements or the medians of the model parameters.

Figure 8 plots in the  $(F_1, F_2)$  plane the formant frequencies measured in the steady-state parts of the [a] and [ε] stimuli correctly identified in the perceptual test. Figure 9 shows a similar plot for the dynamic parameter  $a_1$  of the linear model (1) fitted to the second-half transitions of  $F_1$  or  $F_2$ , respectively, for the *x* or *y* axis. The two ellipses formed by each vowel in the  $(F_1, F_2)$  space are basically disjointed, while the two vowels completely overlap in the dynamic parameter space. The same overlap was found for the parameter  $a_1$  fitted to the first-half formant transitions and for the parameter *d* of the nonlinear model (5) fitted to complete formant transitions.

### III. DISCUSSION

As explained in the Introduction, fitting a dynamic or an equivalent kinematic model can lead to different results. Both fits were compared for each model, and both fits produced similar results every time. This simply means that the minimization processes of the merit functions of each dynamic and corresponding kinematic models found a minimum in the same area of their parameter spaces. Hence, both kinds of models were equivalent as far as the analysis of their fitted parameters was concerned.

Figure 4 shows that the first two formant frequencies measured at the slowest speaking rate were close to (400, 1800) Hz for both speakers’ [ε] and (600–700, 1400–1500) Hz for speakers’ [a]. Table II shows that the first two formant frequencies of the context were close to (250, 1950) Hz with small variation around those values. A contextual assimilation should therefore decrease [a] and [ε]  $F_1$ , and increase their  $F_2$  for high speaking rates. In contrast, a centralization effect should not modify the [a] formant values much, but

TABLE V. Results of analyses of variance of  $a_1$  [linear model (1)] estimates fitted to the second-half formant transitions for speaking rate and contrastive stress factors. One analysis was carried out for each combination of formant, speaker, and vowel. Labels “DF,” “S:R,” and “Resid” stand for degree of freedom, interaction between stress and rate, and residuals, respectively. When  $p < 1e-10$ , 0 was reported in the table.

	Speaker A						Speaker B					
	[ε]			[a]			[ε]			[a]		
	DF	<i>F</i>	<i>p</i>	DF	<i>F</i>	<i>p</i>	DF	<i>F</i>	<i>p</i>	DF	<i>F</i>	<i>p</i>
	$F_1$											
Stress	1	3.42	0.065	1	0.75	0.39	1	2.48	0.12	1	51.80	0
Rate	9	2.35	0.013	9	1.20	0.29	8	1.29	0.24	8	18.12	0
S:R	9	1.99	0.038	9	1.06	0.39	8	2.67	0.0069	8	8.88	0
Resid	570			569			548			538		
	$F_2$											
Stress	1	2.89	0.089	1	0.88	0.35	1	0.46	0.50	1	3.11	0.079
Rate	9	1.35	0.21	9	1.80	0.066	8	1.62	0.12	8	3.90	$1.8e-4$
S:R	9	0.99	0.45	9	1.57	0.12	8	0.85	0.56	8	1.72	0.091
Resid	554			596			385			448		

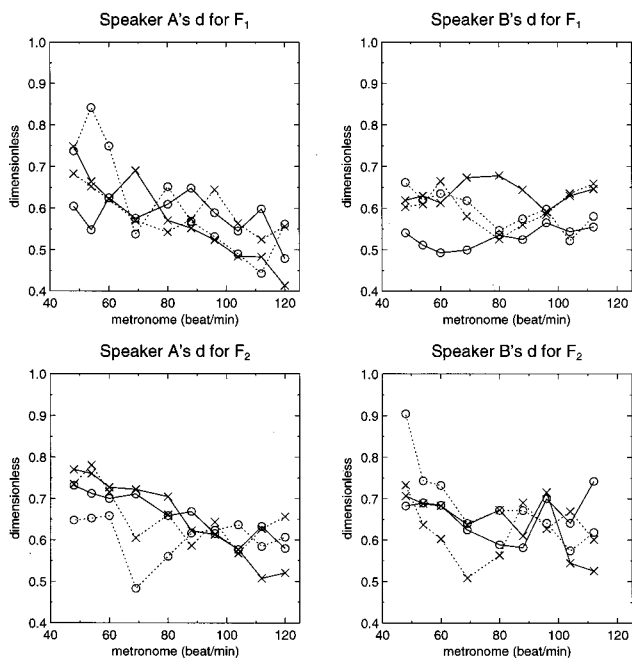


FIG. 7. Medians of parameter  $d$  of the logistic model (5) fitted to the whole formant transitions of speakers' [iε] or [ia] for each combination of speaking rate and contrastive stress. The data are plotted as functions of speaking rate measured in metronome rhythm: "o" stressed vowels, "x" unstressed vowels, solid lines for [a], and dotted lines for [ε]. Low speaking rates are on the left-hand side of each graph.

should increase [i] and [ε]  $F_1$ , and decrease their  $F_2$ . Figure 4 shows a contextual assimilation for both speakers because [ε] and [a] formant frequencies were shifted away from a central position toward [i] formant values for high speaking rates. Table III shows that the dependence of the formants on speaking rate and contrastive stress was statistically significant at the 0.05 level in all cases.

This conclusion should not be generalized to any phoneme and context. The corpus was purposely chosen to make [i] stable and [ε] and [a] varying. First, French [i] is known to need more articulatory accuracy and to have more stable formants than other vowels (Majid *et al.*, 1987). Second, the

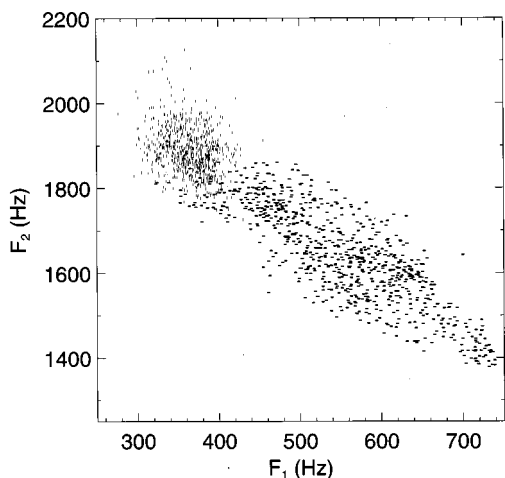


FIG. 8. Formant measurements for the [a] and [ε] correctly identified in the perceptual experiment. The | signs correspond to the [ε], the — to the [a]. The axes are linearly graduated in Hz.

carrier sentence [iliViimedia] was built to favor the context instead of V.

The speakers did not manage their contrastive stress in the same way. First, Fig. 4 shows that speaker A's formants were less influenced by contrastive stress than speaker B's, specially for [a] (each pair of solid or dotted line was closer for speaker A than B). Second, speaker A's production seemed more regular than speaker B's. This can be seen in Fig. 4 because speaker B's unstressed [a] uttered at the six highest speaking rates and those uttered at the three lowest ones seem different. The first two formants of the former were similar to the speaker's [ε] formants, while the formants of the latter were clearly distinct. The perceptual results support this hypothesis because the unstressed [a] identification scores were less than 50% correct for the six highest speaking rates, while they were equal to 100% in the other three cases. On the other hand, speaker B's stressed [a] formants were shifted more gradually toward the contextual formant values when speaking rate was increased. This pattern suggests that speaker B did not produce his stressed and unstressed [a] in the same way, like the other speaker did.

It is common to find data in the literature showing that speakers can increase, keep constant, or decrease their articulatory strength when they stress a syllable or speak faster (Gay *et al.*, 1974; Kuehn and Moll, 1976; Sonoda, 1987; Guenther, 1995; Marchal and Meynadier, 1996). Figure 6 suggests that the parameter  $a_1$  of the linear model (1) fitted to the second-half transitions tended to be smaller for high speaking rates but did not exhibit any consistent trend for contrastive stress. This result means that formants approached their  $F_1$  maximum or  $F_2$  minimum faster for high speaking rates. Even if the trend was consistent, Table V shows that it was seldom significant at the 0.05 level. It could mean that the trend was partly masked by the noise and that processing a bigger corpus would confirm the results. Figure 5 shows that the results regarding the first-half transitions were less regular.

Figure 7 shows that modeling whole transitions by means of the logistic model (5) did not change the previous trend, although this trend appeared less clearly. The parameter  $d$  was smaller for high speaking rates and the results were less clear for stress. This means that the formants tended to leave their initial values and reached their final values more abruptly for high speaking rates. Once again, the trend was not always significant, but consistent. The trend suggests that the speakers used more force in their articulatory gestures for high speaking rates. But, the nonlinear relationships between the articulatory and acoustical levels may be misleading and this conclusion must be confirmed by articulatory or muscle activity measurements.

Experiments showed that vowels can be better categorized by dynamic than static information in some situations (Strange, 1989b, 1989a; Casidy and Harrington, 1992; Beauteemps, 1993; Jenkins *et al.*, 1999). Table VII shows that the absolute values of the linear correlations between the perceptual scores and the medians of speakers' formants measured in the vowel steady-state parts were always greater than 0.8 for [a]. Because the perceptual scores of [ε] were always close to 100%, the correlations between those scores and the

TABLE VI. Results of analyses of variance of  $d$  [nonlinear model (5)] estimates fitted to complete formant transitions for speaking rate and contrastive stress factors. One analysis was carried out for each combination of formant, speaker, and vowel. Labels “DF,” “S:R,” and “Resid” stand for degree of freedom, interaction between stress and rate, and residuals, respectively. When  $p < 1e-10$ , 0 was reported in the table.

	Speaker A						Speaker B					
	[ε]			[a]			[ε]			[a]		
	DF	$F$	$p$	DF	$F$	$p$	DF	$F$	$p$	DF	$F$	$p$
	$F_1$											
Stress	1	1.37	0.24	1	1.23	0.27	1	2.26	0.13	1	237.16	0
Rate	9	12.70	0	9	20.63	0	8	8.02	0	8	2.68	0.0067
S:R	9	3.77	1.2e-4	9	10.89	0	8	7.57	0	8	6.51	3.9e-8
Resid	604			615			567			586		
	$F_2$											
Stress	1	14.84	1.3e-4	1	0.037	0.85	1	2.09	0.15	1	0.99	0.32
Rate	9	5.05	1.4e-6	9	18.07	0	8	1.46	0.17	8	1.17	0.32
S:R	9	2.31	0.015	9	1.68	0.091	8	0.56	0.81	8	1.05	0.40
Resid	602			616			545			568		

formant values are not informative. Only the [a] scores will be discussed.

It is interesting to find such a high correlation between the perceptual scores and the formant measurements for [a] because those measurements were made in the steady-state part of the vowels. It suggests that static information was more important than expected. To investigate this issue, the formant measurements of the 1900 vowels [a] and [ε] correctly identified in the perceptual experiment were plotted in the  $(F_1, F_2)$  plane in Fig. 8. In this acoustical representation, all but 40 vowels [a] and [ε] out of 1900 formed two ellipses without intersection, i.e., [a] and [ε] could be classified in this representation.

This result contrasts with Beautemps’ (1993) results, showing greater overlap between [a] and [ε] of a smaller but similar corpus. However, he only used 40 stimuli produced by one speaker, so his results should not be generalized.

On the other hand, my results corroborate one Peterson and Barney (1952) finding. They showed many measurements of the first two formants of ten English vowels in the  $(F_1, F_2)$  plane, and the vowels overlapped in this space. Then, they carried out a perceptual experiment presenting the vowels with minimal context to listeners. When only the tokens correctly identified in the perceptual assignment were

plotted in the  $(F_1, F_2)$  plane, the overlap between the vowels almost disappeared. Similarly, as long as a vowel in my corpus was correctly perceived, the first two formants were sufficient to classify it.

To summarize the perceptual results so far, static information was sufficient to explain them. But, we still do not know if dynamic information could be used by the listeners to identify the vowels in the perceptual task. Figure 9 shows that the parameter  $a_1$  of the linear model (1) fitted to the second-half formant transitions could not distinguish the [a] from [ε] speech signals for the subset of stimuli correctly identified in the perceptual assignment. The same results were also found for the first-half transitions and for the parameter  $d$  of the nonlinear model (5) fitted to whole transitions. This suggests that the part of formant dynamics encoded in the model parameters did not play a major role in the perceptual task, i.e., the models did not isolate an important dynamic parameter for the perception.

Furthermore, the linear correlations between the perceptual scores and the medians of the dynamic parameter estimates are shown in Table VII. The correlations are lower and less consistent for the model parameters than formant measurements.

It is not perfectly clear why the dynamic parameters did

TABLE VII. Standard linear Pearson correlations between the medians of the formant measurements or the medians of the model parameters and the perceptual results for [a]. The medians are shown in Figs. 4, 5, 6, and 7; the perceptual scores are shown in Fig. 3. Labels “u” or “s” stand for unstressed or stressed [a], “ $F_i$ ” for the  $i$ th formant, “ $a_1(1)$ ” or “ $a_1(2)$ ” for the parameter  $a_1$  of the linear model (1) estimated on the first- or second-half transitions, and  $d$  for the parameter  $d$  of the logistic model (5). Correlations printed in bold were significantly different from 0 at the 0.05 level according to a two-tail standard Pearson sample correlation test. The number of degrees of freedom were 8 and 7, respectively, for speakers A and B.

	$F_1$				$F_2$			
	Speaker A		Speaker B		Speaker A		Speaker B	
	u	s	u	s	u	s	u	s
$F_i$	<b>0.82</b>	<b>0.87</b>	<b>0.90</b>	<b>0.89</b>	<b>-0.85</b>	<b>-0.87</b>	<b>-0.94</b>	<b>-0.81</b>
$a_1(1)$	-0.52	<b>0.84</b>	0.35	<b>0.74</b>	<b>-0.73</b>	-0.21	<b>0.75</b>	<b>0.86</b>
$a_1(2)$	<b>0.79</b>	0.46	0.02	0.49	<b>0.89</b>	<b>0.96</b>	<b>0.68</b>	0.53
$d$	<b>0.79</b>	0.51	-0.22	<b>-0.78</b>	<b>0.83</b>	<b>0.82</b>	<b>0.68</b>	-0.56

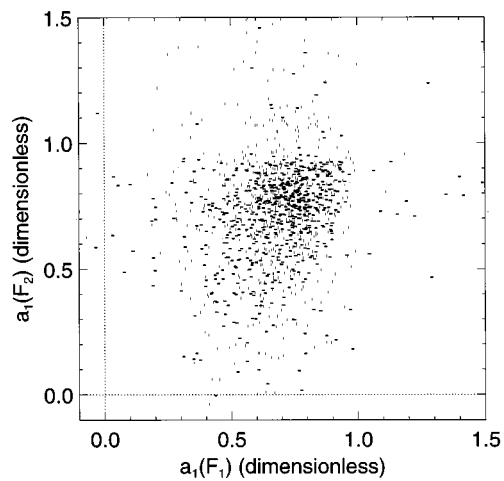


FIG. 9. Parameter  $a_1$  of the linear model (1) fitted to the second-half  $F_1$  or  $F_2$  transitions of the [ia] or [ie] stimuli correctly identified in the perceptual experiment. The  $x$  axis corresponds to  $F_1$ , the  $y$  axis to  $F_2$ , the | signs to [ε] and — to [a] stimuli.

not capture any feature correlated with the perceptual results. A first hypothesis is that the experiment could correspond to a situation where static information was primarily used. A second hypothesis is that the dynamic models inadequately described formant dynamics or that the fitting procedures were inaccurate. Visual inspection of modeled and real formant contours suggested that dynamics was properly described by the models (see an example in Fig. 4), but future work should include a study of modeling error. Because the formant values measured in the steady-state part of the vowels were sufficient to explain the perceptual results, it seems plausible that formant dynamics could have played a secondary role in the perceptual assignment.

In summary, the first two formants measured in the steady-state parts of the vowels were highly correlated with the perceptual results and were sufficient to explain them. Conversely, the dynamic parameters of the models were poorly correlated with the perceptual results and were not sufficient to explain them. This does not imply that dynamic information was not used in the perceptual task, but it does not back up dynamic specification of the vowels.

#### IV. CONCLUSION

A corpus of [iai] and [iei] was produced by two speakers in a carrier sentence at different speaking rates with two contrastive stress patterns. The corpus was perceptually analyzed, the first two formant frequencies were measured in the static parts of [a] or [ε], and the formant time series were dynamically and kinematically modeled. Relationships between the perceptual results, the formant measurements, and the model parameters were assessed.

The analysis of the first two formant frequencies measured in the steady-state part of the vowels showed a contextual assimilation instead of a centralization effect. The same data also suggest that the speakers did not manage their contrastive stress in the same way.

The analysis of model parameters showed that dynamic or kinematic modeling led to the same results. The formant

transitions were sharper for a high speaking rate, but no systematic trends were found for contrastive stress.

Relationships between the perceptual scores and the formant measurements or model parameters did not support the importance of dynamic specification of vowels.

#### ACKNOWLEDGMENTS

This research was mainly supported by a grant from the Fondation Van Buuren and mainly carried out in the Institute of Modern Languages and Phonetics in Bruxelles. Some parts of the work were also done in the Institute of Phonetics in Aix-en-Provence with the support of a TMR grant from the European Community, and at Queen's University in Kingston, Canada with the support of a Natural Sciences and Engineering Research Council grant from the Government of Canada No. 388 429 43. I am grateful to Jean Caelen and Jean-Luc Schwartz for selecting the subject and for their valuable suggestions. I also thank Jean Schoentgen for his helpful discussions, as well as Anders Lofqvist and three anonymous referees for their important suggestions. I express my gratitude to Kevin Munhall and Jeff Jones for their pertinent and constructive comments and for translating my personal English into a real language.

- Andruski, J. E., and Nearey, T. M. (1992). "On the sufficiency of compound target specification of isolated vowels and vowels in /bVb/ syllables," *J. Acoust. Soc. Am.* **91**, 390–410.
- Atal, B. S., and Hanauer, S. L. (1971). "Speech analysis and synthesis by linear prediction of the speech wave," *J. Acoust. Soc. Am. Suppl.* **1** **50**, 637–655.
- Beautemps, D. (1993). "Récupération des gestes de la parole à partir de trajectoires formantiques: Identification de cibles vocaliques non-atteintes et modèles pour les profils sagittaux des consonnes fricatives," Ph.D. thesis, INPG, Grenoble.
- Broad, D. J., and Clermont, F. (1987). "A methodology for modeling vowel formant contours in CVC context," *J. Acoust. Soc. Am.* **81**, 155–165.
- Carré, R., Chennoukh, S., Jospa, P., and Maeda, S. (1996). "The ears are not sensitive to certain coarticulatory variations: Results from VCV synthesis/perceptual experiments," in *1st ESCA Tutorial and Research Workshop on Speech Production Modeling: From Control Strategies to Acoustics* (European Speech Communication Association, Autrans, France), pp. 13–16.
- Carré, R., Chennoukh, S., Lindblom, B., and Divenyi, P. (1994). "On the perceptual characteristics of 'speech gestures,'" *J. Acoust. Soc. Am.* **96**, 3326(A).
- Carré, R., and Mrayati, M. (1992). "Distinctive regions in acoustic tubes, speech production modelling," *J. Acoust.* **5**, 141–159.
- Cassidy, S., and Harrington, J. (1992). "Investigating the dynamic structure of vowels using neural networks," in *Proceedings of the Fourth Australian International Conference on Speech Science and Technology* (Australian Speech Science and Technology Association, University of Queensland, Brisbane, Australia), pp. 495–500.
- Delattre, P., Liberman, A. M., Cooper, F. S., and Gerstman, L. J. (1952). "An experimental study of the acoustic determinants of vowel color: observations on one- and two-formant vowels synthesized from spectrographic patterns," *Word* **8**, 195–210.
- Divenyi, P., Lindblom, B., and Carré, R. (1995). "The role of transition velocity in the perception of  $V_1V_2$  complexes," in *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Department of Speech Communication and Music Acoustics, KTH (Royal Institute of Technology) and the Department of Linguistics, Stockholm, University, Stockholm, Sweden), pp. 258–261.
- Gay, T., Ushijima, T., Hirose, H., and Cooper, F. S. (1974). "Effect of speaking rate on labial consonant-vowel articulation," *J. Phonetics* **2**, 47–63.

- Guenther, F. H. (1995). "Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production," *Psychol. Rev.* **102**(3), 594–621.
- Helmholtz, H. F. (1885). *On the Sensations of Tone*, translated by A. J. Ellis (Dover, New York, reprinted in 1954).
- Imaizumi, S., and Kiritani, S. (1989). "Effect of speaking rate on formant trajectories and inter-speaker variations," *Ann. Bull. RILP* **23**, 27–37.
- Jenkins, J. J., Strange, W., and Trent, S. A. (1999). "Context-independent dynamic information for the perception of coarticulated vowels," *J. Acoust. Soc. Am.* **106**, 438–448.
- Joos, M. (1948). "Acoustic phonetics," in *Language Monograph No. 23*, edited by L. S. of America (Waverly, Baltimore).
- Koopmans-van Beinum, F. J. (1980). "Vowel contrast reduction, an acoustic and perceptual study of Dutch vowels in various speech conditions," Ph.D. thesis, University of Amsterdam.
- Kuehn, D. P., and Moll, K. L. (1976). "A cineradiographic study of VC and CV articulatory velocities," *J. Phonetics* **4**, 303–320.
- Kuwabara, H. (1985). "An approach to normalization of coarticulation effects for vowels in connected speech," *J. Acoust. Soc. Am.* **77**, 686–694.
- Laboissière, R., Ostry, D., and Perrier, P. (1995). "A model of human jaw and hyoid motion and its implications for speech production," in *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Department of Speech Communication and Music Acoustics, KTH (Royal Institute of Technology) and the Department of Linguistics, Stockholm University, Stockholm, Sweden), pp. 60–67.
- Lindblom, B. E. F. (1963). "Spectrographic study of vowel reduction," *J. Acoust. Soc. Am.* **35**, 1773–1781.
- Lindblom, B. E. F., and Studdert-Kennedy, M. (1967). "On the role of formant transitions on vowel recognition," *J. Acoust. Soc. Am.* **42**, 830–843.
- Loevenbruck, H., and Perrier, P. (1993). "Vocalic reduction: Prediction of acoustic and articulatory variabilities with invariant motor commands," in *Proceedings of Eurospeech'93*, Vol. 1, Catalyst Consult (ESCA, Berlin, Germany), pp. 85–88.
- Majid, R., Abry, C., Boe, L.-J., and Perrier, P. (1987). "Fonctions de macrosensibilité articulatoire-acoustique, pour le système vocalique du français, à partir d'un modèle articulatoire du conduit vocal," *Bulletin du Laboratoire de la Communication Parlée de Grenoble 1A*, pp. 133–157.
- Makhoul, J. (1975). "Linear prediction: A tutorial review," *Proc. IEEE* **63**(4), 561–580.
- Marchal, A., and Meynadier, Y. (1996). "Coarticulation in /kl/ sequences in French: A multisensor investigation of the timing of lingual gestures," in *Proceedings of the 6th Australian International Conference on Speech Science and Technology* (Australian Speech Science and Technology Association, Inc., Adelaide, Australia), pp. 434–438.
- Moon, S.-J., and Lindblom, B. (1994). "Interaction between duration, context, and speaking style in English stressed vowels," *J. Acoust. Soc. Am.* **96**, 40–55.
- Nábělek, A. K., and Ovchinnikov, A. (1997). "Perception of nonlinear and linear formant trajectories," *J. Acoust. Soc. Am.* **101**, 488–497.
- Nearey, T. M. (1989). "Static, dynamic, and relational properties in vowel perception," *J. Acoust. Soc. Am.* **85**, 2088–2113.
- Nearey, T. M., and Assmann, P. F. (1986). "Modeling the role of inherent spectral change in vowel identification," *J. Acoust. Soc. Am.* **80**, 1297–1308.
- Peterson, G., and Barney, H. (1952). "Control methods used in a study of vowels," *J. Acoust. Soc. Am.* **24**, 175–184.
- Pitermann, M. (1996). "Évaluation expérimentale de la théorie des cibles formantiques dans le cadre de la production des voyelles," Ph.D. thesis, Université Libre de Bruxelles.
- Polka, L., and Bohn, O.-S. (1996). "A cross-language comparison of vowel perception in English-learning and German-learning infants," *J. Acoust. Soc. Am.* **100**, 577–592.
- Pols, L. C. W., and van Son, R. J. J. H. (1993). "Acoustics and perception of dynamic vowel segments," *Speech Commun.* **13**, 135–147.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C*, 2nd ed. (Cambridge University Press, Cambridge), Chap. 15, pp. 681–688.
- Sonoda, Y. (1987). "Effect of speaking rate on articulatory dynamics and motor event," *J. Phonetics* **15**, 145–156.
- Soquet, A. (1995a). "A cooperative approach to formant extraction," in *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Department of Speech Communication and Music Acoustics, KTH (Royal Institute of Technology) and the Department of Linguistics, Stockholm University, Stockholm, Sweden), pp. 448–451.
- Soquet, A. (1995b). "Étude comparée de représentations acoustiques et articulatoires du signal de parole pour le décodage acoustico-phonétique. Application à la classification de voyelles et à la détermination du lieu d'articulation des occlusives," Ph.D. thesis, Université Libre de Bruxelles.
- Spiegel, M. R. (1988). *Statistics*, 2nd ed., Schaum's Outline Series (McGraw-Hill, New York).
- Stevens, K. N., and House, A. S. (1963). "Perturbation of vowel articulations by consonantal context: An acoustical study," *J. Speech Hear. Res.* **6**(2), 111–128.
- Strange, W. (1989a). "Dynamic specification of coarticulated vowels spoken in sentence context," *J. Acoust. Soc. Am.* **85**, 2135–2153.
- Strange, W. (1989b). "Evolving theories of vowel perception," *J. Acoust. Soc. Am.* **85**, 2081–2087.
- Strange, W., and Bohn, O.-S. (1998). "Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies," *J. Acoust. Soc. Am.* **104**, 488–504.
- van Son, R. J. J. H. (1993). "Vowel perception: A closer look at the literature," in *Proceedings of the Institute of Phonetic Sciences* (University of Amsterdam, Amsterdam), Vol. 17, pp. 33–64.
- van Son, R. J. J. H., and Pols, L. C. W. (1992). "Formant movements of Dutch vowels in a text, read at normal and fast rate," *J. Acoust. Soc. Am.* **92**, 121–127.
- Viterbi, A. (1967). "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory* **IT-13**, 260–269.
- Watson, C. I., and Harrington, J. (1999). "Acoustic evidence for dynamic formant trajectories in Australian English vowels," *J. Acoust. Soc. Am.* **106**, 458–468.



# Acoustic effects of variation in vocal effort by men, women, and children

Hartmut Traunmüller<sup>a)</sup> and Anders Eriksson

Department of Linguistics, Stockholm University, S-106 91 Stockholm, Sweden

(Received 7 September 1999; accepted for publication 23 February 2000)

The acoustic effects of the adjustment in vocal effort that is required when the distance between speaker and addressee is varied over a large range (0.3–187.5 m) were investigated in phonated and, at shorter distances, also in whispered speech. Several characteristics were studied in the same sentence produced by men, women, and 7-year-old boys and girls: duration of vowels and consonants, pausing and occurrence of creaky voice, mean and range of  $F_0$ , certain formant frequencies ( $F_1$  in [a] and  $F_3$ ), sound-pressure level (SPL) of voiced segments and [s], and spectral emphasis. In addition to levels and emphasis, vowel duration,  $F_0$ , and  $F_1$  were substantially affected. “Vocal effort” was defined as the communication distance estimated by a group of listeners for each utterance. Most of the observed effects correlated better with this measure than with the actual distance, since some additional factors affected the speakers’ choice. Differences between speaker groups emerged in segment durations, pausing behavior, and in the extent to which the SPL of [s] was affected. The whispered versions are compared with the phonated versions produced by the same speakers at the same distance. Several effects of whispering are found to be similar to those of increasing vocal effort. © 2000 Acoustical Society of America.

[S0001-4966(00)01306-0]

PACS numbers: 43.70.Gr [AL]

## INTRODUCTION

The acoustic properties of speech sounds vary not only with their linguistic phonetic value, but also as a function of expressive, organic, and perspectival factors. The present investigation is concerned with the effects of expressive variation, mainly in vocal effort, but also in mode of phonation (whispering vs phonating). These effects are investigated in the speech of men, women, and children, so that some information about effects of organic variation associated with speaker age and sex is also obtained.

Variation in vocal effort obviously affects the sound pressure  $p$  of speech signals, but it also has substantial effects on several additional acoustic properties. “Vocal effort” is basically a subjective physiological quantity that has to be kept distinct from “voice level.” The latter could be defined as the sound-pressure level (SPL) of a standard utterance as measured at a given distance from the speaker’s mouth. However, there are alternative ways in which speakers can control the level of their voice. It is well-known that trained singers and actors are capable of doing it without much affecting  $F_0$ , the fundamental frequency of the voice. They vary their pulmonic pressure more than ordinary speakers do (Åkerlund and Gramming, 1994). However, in ordinary speech communication, there is no need for keeping the pitch at a constant level, and when these techniques are used in speaking, they lend a theatrical quality to the speech. Ordinary speakers can do with smaller variations in pulmonic pressure by using more variation in the tension of their vocal folds in controlling their vocal effort. This economizes the airflow, but it affects  $F_0$ .

Within the framework of the present investigation, it was essential to define and quantify vocal effort in a way that is appropriate for ordinary speech. According to the definition adopted, *vocal effort is the quantity that ordinary speakers vary when they adapt their speech to the demands of an increased or decreased communication distance.* “Ordinary,” in this sense, are speakers who are free from any relevant pathology and who have not received any special training of their voice.

In some investigations of the aerodynamics of glottal airflow, such as that by Holmberg, Hillman, and Perkell (1988), subjects were asked to adjust their “loudness level” in order to produce syllables with normal, loud, and soft voice. The observed variation in  $F_0$  was not very large. However, in this situation it is questionable whether the speakers controlled their vocal effort in a way consistent with our definition, since they may have adopted the singer’s technique to some extent.

Very large effects on  $F_0$  have been observed in investigations of shouted vs normal speech (e.g., Rostolland, 1982a). Increases in  $F_0$  were also observed as an effect of speaking in noise (Rastatter and Rivers, 1983; Loren, Colcord, and Rastatter, 1986; Summers *et al.*, 1988; Bond, Moore, and Gable, 1989; Junqua, 1993). Speakers react to increased or decreased environmental noise (the Lombard effect) similarly to the way they react to an increased or decreased communication distance, but we must not assume complete equivalence. Speakers’ responses to noise with a given sound pressure might differ according to its spectrotemporal properties. Rivers and Rastatter (1985) actually observed such differences between the effects of white and multitalker noise.

Variation in vocal effort is known to affect the shape of

<sup>a)</sup>Electronic mail: hartmut@ling.su.se

the glottal pulses, i.e., the course of the airflow volume velocity and pressure as a function of time within a glottal cycle (Holmberg *et al.*, 1988; Holmberg *et al.*, 1995; Södersten, Hertegård, and Hammarberg, 1995). The vocal-fold closing velocity and the relative duration of the closed interval of the glottal cycle was found to be higher in loud, as compared with the normal, voices. We know from other investigations (Gauffin and Sundberg, 1989; Childers and Lee, 1991) that in the spectral domain, this is reflected in an emphasis of the higher partials. The level of the lowest partials, in particular that of the first, does not vary as much when speakers vary their vocal effort.

It has also been reported that formant frequencies, in particular  $F_1$ , increase with increasing vocal effort (Frøkjær-Jensen, 1966; Rostolland and Parant, 1974; Rostolland, 1982b; Schulman, 1985; Bond *et al.*, 1989; Junqua, 1993; Liénard and Di Benedetto, 1999; Huber *et al.*, 1999). However, measurements of formant frequencies at high  $F_0$ s are often not very reliable, since they tend to be affected by a bias towards the partial that is closest to the resonance peak. In any case, it is well-known (Schulman, 1989) that speakers open their mouths more when shouting than when speaking with more moderate vocal effort, and the observed formant frequencies, in particular the substantially increased values for  $F_1$ , agree with what can be expected for such an articulation on the basis of the acoustic theory of speech production (Fant, 1960; Lindblom and Sundberg, 1971).

In the investigations by Fónagy and Fónagy (1966), Rostolland (1982a), and Bonnot and Chevrie-Muller (1991), the duration of vowels increased with vocal effort, while the duration of consonants decreased somewhat.

Since the sound pressure of an utterance is the acoustic variable that is affected most by variations in vocal effort, everything else being equal, it can be used as an index of vocal effort, as in the recent study by Liénard and Di Benedetto (1999), who had evoked variations in vocal effort by varying the distance between speaker and interlocutor. However, perception experiments (Rundlöf, 1996; Traunmüller, 1997; Eriksson and Traunmüller, 1999) have demonstrated that the SPL actually plays no major part in judgments of vocal effort. The sound pressure varies widely as a function of perspectival factors such as the distance between a listener and the source of the sound, which often does not substantially affect anything else. Listeners are, nevertheless, capable of distinguishing variations in a speaker's vocal effort from variations in their own distance from the speaker. This was already implied by the results of Wilkens and Bartel (1977), who showed that listeners were capable of recreating the original SPL of a speaker with high precision from a recording. More recently, Eriksson and Traunmüller (1999) showed that listeners are capable of separating and judging the two variables on the basis of single vowels.

It would, therefore, be preferable to have a more generally valid measure of vocal effort. The measure used in the analysis of the results of the present experiments is based on the average rating, by a group of listeners, of the communication distance for each stimulus.

While previous research has produced useful knowledge about various acoustic effects of variations in vocal effort, it

is our ambition to obtain a description that captures the whole set of effects that might be of importance for its perception. Detailed quantitative knowledge of these effects is required in order to be able to simulate human behavior and capabilities in speech synthesis and in automatic recognition of the various kinds of information that are transmitted by speech. In addition to the linguistic information (what is said), this includes expressive information (how it is said), organic information (by whom it is said), and perspectival information (where it is said). Within this scheme, vocal effort is the most prominent expressive variable, while sound pressure  $p$  is more closely associated with perspectival information. Temporal variations in vocal effort also serve the expression of linguistic stress (Sluijter and van Heuven, 1996).

The present descriptive study was designed to provide data against which general theories of speech communication, such as the modulation theory (Traunmüller, 1994, 1998), which must handle the production and perception of all kinds of information transmitted by speech, can be tested. The acoustic characteristics that will be considered are those that can be expected to be most important for the perception of this information: segment durations,  $F_0$ , formant frequencies, SPLs, and spectral emphasis. No attempt is made here to study periods of glottal excitation in the temporal domain.

## I. METHOD

### A. Subjects

Twenty subjects served as speakers: six male adults, aged 20–51 years, six female adults, aged 20–38 years, as well as four boys and four girls, all 7 years of age. The adult subjects were students and staff members at the department of linguistics. The children were from a primary school close to the recording site. All subjects were speakers of Stockholm Swedish.

None of the subjects had any known speech or hearing disorder. Some of the children may have had irregularities in their dentition. While this was not checked in detail, their speech was not noticeably affected by such a factor. However, the speech of 7-year-olds is typically not yet developed to full maturity and shows dysfluencies and individual deviations from the adult norm. Such deviations were also present in the speech of these children. None of the subjects was a trained singer or had been subject to any extensive voice training.

### B. Speech material

The speech material consisted of several versions of a single sentence that had been chosen so as to provide suitable data (representative and easy to measure) of segment durations, fundamental frequencies, formant frequencies, spectral slopes, and the levels of voiced and voiceless segments. Nasals were avoided. The sentence used was *Jag tog ett violett, åtta svarta och sex vita*, [jɛtʊgɛtvɪøɛt ɔtəsvaʃa ɔsɛksvi:tɑ] ‘I took one violet, eight black, and six white.’ Each subject produced seven versions of this sentence, two

TABLE I. Distances between speaker and addressee. The full range was used for phonated speech. Whispered speech was used only at the two shortest distances.

Version	1	2	3	4	5
Distance (m)	0.3	1.5	7.5	37.5	187.5

whispered and five phonated at increasing levels of vocal effort in order to bridge an increasing communication distance.

One of the children consistently said “*viloett*” instead of “*violet*” and another on two occasions said “*en*” where “*ett*” would be required. Some of the speakers, including several adults, skipped the initial [j] of the utterance. On the other hand, there were also speakers who realized the final [g] in the word “*jag*,” which most speakers drop even in their clearest style of spontaneous speech.

### C. Recording procedure

The experiment took place in an open field in a reasonably quiet area on the island of Lidingö outside Stockholm. The surface of the field was soft soil covered with grass. Beside some trees at a distance of at least 25 m behind the speakers, there were no other nearby objects that might have caused noticeable reflections. The utterances were recorded on a portable DAT recorder, using a small, high-quality microphone (Sennheiser Red Dot MKE2) that was provided with a cover to reduce wind noise. A stepwise attenuator was used to keep variation in recording levels within tolerable limits and under control. The steps were 0, 8, 16, 24, and 32 dB. The microphone had been fitted to a fixture on a headphone hoop with the headphones removed, and it was adjusted to remain at a horizontal distance of 50 mm from the speaker’s upper lip and approximately 10 mm above it, with its center line pointing towards the mouth opening. This was ensured by using a plastic gauge specifically manufactured for this purpose. A male experimenter gave the subjects instructions as to whether the utterance was to be phonated or whispered, and controlled the attenuator and recorder. A female experimenter served as the addressee of each subject. She asked each subject “*Hur många kort tog du av varje färg?*” “How many cards of each color did you take?,” standing at five previously measured distances from the subject (see Table I) and adapting her vocal effort to the communication distance. The subjects were instructed to produce each utterance with the vocal effort considered necessary for the specific distance. For an utterance to be satisfactory, the addressee had to accept it as clearly audible against the background noise and fully intelligible.

For all speakers, the first condition was communicating over the shortest distance, and the distance was subsequently increased in four steps. At the distances of 0.3 and 1.5 m, whispered versions were produced after the voiced, but not in immediate succession. Between all versions, the male experimenter stated the distance, the desired mode of phonation, and the chosen attenuation. In asking her question, the addressee used in each case the same mode of phonation that was required from the subjects.

All recordings followed the same procedure, and the addressee was always the same person. There were several recording occasions, and the children were actually recorded 1 year later than the adults, in autumn 1995 and 1994, respectively. On all occasions, it was slightly windy.

Due to a mistake, one utterance by an adult male speaker went unrecorded, and one by a child was skipped because it was affected by extraneous noise. This left 138 utterances to be analyzed.

### D. Acoustic measurements

The utterances were digitized at 16 kHz, using 16 bits per sample. In order to remove some of the low-frequency wind noise that was present in most of the recordings, the signals were high-pass filtered at 70 Hz, 48 dB/octave. Most of the acoustic analyses were done using the ESPS/WAVES program package (Entropic Research Lab.).

#### 1. Sound-pressure levels ( $SPL_v$ , $SPL_s$ )

Sound-pressure levels were measured, and the data found to be most informative were the levels of selected segment types within each utterance: the voiced segments ( $SPL_v$ ) or the potentially voiced segments of the whispered versions, and the three [s]-es ( $SPL_s$ ). In order to make these measurements, two excerpts of each utterance were produced by concatenating the relevant segments of each original utterance, excluding everything else. These concatenations were initially based on markings of segment boundaries that had been used for the measurement of segment durations (see below), but they were edited further to remove some audible traces of inappropriate acoustic features such as traces of voicing in some of the [s] segments.

#### 2. Spectral emphasis (*Emph*)

The concatenated voiced segments of each phonated utterance were low-pass filtered with a cutoff frequency of  $1.5 F_{0\text{mean}}$ , 18 dB/octave, and the average SPL of the low-pass-filtered signal ( $SPL_0$ ) was measured. A measure of spectral emphasis was then obtained by calculating the difference,  $Emph = SPL_v - SPL_0$ . This measure is equal to zero when partials above the first are totally absent and it is +3 dB when there are equal amounts of energy below and above  $1.5 F_{0\text{mean}}$ . Similar measures were used by Childers and Lee (1991), and Granström and Nord (1992), while Titze and Sundberg (1992) considered the difference between the second and first partial.

#### 3. Fundamental frequency ( $F_0$ )

$F_0$  was computed using the pitch extraction algorithm of the ESPS/WAVES analysis package. The results were inspected manually and only data from voiced segments were considered. Sections were excluded where the signal was too weak (e.g., at the very end of utterances) for the pitch analysis to be fully reliable. Sections with creaky voice, defined as segments beginning with an abrupt frequency halving, were treated separately. Such sections are usually not perceived as lower in pitch but rather as having a coarser voice quality. If they were included in the data from which mean values were

calculated, it would give a misleading impression of a lower pitch used by the speakers for those particular renditions. Mean values and standard deviations were calculated for each individual speaker and for each group of speakers.

#### 4. Formant frequencies ( $F_{1a}, F_3$ )

Any measurement of formant frequencies in phonated speech is hampered by the fact that the spectral envelope is only sampled at the frequencies of the harmonics. In the louder versions,  $F_0$  was often very high. This would cause serious measurement errors in the form of a bias toward the harmonic closest to the formant peak, using any of the customary formant-tracking algorithms. Therefore, average frequency positions of three formants were measured by means of a linear predictive coding (LPC) analysis in which the whole string of voiced sounds of each utterance was placed in one analysis window of the same length as the signal. In the frequency region above  $2F_0$ , this can be expected to produce fairly reliable values, since most of the mentioned bias is avoided thanks to the segmental variation in the formant frequencies and the prosodic variation in  $F_0$  within each string. However, this kind of measurement suffers from another kind of bias: The measured average value is sensitive to the variation in spectral emphasis caused by vocal effort, especially below  $2F_0$ . Therefore, the only  $F_1$  measurement that will be reported is the average frequency value of this formant in a concatenation of the four occurrences of the vowel [a] of each utterance. In these segments,  $F_1$  was always above  $2F_0$ , except for the formant transitions out of or into consonants. An average value for  $F_2$  will not be reported, since the results of previous research (see Traunmüller, 1988) suggest that its behavior in front vowels will be different from that in back vowels. An overall average value would therefore lack representativity.

In measuring formant frequencies one also encounters the additional problem of uncertainty of the formant number. Automatic methods often miss a formant or find a spurious one, but this problem is minimized by the use of a long analysis window. Using the LPC method, it is necessary to specify to the system beforehand how many formants there are. For a given sampling frequency, this number will be different for men, women, and children. In an attempt to capture exactly three formants, the utterances were resampled at the lower sampling frequencies of 6.4 kHz for men, 8 kHz for women, and 10.667 kHz for children. The three conjugate pole pairs found below half the sampling frequency for each speaker category were considered to represent the formants  $F_1$ ,  $F_2$ , and  $F_3$ . Seven reflection coefficients were used for all utterances. The preemphasis factor used was 0.92 in all analyses. For some tokens of [a] produced with a high vocal effort, the formant measurements were compared with the result of a period-by-period LPC analysis. The formant frequencies obtained in this way were close to those obtained with the simpler method.

#### 5. Segment durations ( $dur_V, dur_C$ )

Segment boundaries were marked using oscillograms and broadband spectrograms. Stop occlusions and pauses

were considered to begin where the formant structure of the preceding segment could no longer be seen. Stops were considered to end at the beginning of a following vowel-like segment or after the burst when it was followed by a pause. Vowel-like segments were considered to begin at the first positive zero crossing after a consonant or a pause and to end at the beginning of a following stop occlusion. In the following, the mean durations of vowel-like segments and consonantal segments will be referred to as ( $dur_V$ ) and ( $dur_C$ ). Pause durations will also be reported.

#### E. The measure of vocal effort

The speech material recorded for the purpose of studying the acoustic properties of speech was also used in perception experiments in which listeners had to rate the distance between speaker and addressee (Rundlöf, 1996; Traunmüller, 1997). Such ratings were obtained from two groups of 20 listeners each. One group listened to the utterances with their original variation in sound pressure  $p$ , while the other group listened to them with a random 6-dB variation imposed after equalization of  $p$ . The mean logarithmic distance ratings obtained in the latter condition were highly correlated with and similar to those obtained with the original variation in  $p$  preserved ( $r=0.993$ , slope=0.93). This shows that a listener is capable of distinguishing between vocal effort and the variations in SPL due to perspectival factors. The result was further corroborated by a later experiment (Eriksson and Traunmüller, 1999) where listeners were shown to be able to distinguish between these factors, with a reasonable degree of precision, on the basis of single-vowel utterances.

The correlation coefficient for estimated against actual distances (also logarithmic) was clearly lower ( $r=0.90$ ). This can be understood as due to some additional factors that influenced the speakers' choice of vocal effort and which will be identified in the following. These were unknown to the listeners who had to estimate the communicative distance. Therefore, they had no choice but to interpret all variation in vocal effort as due to variation in this distance. The results showed a consistent and substantial underestimation of the larger distances. The geometric means of the estimated distances were 0.47, 0.69, 1.9, 7.5, and 31 m for the actual 0.3, 1.5, 7.5, 37.5, and 187.5 m. Visual estimates of a similar range of distances by the same subjects were quite close to the actual distances (0.33, 1.54, 7.2, 34, and 159 m). The discrepancy will be explained in Sec. II C. Similar auditory underestimations were also obtained by Eriksson and Traunmüller (1999).

It can be expected that the expressive increase in vocal effort will not fully compensate for the perspectival decrease in  $p$  as a function of distance. Thus, at the ears of the addressee,  $p$  will actually decrease with increasing vocal effort, which again makes it appear inappropriate to take  $p$  as a measure of vocal effort. We shall take the geometric mean of the distance estimates  $d_e$  obtained for each utterance as heard at a constant distance from the speaker as a measure of the vocal effort invested into its production. In Sec. II C and the subsequent analyses, vocal effort level, defined as  $VEL = {}^2\log(d_e)$ , will be used as an independent variable, and the

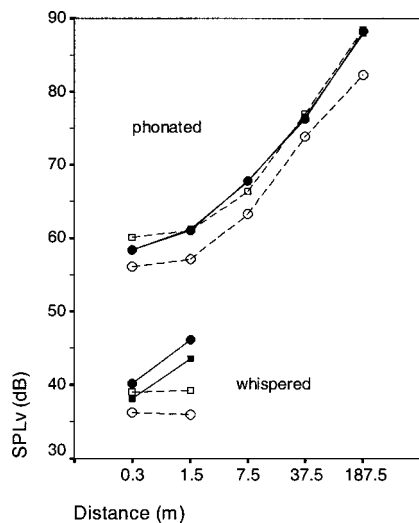


FIG. 1. The average sound-pressure level ( $SPL_v$ ), with an arbitrary reference, of the voiced and potentially voiced segments in the phonated and whispered utterances produced by men (■), women (●), boys (□), and girls (○). (Phonated versions of men and women overlap completely.)

effects of its variation on signal levels, frequencies, etc. will be analyzed. A unit increase in VEL reflects an increase in vocal effort such that listeners judge the communication distance to be increased by 100%.

## II. RESULTS AND DISCUSSION

### A. The extrinsic factors that influenced the acoustic properties

Most of the variation in vocal effort was of course due to the large variation in the distance between speaker and addressee, but there were some additional environmental and personal factors that must be considered. The factors that were recognized to be relevant are listed here below. They describe either the situation (1 to 3) or the speaker (4 to 7).

#### (1) Communication distance, ${}^2\log(d)$

The distance between speaker and addressee, to which speakers had to adapt their vocal effort, was the factor that could be expected to have the most profound influence on quantities related to signal levels. In the following analysis, this distance is expressed in base-2 logarithms,  ${}^2\log(d)$ , with  $d$  in meters. If expressed in this way, the variable was expected to be linearly related to acoustic variables such as levels in dB and frequencies expressed in octaves or semitones, assuming relationships of this kind to be best described by power functions (Stevens, 1965). In a free field, there is a 6-dB loss in sound pressure for each doubling of distance from the sound source.

#### (2) "Closeness," $e^{(1-n)}$

The expectation that the acoustic variables would be linearly related to the base-2 logarithm of the distance was, however, confirmed only for the larger distances. This can be seen in Fig. 1 for the SPL of the voiced segments. The difference between the versions produced at 0.3 and 1.5 m was relatively small in all respects. This can be thought of as a habitual floor effect: Speakers appear to retain their habitual

vocal effort for this range of distances and increase it appreciably only when clearly required. Speakers would probably decrease the level of their voice at still shorter distances, in order not to cause discomfort to their addressees, but we have no data to confirm this.

As for speech rate, there may also have been an effect of sequential order. The data from the children indicate an especially slow speech rate at the shortest distance, which may have been due to the novelty of the utterance they were required to produce.

In order to account for both types of deviation from the linear relationship between  ${}^2\log(d)$  and the acoustic parameters, a supplementary "closeness" variable was introduced, defined as  $e^{(1-n)}$ , where  $n$  is the version number (proportional to  ${}^2\log(d)$ , see Table I). This was found to compensate in a nearly optimal way for the discrepancy in most acoustic parameters.

### (3) Wind noise

Due to variations in the noise produced by the wind, a higher vocal effort was necessary when the wind velocity was higher and its direction is also relevant. However, since no on-site measurements were taken, the analysis of this effect had to remain somewhat superficial. The values used represent the mean wind velocity in m/s, measured at a nearby meteorological observatory during the hours when the recordings were made. The wind velocity varied between 1 and 7 m/s for the different recording events. It is not expected to be linearly related to any acoustic variables, but attempts to find a more appropriate scaling remained inconclusive. The experiment was not planned to investigate this kind of disturbance.

### (4) Speaker age, ${}^2\log(\text{age in years})$

The analysis was done based on the base-2 logarithm of each speaker's age in years, but care has to be taken in generalizing from this. The data allow us only to say with some confidence how the speech of 7.5-year-olds differs from that of the adults, whose average age was about 30 years, thus two units larger. It is well-known that, for males in particular, the developmental changes between these ages are far from linearly related to the base-2 logarithm of age. For data, see Lee, Potamianos, and Narayanan (1999) and Huber *et al.* (1999).

### (5) Boyhood and (6) manhood

It is not very useful to classify speakers according to sex when adults and children are lumped together. We know *a priori* that the acoustic differences due to sex are much smaller among children than among adults. Unless we make an additional bipartition, the results will vary according to how many speakers of each age group are included. While we can let the age factor account for the difference between girls and adult females, we need the two additional parameters, "boyhood" and "manhood." These have a value of 1 for boys and adult males, respectively, and zero for all other speakers.

TABLE II. Result of a multiple regression analysis of the effects of environmental and speaker specific factors (rows) on acoustic variables (columns) measured in the phonated utterances. Listed are the variance explained ( $r^2$ ) before and after correction for constant speaker-specific factors, a reference value representing a 30-year-old female speaker at a distance of 1.5 m from her addressee when there is no wind, and the effect of the independent variables (1) to (6) mentioned in Sec. II A. The dependent variables were  $SPL_v$ ,  $SPL_0$ , spectral emphasis ( $SPL_v - SPL_0$ ), and  $SPL_s$ , for all of which the effect is expressed in dB and  $F_0, F_1$  of the [a]-segments, and  $F_3$  of the voiced segments, as well as the durations of vowel-like ( $dur_v$ ) and consonantal segments ( $dur_c$ ), for which the effect is expressed as a factor. An asterisk marks values that are significantly different from 0 dB and 1, respectively.

	$SPL_v$	$SPL_0$	Emph	$SPL_s$	$F_0$	$F_{1a}$	$F_3$	$dur_v$	$dur_c$
$r^2$	0.94	0.92	0.79	0.79	0.91	0.84	0.93	0.66	0.28
$r^2$ , speaker specific	0.98	0.96	0.90	0.88	0.97	0.93	0.97	0.88	0.63
Reference value	58.6 dB	53.8 dB	4.9 dB	47.7 dB	175 Hz	580 Hz	2687 Hz	58 ms	70 ms
(1) Distance doubled	+4.6*	+3.3*	+1.4*	+2.0*	1.13*	1.08*	1.00	1.11*	1.02
Distance fivefolded	+10.8*	+7.6*	+3.2*	+4.6*	1.37*	1.19*	1.01	1.27*	1.04
(2) "Closeness" 0.3 vs 1.5 m	+9.5*	+7.0*	+2.6*	+2.6	1.36*	1.09	1.01	1.35*	1.17*
Closeness 0.3 m vs $\infty$	+15.0*	+11.0*	+4.1*	+4.1	1.63*	1.15	1.02	1.61*	1.27*
(3) Wind velocity +1 m/s	+0.6*	+0.5*	+0.1	+0.6*	1.04*	1.03*	1.01*	1.05*	1.00
(4) Speaker age 30 vs 7.5 yrs.	+3.7*	+2.7*	+1.0	+9.0*	0.74*	0.79*	0.75*	0.69*	0.85*
(5) Boy	+4.2*	+4.2*	-0.0	+6.4*	1.00	1.05	1.00	0.99	0.93*
(6) Man	-0.5	-1.1	+0.7	+2.2	0.61*	0.84*	0.88*	1.04	1.00

It would be less adequate to use "girlhood" and "womankind" in addition to "age," since it is essentially the adult male speakers who are different from the others. They have gone through a development in which the vocal folds have been elongated substantially, while the length of the pharyngeal cavity continues to increase.

### (7) Speaker-specific constants

Speakers voices may have particularities that contribute to the total variation observed, and speakers of the same age and sex will not necessarily choose the same vocal effort when addressing somebody at a given distance under the same environmental conditions. In order to gain knowledge about the "sum" of these factors, the speaker-specific average prediction error was calculated for each of the measured acoustic properties. This was done in a linear regression model in which the six above-mentioned variables were used as independent variables. The calculated average error of each speaker was subsequently used as an additional predictor in the analysis of the acoustic data.

## B. The acoustic properties and how they were influenced by the extrinsic factors

Table II shows the results of a number of linear regression analyses in which each of the mentioned extrinsic factors (1) to (6) was considered an independent variable, while the acoustic properties that had been measured were consid-

ered as the dependent variables. Consideration of the speaker-specific constants (7) does not appreciably affect the values shown in this table, except for  $r^2$ , which is a measure of the variance explained, and which is listed with and without consideration of this factor. The first four columns contain the data that refer to levels, the next four concern frequency positions, and the last two concern segment durations. In Fig. 2, the standardized coefficient  $\beta$  is shown for each acoustic and extrinsic (environmental and speaker-specific) variable. This is a measure of the relative contribution of each variable.

### 1. Sound-pressure levels

In Table II it can be seen that  $SPL_v$ , the SPL of the voiced sections, increased by 4.6 dB for a doubling of the distance. In order to keep the SPL constant at the addressee's ear, an increase of 6 dB would have been required. From the values listed in the second row under "distance" and the effect of "closeness" listed in the next row, it can be calculated that the actual increase between 0.3 and 1.5 m was only 1.3 dB. On average,  $SPL_v$  was 32.7 dB higher in version 5 as compared with version 1. See also Fig. 1.

The first partial is less affected by variations in vocal effort. Thus, the SPL of the low-pass-filtered signal ( $< 1.5 F_0$ ),  $SPL_0$ , increased by only 3.3 dB for a doubling of the distance, and the effects of closeness and wind noise were also lower. A separate calculation showed that for com-

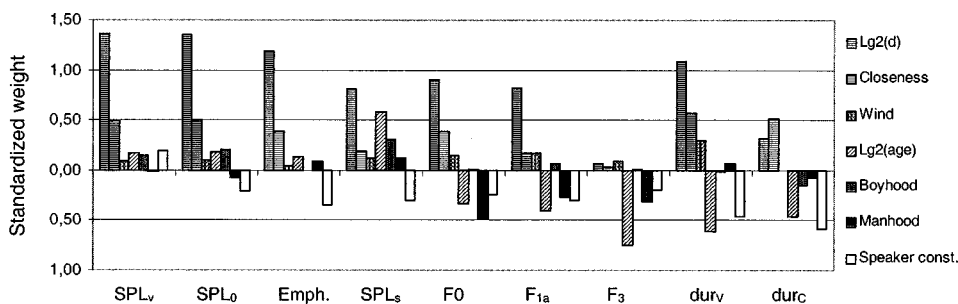


FIG. 2. The contribution of the environmental and speaker specific factors (1) communication distance, (2) "closeness," (3) wind noise, (4) speaker age, (5) boyhood, (6) manhood, and (7) speaker-specific constants, to the variation in acoustic variables measured in the phonated utterances. These variables were (from left to right)  $SPL_v$ ,  $SPL_0$ , spectral emphasis ( $SPL_v - SPL_0$ ),  $SPL_s$ , utterance average  $F_0$ ,  $F_{1a}$ ,  $F_3$ , and the durations of vowel-like ( $dur_v$ ) and consonantal segments ( $dur_c$ ).

TABLE III. Occurrence of creaky voice, in % of the total duration of the voiced segments.

	0.3	1.5	7.5	37.5	187.5	m
Men	1.4	7.8	0.5	0.7	0.0	2.1
Women	7.1	4.4	1.7	0.5	0.0	2.7

ponents at  $f > 1.5 F_{0\text{mean}}$ , there was a nominal increase of 5.1 dB for a doubling of the distance. Speaker age also had a smaller effect on  $\text{SPL}_0$  than on  $\text{SPL}_v$ . For the fourfold increase in age from 7.5 to 30 years, there was an increase of 2.7 and 3.7 dB, respectively.

There was no significant difference in average  $\text{SPL}_v$  between adult male and female speakers. As can be seen in Fig. 1, it was almost the same at any distance. The various effects on SPL that must be assumed to result from the physiological differences between men and women thus compensate each other quite precisely. The values observed for boys were also similar, on average slightly higher than those of the adult speakers. In Table II, this is reflected in the difference between the factors ‘‘boy’’ (+4.2 dB) and ‘‘age’’ (+3.7 dB). Our girls had slightly lower values of  $\text{SPL}_v$  and  $\text{SPL}_0$  at all distances. This is similar to data obtained by Huber *et al.* (1999) from 6-year-olds who sustained an [a] at a comfortable, high, and low level of vocal effort. However, their data obtained from 8-year-olds showed the contrary pattern.

For the three [s]-segments, the increase in level ( $\text{SPL}_s$ ) as a function of distance was smaller than that of the voiced segments ( $\text{SPL}_v$ ).

The five extrinsic factors contributed to a similar extent to  $\text{SPL}_v$  and  $\text{SPL}_0$ . In these cases, the communication distance was the dominant factor. It was somewhat less dominant as a contributor to  $\text{SPL}_s$ , for which age and ‘‘boyhood’’ were more important than for  $\text{SPL}_v$  and  $\text{SPL}_0$ .

While most of the variance in the data was explained by the primary extrinsic factors (see  $r^2$  for  $\text{SPL}_v$ ,  $\text{SPL}_0$ , and  $\text{SPL}_s$  in Table II), additional consideration of the speaker-specific constants led to a significant further increase in  $r^2$ . The value of  $r^2$  for  $\text{SPL}_s$  was reduced by a statistical interaction between speaker group (boys, girls, men, women) and distance or, more properly, vocal effort level. This will be shown in detail further below.

## 2. $F_0$ and formant frequencies

The contribution of the extrinsic factors 1 to 6 to  $F_0$ ,  $F_{1a}$ , and  $F_3$  is shown more to the right in Fig. 2. It can be seen that  $F_0$  as well as  $F_{1a}$  was mainly affected by the communication distance, while speaker age and manhood also contributed appreciably.  $F_3$ , however, remained nearly unaffected by variations in distance, but it varied mainly as a function of speaker age, and also to some extent as a function of manhood. Here, closeness affected only  $F_0$  to a significant extent.

Table III shows the occurrence of creaky voicing in the speech of the adults. Creaky phonation was roughly equally common in the speech of the men and the women, but it was absent in the children’s speech and at very high levels of

TABLE IV. Mean values and standard deviations of  $F_0$  as a function of distance. Standard deviations also expressed in semitones.

		0.3	1.5	7.5	38.5	187.5
Men	Hz	110±17	115±19	138±23	190±28	274±28
	st	2.7	2.8	3.0	2.8	2.0
Women	Hz	213±24	220±28	250±32	311±38	423±235
	st	2.0	2.2	2.2	2.2	1.6
Boys	Hz	276±30	276±28	296±30	391±40	532±48
	st	1.8	1.7	1.8	1.8	1.6
Girls	Hz	287±37	296±39	321±42	406±38	496±39
	st	2.3	2.3	2.3	1.6	1.4

vocal effort. In nearly all cases it occurred in prepausal positions, where the subglottal pressure can be assumed to have been reduced to a level below that required to abduct the vocal folds for each period.

In Table IV, mean values of  $F_0$  are listed together with mean values of a measure of pitch dynamics: the standard deviation in the distribution of  $F_0$  values for each utterance. Sections with creaky phonation were excluded from this analysis. We can see here that in the loudest versions produced by the men,  $F_0$  reached the same value as in the conversational versions produced by the 7-year-olds. Expressed in semitones, the standard deviation of  $F_0$  was roughly constant for each speaker group, but somewhat lower in the loudest versions. This reduction in pitch dynamics is likely to be due to physiological restrictions.

## 3. Segment durations

An analysis of the durations of all segments revealed two different patterns, one that held for all vowels and also for [j] and [v] and another one that held for all consonantal occlusions. The values reported in the following are the mean durations  $\text{dur}_v$  of the vowel-like segments, 14 vowels, 1 [j], and 3 [v], and  $\text{dur}_c$ , of the consonantal occlusions, 8 stop occlusions+voice onset time, 3 [s] and 1 [l], in each utterance.

The contribution of the extrinsic factors 1 to 7 to  $\text{dur}_v$  and  $\text{dur}_c$  is shown to the right in Fig. 2 and Table II. It can be seen that only the duration of the vowel-like segments was significantly affected by the variations in communication distance (and in wind velocity). Closeness affected the duration of all segments to a similar degree. There was also a large effect of age on the durations of all kinds of segments. On average, the vowels produced by the adults had only 69% of the duration of those produced by the children.

## 4. Pausing

The average time all speakers spent pausing between the words in the phonated and whispered versions of the utterance used is shown in Table V. It can be seen there that pausing was avoided within noun phrases, but also between the subject *jag* ‘‘I’’ and the verb *tog* ‘‘took.’’ After the verb, the girls inserted a pause in 27 out of 28 cases. In all other speaker categories, this was much less common. The women paused after the first of the three coordinated object noun phrases, at the place marked with a comma in writing, in all 42 cases, but this place was very popular among all catego-

TABLE V. The mean pausing time, in ms in all phonated and whispered utterances after the word listed in the first column.

Position	Girls	Boys	Women	Men
Jag	0	0	0	0
tog	252	68	51	29
ett	10	12 <sup>a</sup>	13	10
violet	167	192	237	146
åtta	3	0	0	0
svarta	64	160	148	65
och	20	16	0	0
sex	6	17	6	15
vita.				
Σ	522	465	455	265

<sup>a</sup>Here, the occurrence of one exceptional pause with a duration of 2.1 s has been neglected.

ries of speakers. They paused there in 117 out of 138 cases. After the next noun phrase, before the conjunction *och* “and,” boys and women often inserted a pause, while this was more unusual among men and girls.

Pausing behavior was also affected by vocal effort, but here too, we find quite pronounced differences between speaker groups, mainly between children and adults. These are illustrated in Fig. 3. In the utterances produced by adults, there was no uniform change in overall pausing time over most of the range of distances used, but it was markedly longer at the maximal distance. In the children’s utterances, on the contrary, pausing time decreased with increasing distance. We do not know the reason for these differences in behavior.

Most of the increases in pausing time were due to the presence of additional pauses, but when pauses increased in number, those that had been there before tended also to become longer.

### C. The effects of variation in vocal effort on the acoustic properties

#### 1. Sound-pressure levels

Figure 4 shows the SPL of the voiced segments ( $SPL_v$ ) and of the three unvoiced [s]-segments ( $SPL_s$ ) as a function

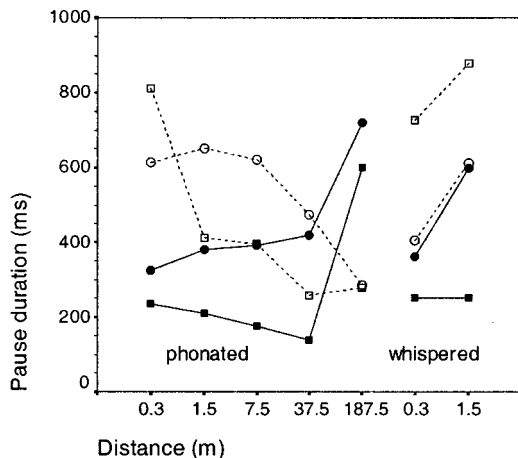


FIG. 3. The mean of the total pause duration (in ms) in phonated and whispered utterances shown as a function of the communication distance for men (■), women (●), boys (□), and girls (○).

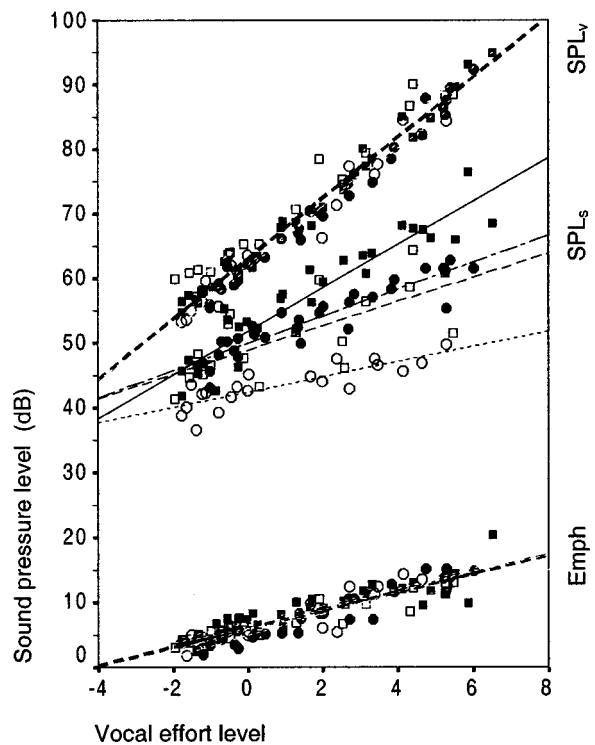


FIG. 4.  $SPL_v$  (above),  $SPL_s$  (middle), and the spectral emphasis  $SPL_v - SPL_0$  (below) shown as a function of vocal effort level  $VEL = 2 \log(d_e)$ , where  $d_e$  is the perceived communication distance in meters. Regression lines fitted to the whole set of data for  $SPL_v$  and emphasis, and to those obtained from each speaker group, men (■, solid lines), women (●, broken), boys (□, dashed), and girls (○, dotted). See also Table VI.

of VEL. As distinct from the nonlinear relation with the logarithm of the communication distance, which can be seen in Fig. 1, it is now evident that  $SPL_v$  is linearly related to VEL, and that there are no apparent differences between speaker groups. Since the differences between speaker groups were so small, only one regression line is shown for the whole set of data ( $r=0.98$ ), but the equations of group-specific regression lines are listed in Table VI. The sound pressure of voiced segments varies as a function of pulmonic pressure, vocal-fold tension, and vocal-tract openness, and the slope of the regression lines reflects the product of these effects.

On average, a doubling of the estimated distance was reflected in an increase in  $SPL_v$  by 4.64 dB. This is quite exactly the same figure as noted in Table II for speakers’ increase in  $SPL_v$  when the distance was doubled. The same also holds for the other regression equations listed in Table VI. Listeners appear to take into account that speakers increase their SPL only by about 4.6 dB instead of the 6 dB that would be required if nothing but SPL were modified. There is, nevertheless, a discrepancy between physical and estimated distances due to the habitual floor phenomenon, for which listeners do not appear to compensate.

In sharp contrast with  $SPL_v$ , there were clear differences between speaker groups in  $SPL_s$ , which varied less than  $SPL_v$  as a function of VEL. This implies that the louder we speak, the more dominant become the voiced parts of the speech signal. This tendency was more pronounced in the speech of women and children than in that of men. For men,



TABLE VI. Equations of the regression lines in Fig. 4 (levels) and 5 (frequencies) and of linear regression lines for  ${}^2\log(\text{dur}_v/\text{dur}_c)$  (cf. Fig. 7) as a function of vocal effort level (VEL).

Men	$\text{SPL}_v = 63.08 + 4.68 \text{ VEL}$	$(r = 0.99)$
Women	$\text{SPL}_v = 61.45 + 4.84 \text{ VEL}$	$(r = 0.99)$
Boys	$\text{SPL}_v = 65.93 + 4.45 \text{ VEL}$	$(r = 0.98)$
Girls	$\text{SPL}_v = 62.11 + 4.49 \text{ VEL}$	$(r = 0.98)$
Men	$\text{SPL}_s = 51.77 + 3.37 \text{ VEL}$	$(r = 0.94)$
Women	$\text{SPL}_s = 49.93 + 2.10 \text{ VEL}$	$(r = 0.92)$
Boys	$\text{SPL}_s = 48.90 + 1.88 \text{ VEL}$	$(r = 0.72)$
Girls	$\text{SPL}_s = 42.41 + 1.18 \text{ VEL}$	$(r = 0.83)$
Men	$\text{SPL}_v - \text{SPL}_0 = 6.85 + 1.31 \text{ VEL}$	$(r = 0.88)$
Women	$\text{SPL}_v - \text{SPL}_0 = 4.87 + 1.66 \text{ VEL}$	$(r = 0.95)$
Boys	$\text{SPL}_v - \text{SPL}_0 = 5.94 + 1.12 \text{ VEL}$	$(r = 0.92)$
Girls	$\text{SPL}_v - \text{SPL}_0 = 5.72 + 1.57 \text{ VEL}$	$(r = 0.92)$
Men	${}^2\log F_0 = 6.92 + 0.22 \text{ VEL}$	$(r = 0.98)$
Women	${}^2\log F_0 = 7.79 + 0.16 \text{ VEL}$	$(r = 0.94)$
Boys	${}^2\log F_0 = 8.25 + 0.15 \text{ VEL}$	$(r = 0.93)$
Girls	${}^2\log F_0 = 8.33 + 0.13 \text{ VEL}$	$(r = 0.95)$
Men	${}^2\log F_{1a} = 9.13 + 0.10 \text{ VEL}$	$(r = 0.91)$
Women	${}^2\log F_{1a} = 9.37 + 0.13 \text{ VEL}$	$(r = 0.95)$
Boys	${}^2\log F_{1a} = 9.76 + 0.16 \text{ VEL}$	$(r = 0.93)$
Girls	${}^2\log F_{1a} = 9.75 + 0.17 \text{ VEL}$	$(r = 0.94)$
Men	${}^2\log F_3 = 11.22 + 0.00 \text{ VEL}$	$(r = 0.12)$
Women	${}^2\log F_3 = 11.47 + 0.02 \text{ VEL}$	$(r = 0.55)$
Boys	${}^2\log F_3 = 11.86 + 0.00 \text{ VEL}$	$(r = 0.02)$
Girls	${}^2\log F_3 = 11.87 + 0.01 \text{ VEL}$	$(r = 0.18)$
Men	${}^2\log(\text{dur}_v/\text{dur}_c) = -0.12 - 0.12 \text{ VEL}$	$(r = 0.82)$
Women	${}^2\log(\text{dur}_v/\text{dur}_c) = -0.07 - 0.15 \text{ VEL}$	$(r = 0.84)$
Boys	${}^2\log(\text{dur}_v/\text{dur}_c) = -0.41 - 0.15 \text{ VEL}$	$(r = 0.90)$
Girls	${}^2\log(\text{dur}_v/\text{dur}_c) = -0.38 - 0.11 \text{ VEL}$	$(r = 0.71)$

$p_s = \text{Effort}^{1.12}$ . While girls showed the lowest variation,  $p_s = \text{Effort}^{0.45}$ , they also showed the lowest absolute values of  $\text{SPL}_s$  throughout.

According to Stevens (1971), the sound pressure of unvoiced fricatives increases as a power function of pulmonic pressure with an exponent of approximately 1.3, but there is one additional factor that is disposed to affect  $\text{SPL}_s$  substantially: dentition and the targeting of the air stream against the teeth (Shadle, 1991). While it may be that adult male speakers have a greater propensity (capacity, inclination, or need) for increasing their pulmonic pressure when increasing their vocal effort, the between-group differences are too large to be explained in this way. The low increase in  $\text{SPL}_s$  in the speech of girls suggests that in this speaker group the targeting of the air stream, which is required for the production of strident fricatives, became less optimal with increasing vocal tract openness. This deterioration was much less pronounced in the speech of men.

Figure 4 also shows the spectral emphasis ( $\text{SPL}_v - \text{SPL}_0$ ) as a function of VEL. In this relation, with  $r = 0.90$ , there were no significant between-group differences. Overall,  $\text{Emph} = 5.9 + 1.42^2 \log(d_e)$ , with Emph in dB and Effort in meters of communication distance.

## 2. $F_0$ and formant frequencies

Figure 5 shows how  $F_0$ ,  $F_1$  of the [a]-segments and  $F_3$

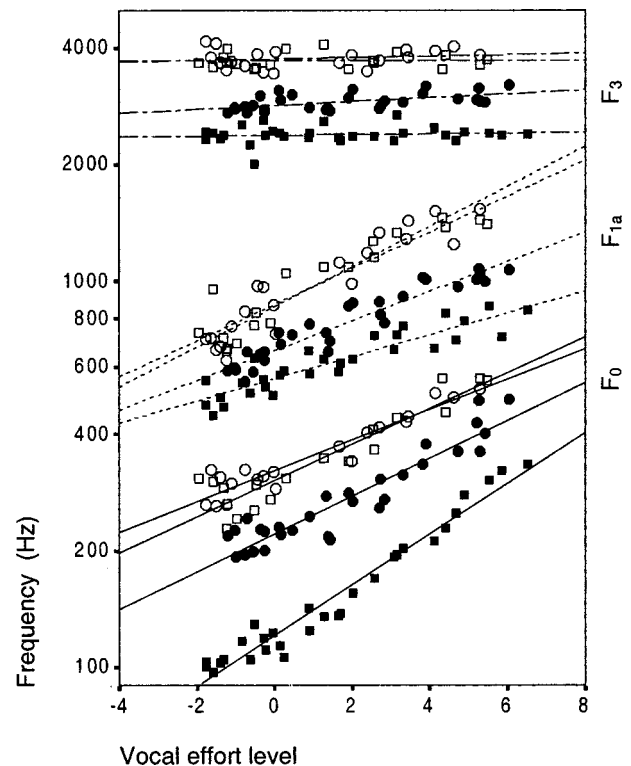


FIG. 5. Mean values of  $F_0$ ,  $F_{1a}$ , and  $F_3$ , shown as a function of VEL for men (■), women (●), boys (□), and girls (○). Regression lines fitted to each variable (solid, dotted, broken lines) and speaker group. See also Table VI.

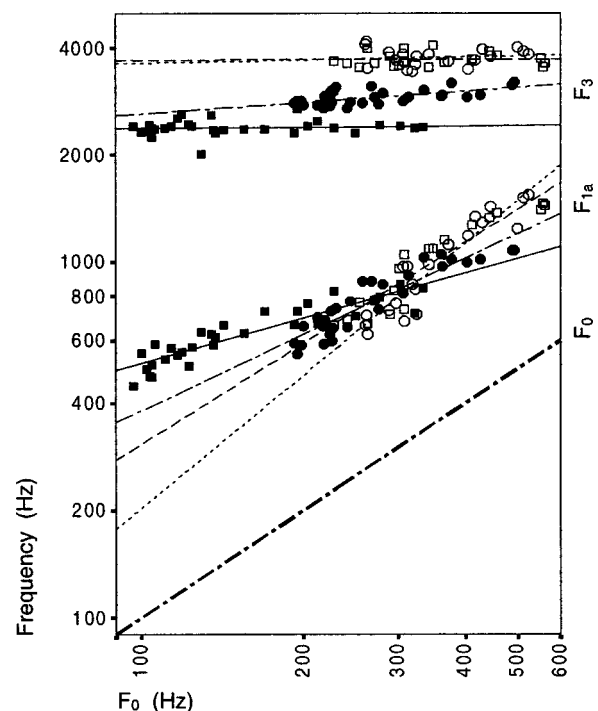


FIG. 6. Mean values of  $F_1$  of the [a]-segments and of  $F_3$  plotted as a function of the mean value of  $F_0$  for each utterance. Regression lines shown for each variable and speaker group, men (■, solid lines), women (●, broken), boys (□, dashed), and girls (○, dotted). Fat broken line:  $F_0$ .

of the voiced segments varied as a function of speaker group and VEL. The increase in  $\log(F_0)$  as a function of VEL was slightly steeper in the men's speech than in that of the other speaker groups, while the opposite was true for the increase in  $F_{1a}$ .

The covariation of  $F_1$  with  $F_0$  can be seen more clearly in Fig. 6, which shows  $F_{1a}$  and  $F_3$  as a function of  $F_0$ , with all variables plotted logarithmically. Here,  $\log(F_{1a})$  can be seen to covary quite closely, but not quite linearly, with  $\log(F_0)$ . The regression lines fitted to  $F_{1a}$  for the four speaker groups differ substantially in slope, but they intersect at an  $F_0$  around 280 Hz. For a 100% increase in  $F_0$ ,  $F_{1a}$  increased by 42% for men ( $r=0.90$ ), 71% for women ( $r=0.92$ ), 95% for boys ( $r=0.94$ ), and 124% for girls ( $r=0.94$ ). The results obtained with the adult speakers are compatible with previous observations, while the effect on  $F_1$  observed in the children's speech is much more drastic. However, these differences in slope need not necessarily be interpreted as genuine differences between the speaker groups, since their major part can also be understood as due to a nonlinear relation between  $F_{1a}$  and  $F_0$ .

The *positive* correlation between  $F_1$  and  $F_0$  in realizations of the same linguistic strings by speakers who differ in age and/or sex, or by the same speakers who alter their pitch register, can be compared with the *negative* correlation that holds between  $F_1$  and  $F_0$  in vowels produced by a given speaker in the same linguistic and paralinguistic context. This has often been referred to as the "intrinsic pitch" of vowels, and it can be understood as, basically, an effect of physiological interactions between tongue and larynx. Observations concerning the magnitude of the intrinsic pitch effect vary considerably, but the almost threefold increase in  $F_1$  from close to open vowels is on average matched by a 10% decrease in  $F_0$  (Whalen and Levitt, 1995). This effect is, thus, much less prominent than the positive correlation between  $F_1$  and  $F_0$  due to paralinguistic factors.

Increases in vocal effort appear to be brought about by a synergetic process that involves an increase in subglottal pressure, an increase in vocal-fold tension, and an increase in the openness of the vocal tract, hence also in  $F_1$ . All these modifications contribute to an increase in the audibility of an utterance, and in ordinary speech, speakers appear to make use of all these possibilities simultaneously. The effects on  $F_0$  and  $F_1$  are substantial and they are similar in magnitude. This contrasts with between-group comparisons of speech produced by men, women, and children at moderate vocal effort, where the factor by which  $F_1$  and the other formant frequencies differ is usually found to be considerably smaller than that for  $F_0$  (Traunmüller, 1988). For a given increase in SPL, the increase in  $F_1$  is larger than that reported for adults by Liénard and Di Benedetto (1999), while it is compatible with the values reported by Rostolland and Parant (1974), but the data are not directly comparable. Rostolland (1982b) reported no increase in  $F_1$  for [a], but large increases in all the other nine vowels included in the study. The increase in  $F_0$  is compatible with the mentioned studies.

A within-speaker increase in  $F_1$  together with  $F_0$  has previously been observed in vowels by Bladon (1982) for men and women, and by Maurer *et al.* (1992), and Klinkert

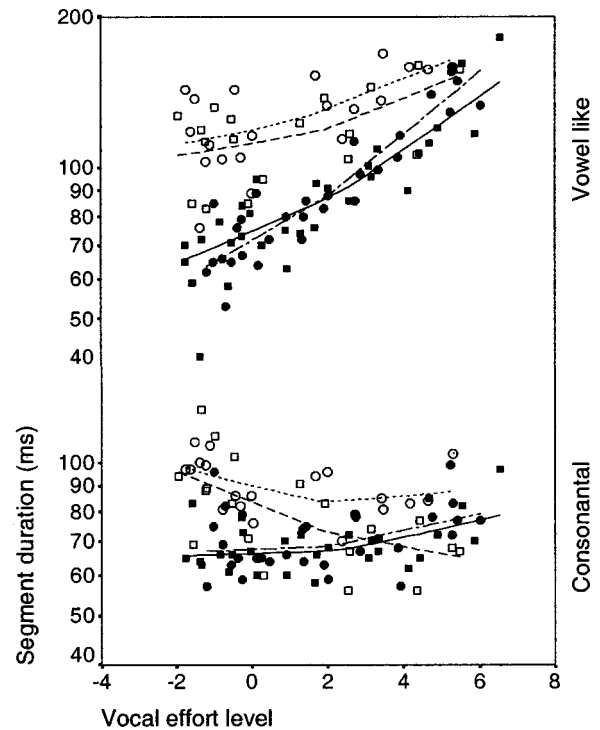


FIG. 7. Mean durations of vowel-like segments (above) and consonantal segments (below) shown as a function of VEL. Locally weighted least-squares regression lines fitted to the data obtained from each speaker group, men (■, solid lines), women (●, broken), boys (□, dashed), and girls (○, dotted). See also Table VI.

and Maurer (1997) for men, women, and children. In these investigations, the expressive variation in  $F_0$  was brought about by asking subjects to modify their pitch. Although this may be more closely related to singing than to varying vocal effort, the covariation of  $F_1$  with  $F_0$  observed under this condition was roughly the same as that observed in the present investigation. We have to conclude that in a given vowel,  $F_1$  tends to be kept at the same frequency position when  $F_0$  is the same, no matter why. This is probably due to the perceptual requirements for constant vowel quality, although a kind of base value of  $F_0$  appears to be more relevant than its instantaneous value (Traunmüller, 1994).

The relation between  $F_3$  and  $F_0$  is of a different kind. In this case, most of the variance was explained by the age factor ( $r^2=0.85$ ).  $F_3$  decreased by 11.5% during the four-fold increase in age from 7.5 to 30 years. Some additional variance was explained by the manhood factor ( $r^2=0.64$  alone and 0.92 combined).  $F_3$  was lower by 11.1% for adult male speakers. There was a small but significant effect of  $F_0$ . For a 100% increase in  $F_0$ ,  $F_3$  increased by 4.1%. In Figs. 5 and 6 it can be seen that most of this average increase in  $F_3$  was due to the group of adult female speakers. However, an analysis of  $F_3$  data obtained from the [a]-segments alone showed again no significant effect in the children, but  $F_3$  increased to the same extent in men and women.

A raised value of  $F_3$  would be expected when speakers shorten their vocal tract by raising their larynx when speaking up, and the data suggest this to happen in adults, but not in 7-year-olds. An explanation of the discrepancies between men and women would require a deeper analysis based on

data from different vowels, but it is to be expected that  $F_3$  in some vowels is more sensitive to variations in vocal effort than in others and that there may be differences in this respect between men and women, since their vocal tracts are not quite proportional, the pharyngeal cavity being relatively longer in men.

### 3. Segment durations

Figure 7 shows the variation in the duration of vowel-like and consonantal segments as a function of VEL. The logarithmic values of the durations of the vowel-like segments can be seen to be positively correlated with VEL, in particular in the adult speakers. In the 7-year-olds, this tendency was much less pronounced and at moderate vocal effort, their vowel-like as well as their consonantal segments had longer duration than in the speech of the adults.

The increase in vowel duration agrees qualitatively with previous observations of the effects of increases in vocal effort (Fónagy and Fónagy, 1966; Rostolland, 1982a; Bonnot and Chevrie-Muller, 1991). However, for consonants these researchers reported a consistent decrease in segment duration with increasing vocal effort. In Fig. 7, an overall decrease can be observed in the children's data (significant for boys), but there is a slight increase in duration in the adults' data (significant for men).

The relationship between the logarithmic values of the segment duration and VEL was nonlinear, but the deviation from linearity was approximately the same for each type of segment and speaker. The slopes of the locally weighted least-squares regression lines are more positive at  $VEL > 2$ , corresponding to distance estimates  $> 4$  m. There appear to be several factors involved: (a) a marked increase in vowel durations, (b) an equal percentage increase in the durations of vowels and consonants at  $VEL > 2$ , and (c) an initially reduced speech rate in children. This has perhaps nothing to do with vocal effort since it may have been due to the novelty of the utterance having caused the children some difficulties.

Here, it is relevant to consider also the ratio  $dur_V/dur_C$ , which is reflected in the difference between  ${}^2\log(dur_V)$  and  ${}^2\log(dur_C)$ . Since the deviation from linearity of the original variables was approximately the same, their difference is linearly related to VEL. Over the whole range in vocal effort, the vowel-to-consonant durational ratio varied by a factor of 2. This raises a question. In Swedish, there is a phonological opposition between sequences realized as [V:C] and [VC:] (Elert, 1964). If this ratio, as well as the absolute segment duration, varies as a function of vocal effort, is there any perceptual constancy in this opposition? This question can only be answered by means of perceptual experiments.

As compared with the adults, the children showed a larger range of variation in segment duration. This is in qualitative agreement with observations such as for the duration of vowels and [s]-segments in the investigation by Lee *et al.* (1999) and with previous observations (Smith, Kenney, and Hussain, 1996).

TABLE VII. Mean values and standard deviations of differences between whispered and voiced versions of the same utterance produced by the same speakers at the same communication distance (0.3 and 1.5 m). The significance level of the difference between the age groups is also indicated.

<i>n</i>	Adults 23	Children 15	Sign.
SPL <sub>v</sub>	-17.7±4.5 dB	-20.8±2.0 dB	**
SPL <sub>s</sub>	-0.7±2.7 dB	-4.6±2.8 dB	***
$F_{1a}$	+24±12%	+26±12%	n.s.
$F_3$	+5.1±4%	+3.3±6.3%	n.s.
dur <sub>V</sub>	+16±17%	+7±23%	n.s.
dur <sub>C</sub>	+11±14%	-14±21%	***

### 4. Recognition of vocal effort

In the results of the present experiment, the SPLs of the voiced segments were highly correlated with VEL,  $r=0.95$  and  $0.98$  for SPL<sub>0</sub> and SPL<sub>v</sub>, respectively. However, this high correlation was due to the exceptional circumstance that the microphone was fixed at a constant distance from the mouth of each speaker. When this is not the case, the emphasis (SPL<sub>v</sub>-SPL<sub>0</sub>) is the best single predictor of VEL ( $r=0.90$ ). It can be assumed to be most important for the perception of vocal effort since it is not significantly affected by speaker age and sex. Since linguistic prominence or stress appears to be realized by means of vocal effort, this measure of spectral emphasis or tilt can also be considered a reliable correlate of this kind of variation, which has been investigated by Sluijter and van Heuven (1996) and Sluijter, van Heuven, and Pacilly (1997). However, instead of comparing the energy in fixed frequency bands, as these investigators did, we prefer to express it by a measure such as our (SPL<sub>v</sub>-SPL<sub>0</sub>), which retains its validity irrespective of  $F_0$ .

If considered together, the formants and  $F_0$  provide similarly reliable information about vocal effort. Various combinations of variables give us correlation coefficients as follows:  $F_0$  and  $F_3$ :  $r=0.87$ ;  $F_0$ ,  $F_3$ , and Emph:  $r=0.96$ ;  $F_0$ ,  $F_3$ , Emph, and  ${}^2\log(dur_V/dur_C)$ :  $r=0.97$ . In the last mentioned case, the standard error of the estimate was 0.64 units, corresponding to factors of 0.64 and 1.56. No further acoustic variables were found to give a significant contribution.

In whispered speech, there is no variation in spectral emphasis since there is no  $F_0$ , and  $F_3$  is hardly affected by variation in vocal effort. However, predicting vocal effort on the basis of  $F_3$ ,  $F_{1a}$ , and  ${}^2\log(dur_V/dur_C)$  results in  $r=0.902$ . This means that it should be possible to recognize variations in vocal effort fairly well also in whispered speech. However, this involves a risk of confusing variations in linguistic and paralinguistic quality. As for the perception of speaker sex, this has been observed experimentally (Traunmüller, 1998).

### D. Whispering compared with phonated speech

Table VII presents the differences between the whispered and the voiced versions of the same utterance produced by the same speakers at the same conversational distances of 0.3 and 1.5 m for adults (23 pairs) and children (15 pairs). Some of these quantities showed a significant differ-

TABLE VIII. Mean perceived and calculated distances between speaker and addressee for the phonated versions compared with distances calculated using the same equations for the whispered versions. The independent variables were  $F_{1a}$ ,  $F_3$ ,  $\text{dur}_V$ , and  $\text{dur}_C$ .

Perc. dist. (m), phonated	0.47	0.69	1.9	7.5	31
Calc. dist. (m), phonated	0.52	0.82	2.0	7.7	22
Calc. dist. (m.) whispered	2.0	3.3			

ence between these age groups, but there was no significant difference between sex groups, and  $\text{dur}_C$  was the only variable that showed a significant difference related to distance (0.3 vs 1.5 m).

In the adults' speech, the durations of all segments tended to be longer in the whispered versions than in the corresponding phonated ones. In the children's whispered versions, the vowels were on average also slightly longer, but the consonants were shorter in duration. This is similar to what happened with the phonated versions when the vocal effort was increased (see Fig. 7). This holds as well for the observed formant frequencies, in particular for  $F_{1a}$ . This suggests that in whispering, speakers adopt an articulation similar to that used for phonated speech produced at a higher vocal effort. This may be caused by a need to increase the tension (stiffness) of the vocal folds in order to prevent vocal-fold vibration. When controlling vocal effort, such an increase in vocal-fold tension is accompanied by modifications in articulation, which might automatically be realized also in whispering. It may also be that speakers attempt to increase the audibility of their whispers by modifying their articulation in the same way as they do for phonated speech.

More precise formant frequency measurements by Eklund and Traunmüller (1996) in vowels produced by adults showed an average difference of 31% in  $F_1$ , 12% in  $F_2$ , and 3% in  $F_3$ , and similar values had been obtained in previous studies (Peterson, 1961; Kallail and Emanuel, 1984a, b).

A linear regression model using no more than the acoustic variables  $\text{dur}_v$ ,  $\text{dur}_c$ ,  $F_{1a}$ , and  $F_3$  can be used to predict the perceived distances. The results of such a calculation, based on the logarithmic values of these four variables, is

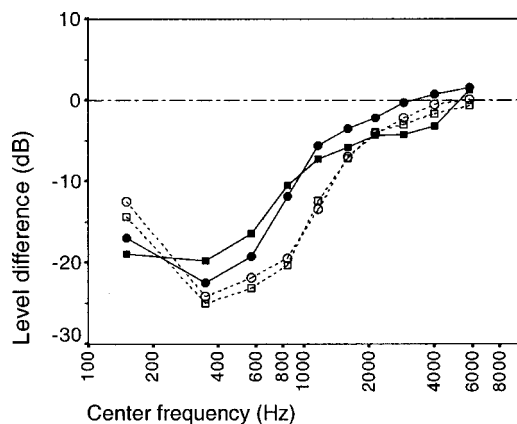


FIG. 8. The gross difference in spectral energy distribution between whispered and phonated versions of the same utterance produced by men (■), women (●), boys (□), and girls (○) at the same communication distance (0.3 and 1.5 m), based on level measurements in frequency bands covering three critical bands with overlap.

shown in Table VIII. For the phonated versions, the calculated values agree reasonably well with the perceived distances ( $r=0.97$ , standard error 0.59 units). The distances calculated with the same equations for the whispered utterances were higher by a factor close to 4.

However, the slightly, and for children significantly, lower SPL of the [s]-segments that was observed in the whispered utterances (see Table VII) suggests that speakers do not increase their subglottal pressure for the purpose of whispering, nor do they necessarily lower it, since the lower  $\text{SPL}_s$  may also result from a less efficient targeting of the air stream (Shadle, 1991), which already has been suggested as an explanation of the between-group differences observed in the relation between VEL and  $\text{SPL}_s$ .

Figure 8 describes the overall difference in spectral energy distribution, as measured in filter bands with a width corresponding to 3 barks, overlapping by 1 bark on each side. The frequency bands were 20–300, 200–510, 400–770, 630–1080, 920–1480, 1270–2000, 1720–2700, 2320–3700, 3150–5300, and 4400–7700 Hz, all with slopes of  $\pm 36$  dB/octave. The figure shows the mean level difference (whispered–phonated) obtained in each band, represented at its midfrequency, for each speaker group. In the region of  $F_1$ , the levels can be seen to be much lower in whispering than in phonated speech. Most of the difference in the energy distribution between the two modes can be attributed to the difference between the spectra of the source signals and to the damping that affects the aperiodic source signal by the coupling to the subglottal cavities when the glottis is open. The results obtained in the lowest frequency band are not reliable, since this band of the whispered versions was in many instances dominated by wind noise while the same band in children's phonated versions barely contained the lowest partial.

Figure 8 shows virtually no difference between girls and boys, and the difference between children and women consists of a shift of the curve in level and frequency. The apparent frequency shift between speaker groups can be understood on the basis of the assumption that the spectrum of the noise source effective in whispering is shifted in frequency to the same extent as the formant frequencies, cf. the values noted in Table II for the effects of speaker age and sex on  $F_3$  and  $F_{1a}$ . The shape of the curve describing the level difference in the speech of men is somewhat different from that observed in women and children. This difference is probably due to differences in phonation rather than in noise production. Similar differences can be seen in the data obtained by Eklund and Traunmüller (1997), who used the same method when comparing the spectra of whispered and phonated vowels produced by men and women.

### III. GENERAL DISCUSSION

Not surprisingly, SPL was strongly affected by vocal effort, but it gives a listener no unambiguous clues. While variation in SPL may be due to expressive factors, such as vocal effort, it may also be due to perspectival factors, such as the listener's distance from the speaker. More reliable information is conveyed by spectral emphasis ( $\text{SPL}_v - \text{SPL}_0$ ), albeit that linguistic factors (the phonetic quality of

segments) also affect emphasis as well as SPL. Several other characteristics of speech were also affected substantially by vocal effort, most notably  $F_0$  and  $F_1$ , but also the duration of vowels and consonants, pausing behavior, and occurrence of creaky voice. While in several respects, men, women, boys, and girls behaved somewhat differently ( $SPL_s$ ,  $F_{1a}$  compared with  $F_0$ , segment durations, and pausing behavior), there were almost no such differences in  $SPL_v$  and  $SPL_0$  (or Emph). Although the occurrence of creaky phonation has been considered typical for male speech (Henton and Bladon, 1987), no such gender difference was observed in the present investigation.

When the distance was increased from 0.3 to 1.5 m, there was generally only a small change in acoustic parameter values. Speakers appear not to change their vocal effort substantially within the range of distances that are most common in everyday conversations. This habitual floor was modeled by means of an *ad hoc* factor (closeness). However, speaker behavior in response to variation in this range of distances requires a more detailed study.

The results were affected by some additional interfering variables such as wind and between-speaker differences. All this interference was eliminated by taking the communication distance estimated by a group of listeners for each utterance as a measure of vocal effort. This measure correlated better with the acoustic variables than the physical distances did. Speakers increased their  $SPL_v$  by an average of only 4.6 dB instead of the 6 dB that in the absence of any variation in other factors would be required in order to compensate for a doubling in distance. Although listeners appear to take this into account with surprising precision, basing their distance estimates on a 4.6-dB increase per doubling, there remains a substantial discrepancy between physical and estimated distances due to the habitual floor phenomenon, for which acoustic signals contain no clues.

Measurements of the frequency position of the formants, considered as the resonances of the vocal tract, are affected by substantial errors when  $F_0$  is as high as it is when people communicate over large distances. This holds for LPC-based methods as well as when using visual inspection of spectrograms. In the measurements of the mean values of  $F_3$  (over all potentially voiced segments) and of  $F_1$  in the [a] segments, which we report, we judged these errors to remain below a reasonable tolerance level. Data on all the different vowels would have been valuable, but this will have to await the development of more accurate measurement techniques.

The whispered versions were compared with the phonated versions produced by the same speakers at the same distance, and it was observed that segment durations and formant frequencies were affected in a similar way by switching over to whispering as when vocal effort was increased. It was suggested that this may be due to a similarity in articulatory settings, but this hypothesis remains to be tested.

The reported data are mainly descriptive of the organic and expressive properties of speakers' voices. This is what is referred to as the "carrier signal" within the framework of the modulation theory (Traunmüller, 1994, 1998). This signal conveys most of the nonlinguistic information, while the

linguistic information is conveyed by its "modulation," i.e., by the way in which the properties of the speech signal deviate from those of the carrier, whose properties are more permanent but not always explicit. With the possible exception of  $F_3$ , we have found variation in vocal effort to cause significant variation in all the acoustic variables investigated. This poses no problem for the modulation theory, but it is hard to reconcile with any theory that assumes absolute properties to be descriptive of the linguistic phonetic quality of speech sounds and with normalization-based models that account only for differences *between* speakers. The modulation theory predicts that speech sounds with the same modulation will have the same linguistic phonetic quality, "sameness" here being a subjective concept. However, in the absence of phone-specific data we are unable to say anything more about this.

The results of the present investigation can be readily used to simulate variations in vocal effort and whispering in speech synthesis. They also allow us to test how well the expression of linguistic stress can be simulated by increasing *vocal effort* on a syllable, with all its acoustic correlates in the right proportions, or by decreasing vocal effort in the surroundings of the syllable.

The results can also be used for automatic recognition of the paralinguistic quality of speech, and they contribute to the solution of the problem of recognizing the linguistic message when the paralinguistic quality is not constant. This would be a fundamental progress, since present-day commercial systems of speech recognition lack the capability of distinguishing the paralinguistic from the linguistic information in speech signals.

## ACKNOWLEDGMENTS

We want to express our thanks to Anita Andersson, Hassan Djamshidpey, Ingegerd Eklund, and Jessika Rundlöf, who have been engaged in various facets of these "cries and whispers." This research has been supported, in part, by a grant from HSEFR, the Council for Research in the Humanities and Social Sciences, and NUTEK, the National Board for Technical and Industrial Development, within the frame of the Swedish language technology program.

- Åkerlund, L., and Gramming, P. (1994). "Average loudness level, mean fundamental frequency, and subglottal pressure: Comparison between female singers and nonsingers," *J. Voice* **8**, 263–270.
- Bladon, A. (1982). "Arguments against formants in the auditory representation of speech," in *The Representation of Speech in the Peripheral Auditory System*, edited by R. Carlson and B. Granström (Elsevier Biomedical, Amsterdam), pp. 95–102.
- Bond, Z. S., Moore, T. J., and Gable, B. (1989). "Acoustic-phonetic characteristics of speech produced in noise and while wearing an oxygen mask," *J. Acoust. Soc. Am.* **85**, 907–912.
- Bonnot, J-F. P., and Chevrie-Muller, C. (1991). "Some effects of shouted and whispered conditions on temporal organization," *J. Phonetics* **19**, 473–483.
- Childers, D. G., and Lee, C. K. (1991). "Vocal quality factors: Analysis, synthesis, and perception," *J. Acoust. Soc. Am.* **90**, 2394–2410.
- Eklund, I., and Traunmüller, H. (1997). "Comparative study of male and female whispered and phonated versions of the long vowels of Swedish," *Phonetica* **54**, 1–21.
- Elert, C-C. (1964). *Phonological Studies of Quantity in Swedish* (Almqvist & Wiksell, Uppsala).

- Eriksson, A., and Traunmüller, H. (1999). "Perception of vocal effort and speaker distance on the basis of vowel utterances," Proceedings of the ICPHS-99, Vol. 3, pp. 2469–2472. (Accessible on the World Wide Web.)
- Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).
- Fónagy, I., and Fónagy, J. (1966). "Sound pressure level and duration," *Phonetica* **15**, 14–21.
- Frøkjær-Jensen, B. (1966). "Changes in formant frequencies and formant levels at high voice effort," *ARIPUC 1* (Inst. Phonet., University Copenhagen), pp. 47–55.
- Gauffin, J., and Sundberg, J. (1989). "Spectral correlates of glottal voice source waveform characteristics," *J. Speech Hear. Res.* **32**, 556–565.
- Granström, B., and Nord, L. (1992). "Neglected dimensions in speech synthesis," *Speech Commun.* **11**, 459–462.
- Henton, C. G., and Bladon, R. A. W. (1987). "Creak as a socio-phonetic marker," in *Language, Speech and Mind: Studies in Honor of Victoria Fromkin*, edited by L. Hyman and C. N. Li (Routledge, London), pp. 3–39.
- Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (1988). "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," *J. Acoust. Soc. Am.* **84**, 511–529.
- Imberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., and Goldman, S. L. (1995). "Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice," *J. Speech Hear. Res.* **38**, 1212–1223.
- Huber, J. E., Statopoulos, E. T., Curione, G. M., Ash, T. A., and Johnson, H. K. (1999). "Formants of children, women, and men: The effect of vocal intensity variation," *J. Acoust. Soc. Am.* **106**, 1532–1542.
- Junqua, J.-C. (1993). "The Lombard reflex and its role on human listeners and automatic speech recognizers," *J. Acoust. Soc. Am.* **93**, 510–524.
- Kallail, K. J., and Emanuel, F. W. (1984a). "Formant-frequency differences between isolated whispered and phonated vowel samples produced by adult female subjects," *J. Speech Hear. Res.* **27**, 245–251.
- Kallail, K. J., and Emanuel, F. W. (1984b). "An acoustic comparison of isolated whispered and phonated vowel samples produced by adult male subjects," *J. Phonetics* **12**, 175–186.
- Klinkert, A., and Maurer, D. (1997). "Fourier spectra and formant patterns of German vowels produced at  $F_0$  of 70–850 Hz," *J. Acoust. Soc. Am.* **101**, 3112(A).
- Lee, S., Potamianos, A., and Narayanan, S. (1999). "Acoustics of children's speech: Developmental changes of temporal and spectral parameters," *J. Acoust. Soc. Am.* **105**, 1455–1468.
- Liénard, J.-S., and Di Benedetto, M.-G. (1999). "Effect of vocal effort on spectral properties of vowels," *J. Acoust. Soc. Am.* **106**, 411–422.
- Lindblom, B., and Sundberg, J. (1971). "Acoustical consequences of lip, tongue, jaw, and larynx movement," *J. Acoust. Soc. Am.* **50**, 1166–1179.
- Loren, C. A., Colcord, R. D., and Rastatter, M. P. (1986). "Effects of auditory masking by white noise on variability of fundamental frequency during highly similar productions of spontaneous speech," *Percept. Mot. Skills* **63**, 1203–1206.
- Maurer, D., Cook, N., Landis, T., and d'Heureuse, C. (1992). "Are measured differences between the formants of men, women and children due to  $F_0$  differences?," *J. Int. Phon. Ass.* **21**, 66–79.
- Peterson, G. (1961). "Parameters of vowel quality," *J. Speech Hear. Res.* **4**, 10–29.
- Rastatter, M., and Rivers, C. (1983). "The effects of short-term auditory masking on fundamental frequency variability," *J. Aud. Res.* **23**, 33–42.
- Rivers, C., and Rastatter, M. P. (1985). "The effects of multitalker and masker noise on fundamental frequency variability during spontaneous speech for children and adults," *J. Aud. Res.* **25**, 37–45.
- Rostolland, D. (1982a). "Acoustic features of shouted voice," *Acustica* **50**, 118–125.
- Rostolland, D. (1982b). "Phonetic structure of shouted voice," *Acustica* **51**, 80–89.
- Rostolland, D., and Parant, C. (1974). "Physical analysis of shouted voice," *Contr. to 8th International Congress on Acoustics*, London.
- Rundlöf, J. (1996). *Perceptuella ledtrådar vid auditiv bedömning av avståndet mellan talare och lyssnare* (Perceptual clues in auditory judgments of the distance between speaker and listener) (D-uppsats, Inst. för lingvistik, Stockholms universitet).
- Schulman, R. (1985). "Articulatory targeting and perceptual constancy of loud speech," *PERILUS* No. 4 (Inst. linguist., Stockholm University), pp. 86–91.
- Schulman, R. (1989). "Articulatory dynamics of loud and normal speech," *J. Acoust. Soc. Am.* **85**, 295–312.
- Shadle, C. H. (1991). "The effect of geometry on source mechanisms of fricative consonants," *J. Phonetics* **19**, 409–424.
- Sluijter, A. M. C., and van Heuven, V. J. (1996). "Spectral balance as an acoustic correlate of linguistic stress," *J. Acoust. Soc. Am.* **100**, 2471–2485.
- Sluijter, A. M. C., van Heuven, V. J., and Pacilly, J. J. A. (1997). "Spectral balance as a cue in the perception of linguistic stress," *J. Acoust. Soc. Am.* **101**, 503–513.
- Smith, B. L., Kenney, M. K., and Hussain, S. (1996). "A longitudinal investigation of duration and temporal variability in children's speech production," *J. Acoust. Soc. Am.* **99**, 2344–2349.
- Södersten, M., Hertegård, S., and Hammarberg, B. (1995). "Glottal closure, transglottal air-flow, and voice quality in healthy middle-aged women," *J. Voice* **9**, 182–197.
- Stevens, K. N. (1971). "Airflow and turbulence noise for fricative and stop consonants: Static considerations," *J. Acoust. Soc. Am.* **50**, 1180–1192.
- Stevens, S. S. (1965). "Matching functions between loudness and ten other continua," *Percept. Psychophys.* **1**, 5–8.
- Summers, V. W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., and Stokes, M. A. (1988). "Effects of noise on speech production: Acoustical and perceptual analyses," *J. Acoust. Soc. Am.* **84**, 917–928.
- Titze, I. R., and Sundberg, J. (1992). "Vocal intensity in speakers and singers," *J. Acoust. Soc. Am.* **91**, 2936–2946.
- Traunmüller, H. (1988). "Paralinguistic variation and invariance in the characteristic frequencies of vowels," *Phonetica* **45**, 1–29.
- Traunmüller, H. (1994). "Conventional, biological and environmental factors in speech communication: A modulation theory," *Phonetica* **51**, 170–183.
- Traunmüller, H. (1997). "Perception of speaker sex, age, and vocal effort," *Phonum 4* (Dept. of Phonetics, Umeå University), pp. 183–186.
- Traunmüller, H. (1998). "Modulation and demodulation in production, perception, and imitation of speech and bodily gestures" in *Proceedings of the Eleventh Swedish Phonetics Conference*, edited by P. Branderud and H. Traunmüller (Department of Linguistics, Stockholm University, Stockholm), pp. 40–43. (Accessible on the World Wide Web.)
- Whalen, D. H., and Levitt, A. G. (1995). "The universality of intrinsic  $F_0$  of vowels," *J. Phonetics* **23**, 349–366.
- Wilkens, H., and Bartel, H.-H. (1977). "Wiedererkennbarkeit der Original-lautstärke eines Sprechers bei elektroakustischer Wiedergabe," *Acustica* **37**, 45–49.

# On operating deflection shapes of the violin body including in-plane motions

Anna Runnemalm<sup>a)</sup> and Nils-Erik Molin

*Division of Experimental Mechanics, Luleå University of Technology, SE-971 87 Luleå, Sweden*

Erik Jansson

*Department of Speech, Music and Hearing, Royal Institute of Technology (KTH), SE-100 44 Stockholm, Sweden*

(Received 12 July 1999; revised 10 November 1999; accepted 24 February 2000)

Earlier investigations have assumed only “out-of-plane” vibrations of the plates of the violin. The violin body can, however, be described as a thin-walled, double-arched shell structure and as such it may very well elongate in one direction as it contracts in another. Therefore, at least two orthogonal vibration components have to be included to describe the vibrations. The operating deflection shapes (ODSs) of a good, professionally made and carefully selected violin were therefore measured in several directions by TV holography to determine both “in-plane” and out-of-plane vibration components of the ODSs. The observations were limited to the frequency range 400–600 Hz, as this interval includes two most-prominent resonance peaks of bridge mobility and sound radiation as well as a third poorly radiating resonance. These three peaks clearly showed orthogonal vibration components in the ODSs. The vibration behavior of the violin body, sectioned in the bridge plane, was interpreted as the vibrations of an “elliptical tube” with nodal diameters. The number of nodal diameters increases from two to three in the selected frequency range. The TV holography measurements were supported by electro-dynamical point measurements of bridge mobility, of air volume resonances, and by reciprocity, of radiation properties. Furthermore, a fourth mode, the air mode, A1, is involved indirectly in the sound radiation via influence on the body vibrations. © 2000 Acoustical Society of America. [S0001-4966(00)01106-1]

PACS numbers: 43.75.De [WJS]

## INTRODUCTION

The violin has interested not only makers, players, and collectors, but also a fair number of scientists. For the scientist it represents a challenge. Even today its fundamental physical function is not well understood—“of all the musical instruments, the violin may be said to be the most studied and the least understood.”<sup>1</sup> Most of the knowledge collected up to 1980 is summarized in the monumental book, *The Physics of the Violin*.<sup>2</sup>

Saunders started a long series of physical experiments on the violin. He investigated its function and tried to relate it to the quality. In this work he mapped properties of the A0 resonance (often called the Helmholtz resonance) and the principal peak P of the violin, related to the wolf-tone phenomenon of bowed strings.<sup>3</sup> A group of researchers around Saunders interested in violin acoustics founded the Catgut Acoustical Society. The results on A0 and P of Saunders were later tested in a new set of bowed stringed instruments, the Violin Octet constructed and built by Hutchins, the violin maker of the group.<sup>4</sup>

In the late 1960's new optical measurement techniques using the newly invented laser were developed. The new techniques for vibration measurements were based on interference and holography, called hologram interferometry. The technique could simultaneously measure the vibrations of the whole visible area of a vibrating surface without limiting

demands of surface flatness. The technique was early applied to investigate vibration modes of violins.<sup>5</sup> Thereby, the labeling T for top-plate modes and B for back-plate modes was selected. Unfortunately, the rigid holding needed for the measurements at that time limited the possibilities for analysis of the assembled violin.

Somewhat later modes of the air volume of the violin were reported on.<sup>6</sup> It was now shown that eigenmodes could be found in the volume above the lowest one, A0. The second air mode A1, at about 460 Hz, is well-developed in the violin family of instruments. Possibly mode A1 shows in the back-plate vibrations of the violin investigated previously by hologram interferometry [Fig. 16(c), Ref. 5]. The resonant frequency of A1 is close to that of the P peaks mapped by Saunders and this fact was later suggested as a quality measure of violins.<sup>7</sup>

The optical techniques were further developed and a method called speckle interferometry could be applied to vibration analysis of the complete violin without the rigid holding necessary for the earlier hologram-interferometry measurements. Thereby, modes of the complete violins could be mapped.<sup>8</sup> The measurements indicated eigenmodes with the top and back plates vibrating in phase, i.e., as a homogeneous thick plate up to 700 Hz. The labeling C for corpus modes was selected. The investigations also indicated that the P peak is not one but two peaks, labeled T1 and C3 in Ref. 8. By now it is well-known that we have at least three resonances in the violin body between 400 and 600 Hz.

The response of the violin to external excitation in the

<sup>a)</sup>Electronic mail: Anna.Runnemalm@mt.luth.se

range of 200 to 600 Hz was investigated with great care.<sup>1</sup> The excitation was incident sinusoidal sound waves and the responses were recorded by transducers at the strings, the bridge feet, and in the air volume of the four violins investigated. The investigated violins exhibited the four well-defined normal modes, conveniently described as linear combinations of four elementary motions: First, a Helmholtz motion of the air in the cavity, then a breathing motion of the body, followed by a sloshing motion of the air along the cavity and at last a bending motion of the violin body, cf. A0, T1, A1, and C3, respectively. Note that the violin was found to bend crosswise in C3 by Alonso Moral and Jansson,<sup>8</sup> but lengthwise by Arnold and Weinreich.<sup>1</sup>

With the development of the computer, successive point measurements could be evaluated simultaneously by the modal analysis technique. This technique made it possible to measure accurately magnitudes and phases of vibration at a large number of points and to evaluate the measurements in the form of normal modes by means of digital processing. In such an analysis, two “similar but reciprocal” modes were found, i.e., the first bending of corpus at 466 and at 574 Hz, top-back transverse-longitudinal and longitudinal-transverse, respectively.<sup>9</sup> The two modes correspond in frequency to the modes labeled T1 and C3 above. Large computers and finite-element models have been used to calculate eigenmodes of the violin, both the plates and the cavity of the assembled violin.<sup>10</sup> The two wall resonances (T1 and C3) were found at 432 and 519 Hz, and the A1 resonance at 482 Hz.

Radiativity has been measured as sound pressure in the far field for a constant amplitude driving force.<sup>11</sup> The constant force was obtained by a special, ideal driver, which excited a bridge corner in the main bowing direction with very small influence on the tested violins. The response of old Italian violins show two prominent peaks between 400 and 600 Hz. The bridge admittance (bridge mobility) of violins of soloist quality have also been measured with very small influence on tested violins.<sup>12</sup> The violins had two very prominent peaks between 400 and 600 Hz. The peak at the higher frequency was a few dB higher than the one at the lower frequency. After electronic manipulation of the sound response and the mobility response to the same frequency and magnitude scales, the two responses showed very similar curves, i.e., bridge mobility indicates properties to be found in radiated sound.<sup>13</sup>

The violin body can be described as a thin-walled, double-arched shell structure. It is prestressed by the tension of the strings so the vibration amplitudes of the shell surface are superimposed to the prestressed state. This eliminates play between the different parts of the violin body. A shell may very well elongate in one direction as it contracts in another. Both in-plane and out-of-plane components of the vibration amplitudes are of importance for a full description of the vibrations in a violin.

Out-of-plane motions of three violins have been recorded with an electronic holography system (or TV holography system).<sup>14</sup> The measurements showed clear out-of-plane vibrations of the violin ribs between 400 and 600 Hz, thus indicating in-plane vibrations at the plate edges in addi-

tion to the out-of-plane vibrations (of the modes T1 and C3, which corresponds to peaks P1 and P2). In following investigations it was shown that the influence of the A1 resonance can be seen in a violin top and back plate<sup>15</sup> and that the radiation of both the T1 and the C3 modes corresponds to those of prominent omnidirectional sources as peaks P1 and P2 are prominent in radiativity response.<sup>16</sup> In the earlier work the influence of the soundpost was investigated for excitation at the bridge, also plotted as deformations of the cross section in the bridge plane.

In a Spanish doctoral work, the modes of the violin body were calculated employing finite-element methods. Vibrations of the top plate, the back plate, and the ribs were calculated using both in-plane and out-of-plane vibrations,<sup>17</sup> published partly in English.<sup>18,19</sup> The air cavity modes were also calculated assuming rigid walls. The resonant frequencies were at 470, 506, and 570 Hz for T1, A1 and C3, respectively.

Thus, there is clear evidence that we should expect three resonances in bridge mobility and sound radiation, i.e., the A1, T1, and C3 resonances in the 400 to 600 Hz range. The top and back plates of the violin are, however, not plates. They are arched just on the limit of a “slightly curved plate,” or in technical terms a shell just on the limit of a “shallow shell.” For a shell, the out-of-plane vibrations are coupled to in-plane vibrations, which lead to the first question of this investigation. What are the ODSs at vibration maxima (ODS—operating deflection shape is a more correct notation, as we cannot really observe eigenmodes) if we include both in-plane and out-of-plane vibrations of the violin plates for sinusoidal excitation at the violin bridge?

With care taken of both in-plane and out-of-plane vibrations, we shall obtain a more complete map of ODSs at resonant frequencies. This leads us to the second question of this investigation. With the improved understanding of the vibrations, is it possible to determine the role of the A1 resonance in supporting radiation properties of the violin?

This present investigation focuses on the presence or nonpresence of in-plane vibrations for the ODSs in the frequency range 400–600 Hz. The study of a complete varnished violin is also a challenge for the TV holography system itself.

We introduce labelings P1 and P2 for the most prominent peaks to avoid previous misunderstandings with different and confusing labeling. There are good reasons to introduce P1 and P2 as new peak labelings. Our T1 and C3 indicate eigenmodes, but the peaks are results of a weighted sum of several eigenmodes forming the ODSs measured. There are other labelings used which add to the confusion when reading texts on ODSs/modes in the 400–600 Hz range of violins. It should be pointed out that the P1 and P2 peaks are well established in frequency responses of both bridge mobility and sound radiation.

## I. EXPERIMENTS

A violin of good professional quality, Bernardel 1909, was selected for the experiments. The violin had previously been evaluated both in playing tests and by acoustical measurements and had been found to have properties typical of a



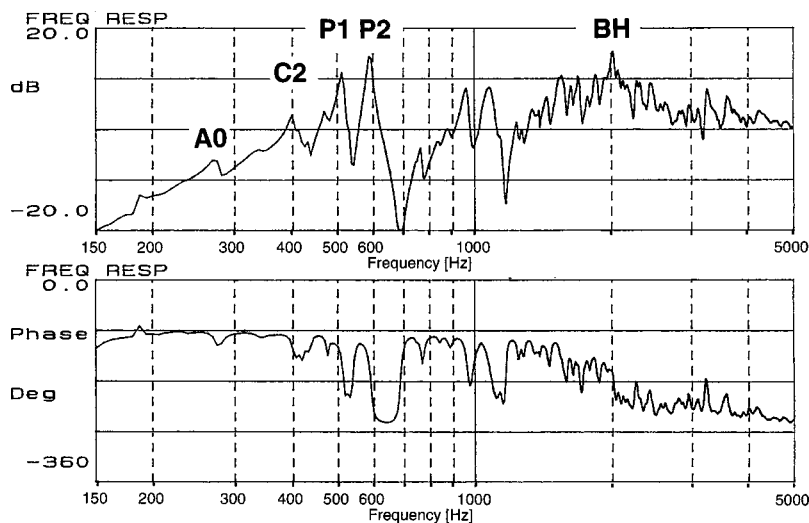


FIG. 1. Frequency response (mobility=velocity/force 150–5000 Hz) at the bridge of the violin (level and phase) measured in bowing direction, and with P1 at 512, P2 at 587, and BH 2010 indicated. Minor peaks at A0 275 Hz, C2 at 400 Hz, and a peak at 469 Hz followed by a dip at the A1 resonance (473 Hz) just below the P1 peak.

good violin. The violin was not prepared in any way and was complete with strings and a shiny and fairly dark varnish.

### A. Acoustical measurements

The bridge mobility of the violin was measured by means of a standard procedure developed at the KTH laboratory.<sup>12</sup> The violin is laid horizontally on top of two felt-covered supports under its back plate. One support is at the very lower end and one at the very upper end. Sympathetic string vibrations are damped out by rubber wedges and felt between the strings. The bass side of the bridge is excited by an impulse force hammer in a pendulum arrangement. The resulting bridge vibrations are recorded by a small magnet (40 mg) waxed to the treble side of the bridge over an air gap by means of an electrical coil. The force signal of the hammer and the velocity signal of the coil are connected to a two-channel fast Fourier transform (FFT) analyser, HP3562A. The support and measuring transducers have very little influence on the bridge mobility of the measured violin. The excitation force and resulting vibration velocity are measured in the bowing direction.

For the optical measurements a more rigid holding of the violin is needed. A series of different holdings was tried and compared to the standard one (very close to free). Thereby, it was found that a clamping of the violin neck perturbs the dynamical properties (measured as bridge mobility) of the violin to a very high degree. After some experimenting it was found that squeezing the violin between four rubber blocks at the ribs, two on each side of the tail-piece pin and two on each side of the foot of the neck, causes minor effects on the bridge mobility. It must be stressed that a careful control of a free neck is important.

The mobility curve with a dB magnitude scale as in Fig. 1 is our common way to present acoustical measurements. However, when measuring ODSs by TV holography the displacement vibration amplitude is measured. Therefore, a curve was plotted corresponding to the sensitivity of the optical measurements, displacement over force plotted with a linear magnitude scale (obtained by integrating the mobility curve). The integrated curve in Fig. 2 (but not the mobility curve) clearly predicts the two peaks P1 and P2 as dominating and well-defined in the 400–600 Hz range.

### B. Optical measurements

At Luleå University of Technology, the optical measuring technique TV holography was used for the vibration shape measurements. TV holography,<sup>20–22</sup> often referred to as electronic holography, electro-optic holography, electronic speckle pattern interferometry (ESPI) or digital speckle pattern interferometry (DSPI), is a real-time, non-contact, and full-field measuring technique. The method is suitable for vibration analysis both for out-of-plane and in-plane vibrations. The measuring principle is based on phase-stepped holographic interferometry. In Fig. 3 the optical unit is presented. The violin is illuminated by a frequency doubled Nd:YAG laser (wave length 532 nm). The image field interferes with a smooth reference beam from an optical fiber onto a charge-coupled device (CCD) detector. The piezoelectric mounted mirror (PZM1) introduces optical phase steps of 90° between subsequent TV frames. The four latest frames are processed on-line by the system. The measurements are presented as high-quality interferograms on a TV

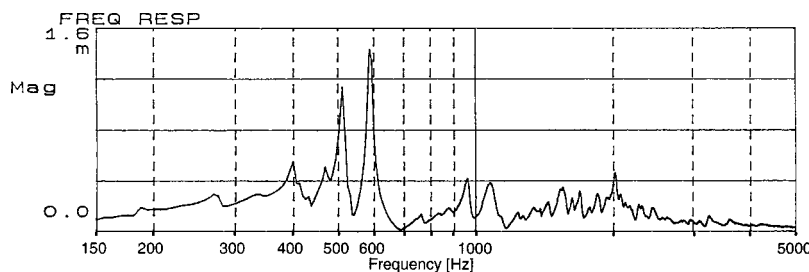


FIG. 2. Frequency response (displacement/force), the mobility curve of Fig. 1 integrated. Note the linear scale of the vertical axis.

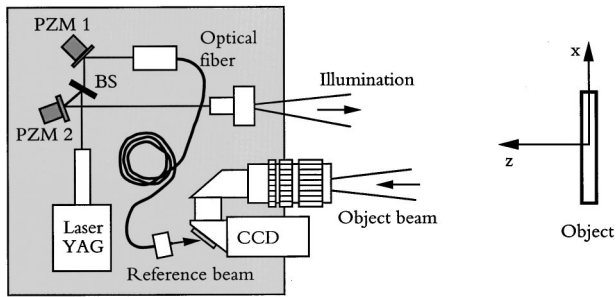


FIG. 3. The TV holography system used for the optical measurements sensitive for displacement in the  $z$  direction. Beam splitter, BS; piezoelectric mounted mirror, PZM. The measurements are performed on a vibration-isolated table.

monitor as in “real time.” The interferograms contain information about the spatial distribution of the vibration amplitude of the object.

With TV holography, vibration analysis of sinusoidally vibrating objects is easily performed. The violin is excited by an external electro-dynamical system at single frequencies. A small magnet (40 mg) is glued to the treble side of the bridge and an electric coil is used to introduce sinusoidal vibrations. The frequency is changed until a strong response is found. This occurs when a maximum number of interference fringes appear in the interferogram.

With the optical setup used here (see Fig. 3), only out-of-plane displacements, which are in the  $z$  direction, are detected. Placing mirrors on both sides of the violin, (see Fig. 4) the out-of-plane motion of the top, the back plate, and the ribs can be studied simultaneously. However, since the top and back plate of the violin body are arched, in-plane movement of the plates can also be detected by the direct viewing direction (see Fig. 4).

The Bernardel violin is mounted on the holographic table as indicated in Fig. 3 in a “cage,” see Fig. 5, with the driving unit fixed to the cage and the violin. The cage made it possible to turn and to view the violin in any direction without shifting the driving. In the cage the violin body is squeezed between the four rubber blocks, two at the bottom end and two at the upper end. The bar supporting the upper blocks is not touching the neck. The strings are damped with small rubber wedges (A- and E-strings with cloth only). It is excited sinusoidally in the frequency range from 300–650 Hz. The response of the violin body (top plate, back plate,

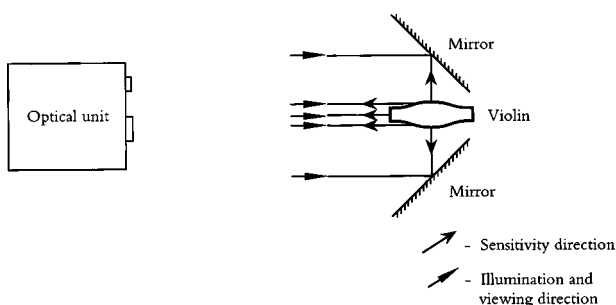


FIG. 4. With two mirrors placed on each side of the violin, both the top and back plates can be studied simultaneously as well as the ribs. Notice that the in-plane components of the vibrations of the arched plates are measured in the direct viewing direction.



FIG. 5. The violin mounted in the measurement cage. The violin is squeezed between four rubber blocks at the top end and bottom end of the ribs, seen in the figure. The driving unit excites a small magnet glued to the treble side of the bridge. The cage made the measurements flexible and more reliable.

and bass side of the ribs) is measured at those frequencies where the peaks in the bridge mobility was found. The phase modulating mirror PZM2 in Fig. 3 was used to determine the relative phase of the in-plane and out-of-plane components of the body vibration. This is repeated for each strong response. The cage is then rotated  $180^\circ$  and the measurements are repeated for the opposite side (top plate, back plate, and treble side of the ribs) with unchanged excitation conditions.

## II. AIR RESONANCES

In the bridge mobility curve a minor peak was found at 400 Hz and two dominating peaks at 512 and 587 Hz; compare Figs. 1 and 2. For practical reasons and in honor of Frederick Saunders (principal peak P) we shall use the working labels P1 and P2 for the dominating peaks. Furthermore, a small peak was found at 469 Hz followed by a small minimum.

The first air resonance, the Helmholtz mode A0, at about 275 Hz is easily found in the bridge mobility curve, being a part of the peak–dip at corresponding frequency in Fig. 1. This mode radiates sound efficiently and acts as a breathing mode. The second air resonance, A1, at about 470 Hz, is not that easy to locate in the bridge mobility curve.<sup>23</sup> The peak related to the ODS of A1 should be close in frequency to the dominating P1 peak. The A1 mode has a node across the waist of the violin, i.e., at the  $f$ -holes. Thus, the A1 mode does not radiate much through the  $f$ -holes and is difficult to separate from the dominating wall radiation at the P1 peak. As A1 does not radiate, its  $Q$  factor is high (magnitude 50). Thereby, the A1 mode may indirectly influence the radiation noticeably by affecting the wall vibrations.

The week after the holographic measurements, the A1 resonance was measured to 473 Hz with a probe microphone and a small sound emitter in the air cavity between lower bouts. Exciting the violin by an external sinusoidal sound field (a loudspeaker 30 cm outside the lower bout on the bass side), the A1 peak was found within the air cavity, i.e., A1 couples to the external sound field. Bridge vibrations also showed strong responses at two peaks, 512 and 592 Hz, the second peak being the higher one. In more detailed measurements, a week later the A1 peak was found at 475 Hz, P1 at 519 Hz, and the P2 at 600 Hz, i.e. the same peaks but at slightly higher frequencies. In bridge mobility measurements a weak pole/zero was found at 469/475 Hz, thus indicating traces of the A1 mode.

### III. OPERATING DEFLECTION SHAPES OF THE VIOLIN

Three strong maxima in the bridge mobility, measured with TV holography, were found at 400, 515, and 591 Hz. The frequency and magnitude of these prominent maxima agree well to the peaks found in the frequency-response measurements. Also, three minor maxima were found at 417, 441, and 465 Hz. The magnitude of each peak is of less importance since the presence or nonpresence of in-plane vibrations are the subject for this investigation (the prominent peaks have higher magnitude than the minor maxima).

In Figs. 6 and 7 the result of the TV holography measurements is summarized. A section sketch of the violin, at the bridge looking towards the upper end, is shown in the center. Above, below, on the right- and left-hand sides, interferograms of the vibration amplitudes are shown. The interferograms are orientated in such a way that the soundpost is placed on the same (the right) side in the top and lower interferograms. The right- and left-hand-side interferograms show the view as seen from the outside of the violin. The information obtained about the in-plane and the out-of-plane components of the vibrations, both amplitude and phase, is summarized in the section sketch in the center. Full lines indicate the section shape at rest and the dashed lines indicate the ODS in principle (highly exaggerated deformations).

The ODSs of the section through the violin at the bridge at the resonance maxima are shown in Figs. 6 and 7. This section of the violin is somewhat elliptic (top and back plates are arched). The longer axis of the ellipse is, however, cut by the ribs. The vibration amplitudes are highly exaggerated and the scales in the vertical and horizontal direction (slightly different) are adjusted to clearly illustrate the ODSs. The deformation sketch is further simplified to give informative pictures. The joints between the ribs and the plates are drawn as hinged, although it is known that clamped is a closer approximation. For the same reason the sections of the ribs are drawn straight, although the ribs must bend somewhat in the deformed shape.

In Figs. 6 and 7 it is seen that if the nodal points at the violin body are connected, nodal diameters will appear, connecting opposite points with zero vibration amplitude. This vibration behavior corresponds to that of a vibrating cylindrical or elliptical tube.

In Fig. 6(a) at 400 Hz (called C2 in Fig. 1) the nodal diameters appear as orthogonal lines. The vibration of the midsection of the violin can therefore be described as if it were an elliptical tube vibrating in a two-diameter mode. There is one horizontal nodal diameter through the violin ribs. The other nodal diameter through the top and back plate is located between the two feet of the bridge, resulting in a rocking motion of the bridge. Moving the cross section from the bridge plane, the vibration amplitude diminishes and approaches zero between the bouts in the upper and lower parts of the violin body.

The ODS at 515 Hz, Fig. 6(b) (at P1 in Fig. 1), also has two nodal diameters. The directions of the nodal diameters are, however, changed and not in right angle. Now, both diameters are passing through the top plate. The whole back plate is vibrating in phase with a larger amplitude at the ribs than in the center. The in-plane movement of the top plate at the ribs is quite high. Note that this mode has no nodal diameter between the two feet of the bridge, but has a node at the treble foot. This means that the bridge is rocking (limping) even if the feet are not moving out-of-phase. At sections through the violin, outside the bridge plane, the nodal diameters will change their position compared to the plane at the bridge. At the upper end the vibration vanishes.

At 591 Hz an ODS with three nodal diameters is found, see Fig. 6(c) (at P2 in Fig. 1). Two of the nodal diameters are going through both the top and back plates and one is passing through the ribs. One nodal diameter is situated between the two feet of the bridge; thus, the ODS is excited by a rocking motion of the bridge. This mode, as the mode at 400 Hz, reminds one of the ODSs of an elliptic tube, now with three nodal diameters. For this ODS, as well as the others, the nodal diameters will change their position when moving the cross section towards the upper and lower bout. At the upper and lower ends of the violin body, the vibration have vanished.

All ODSs in Fig. 6 agree well with the eigenmodes obtained numerically of a complete violin model with plates, ribs, *f*-holes, bassbar, etc. but without air,<sup>17,18</sup> compare Fig. 6 of this paper and Fig. 6 in the latter reference. Information about the entire violin body (top plate, back plate, and the ribs) is found more in detail in Ref. 17.

In addition to the ODSs at the main peaks of Figs. 1 and 2, there are a number of minor ODS maxima shown in Fig. 7. The ODSs are sketched with dashed lines. Note that the scales of the horizontal and the vertical vibrations are not equal. All these ODSs recall the ODS at 400 Hz, compare Fig. 6(a). The difference is that the horizontal nodal diameter through the ribs is slightly angled and at 441 Hz the horizontal nodal diameter is moved towards the top plate.

### IV. DISCUSSION

Only one violin is studied in this investigation. Since the selected violin has typical properties of a good violin, we expect the results to be quite general.

Earlier, Saunders found a main resonance peak in the violin, which was labeled P as the principal resonance peak. Later measurements have shown that there are at least two resonances and we have introduced P1 and P2 as working

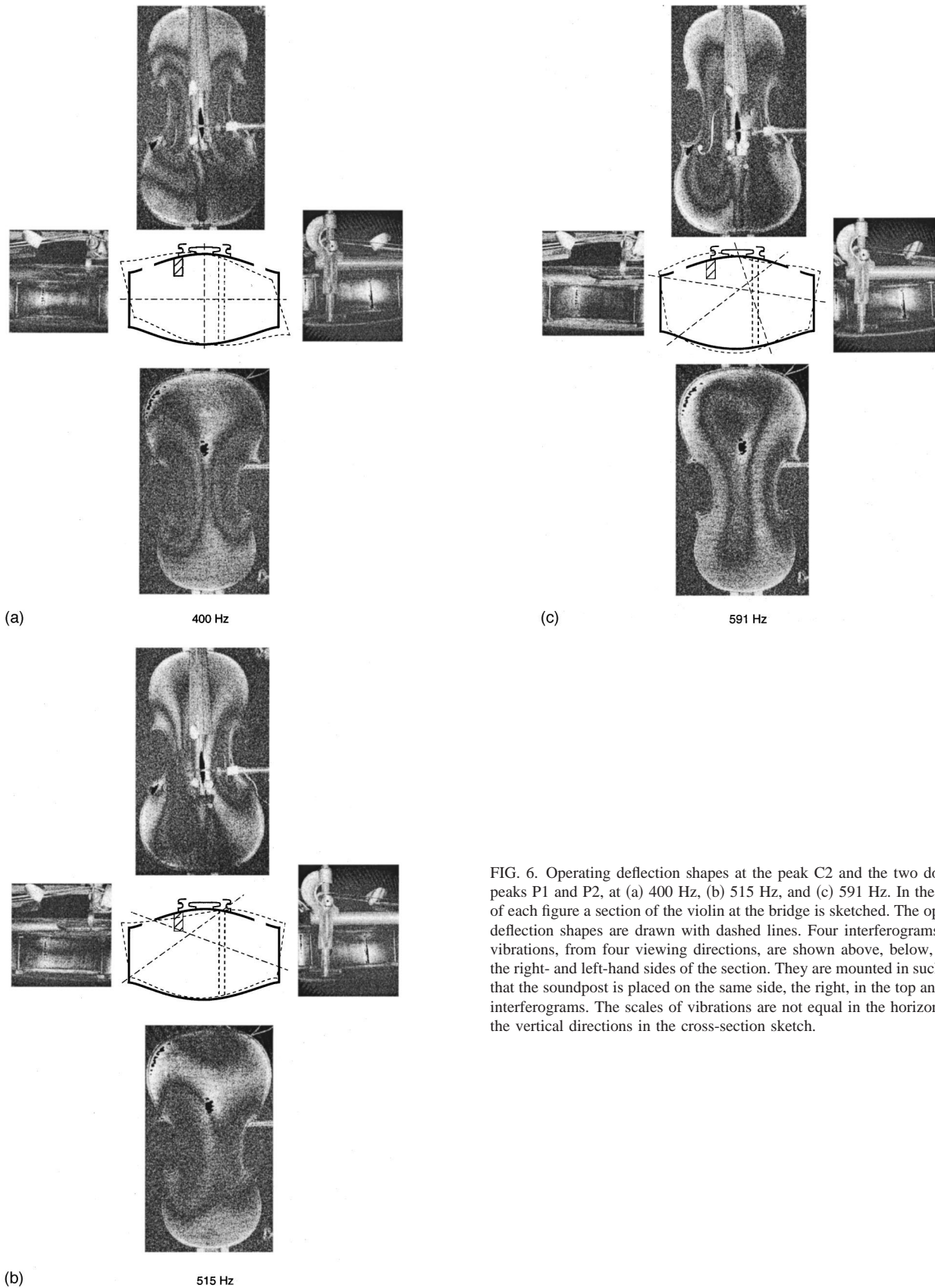


FIG. 6. Operating deflection shapes at the peak C2 and the two dominant peaks P1 and P2, at (a) 400 Hz, (b) 515 Hz, and (c) 591 Hz. In the middle of each figure a section of the violin at the bridge is sketched. The operating deflection shapes are drawn with dashed lines. Four interferograms of the vibrations, from four viewing directions, are shown above, below, and on the right- and left-hand sides of the section. They are mounted in such a way that the soundpost is placed on the same side, the right, in the top and lower interferograms. The scales of vibrations are not equal in the horizontal and the vertical directions in the cross-section sketch.

labels for the two dominant peaks. The two peaks are well-established in frequency-response curves both of bridge mobility,<sup>2,12</sup> and of sound radiation.<sup>2,11</sup> Therefore, the labeling P1 and P2 is introduced and suggested to replace our

earlier, somewhat misleading, labeling and more or less confusing labelings by others. It is, however, difficult to describe corresponding resonances, or rather the observed ODSs. This difficulty leads to the first question of this investigation.

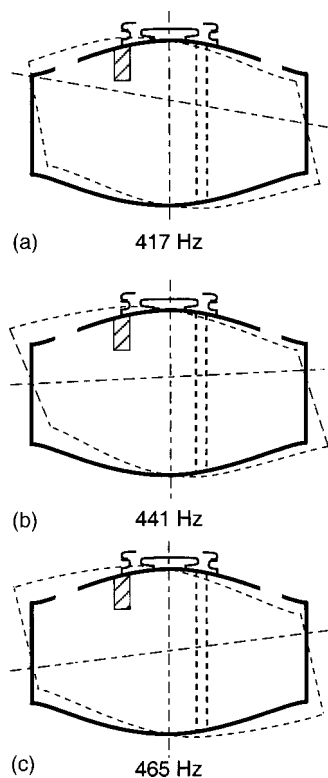


FIG. 7. Operating deflection shapes of the three minor peaks studied at (a) 417 Hz, (b) 441 Hz, and (c) 465 Hz. The section is through the violin at the bridge and the soundpost as indicated. Dashed lines are the operating deflection shapes at this section. The scales of vibrations are not equal in the horizontal and the vertical directions in the cross-section sketch.

What are the ODSs at the frequencies of vibration maxima, if one includes both in-plane and out-of-plane vibrations of the violin plates (shells) for sinusoidal excitation at the violin bridge in the main bowing direction?

Usually the acoustics of the violin body is treated assuming only out-of-plane vibrations. The violin body is regarded as two plates, the top and the back plate, which are joined to sides, the ribs. This picture of the violin follows the typical steps of production. Mechanically it is an oversimplification of the construction and the function of the violin. The violin plates are arched, i.e., shells, joined at the ribs. The violin body forms a rather elliptical tube (in the cross section) in the important bridge region. Thus, both out-of-plane and in-plane vibrations should be involved. The measurements also showed that both out-of-plane and in-plane vibrations of the “plates” and the ribs are of the same magnitude and thus both are important. With care taken of both out-of-plane and in-plane vibrations a more complete map of the ODSs at vibration maxima is obtained; see Figs. 6 and 7.

The ribs are integrated parts of the violin body and are participating in the ODSs of the violin body. In the low-frequency range a nodal line is often found along the center of the ribs. The vibration behavior of the violin body at a cross section at the bridge is very much like an elliptic cylinder; compare vibrational modes in noncircular cylinders.<sup>24</sup> In experiments by Niewczyk and Jansson, not yet reported, this section was found to be of the greatest importance for the P2 peak. The ODSs can be described with a number of nodal diameters, located in different ways. The increase from

two nodal diameters (at 400 Hz) to three (at 591 Hz) in the frequency range 400–600 Hz is quite interesting, since this increases the number of modes present in this important interval. For a tubular section it is not possible to have an odd number of nodes around the “tube” for an eigenmode, since this would mean that there are two neighboring antinodes with a nodal line between them having equal and not opposite phases. The reported experimental results are in good agreement with finite element calculations including both out-of-plane and in-plane vibrations.<sup>17</sup>

A nodal line at the center point between the two feet of the bridge indicates a pure rocking (rotational) motion of the bridge, a motion which is efficiently excited in the main (horizontal) bowing direction but not in the vertical direction (especially not in the center of the bridge). With the nodal line not at the center point between the bridge feet, the bridge vibrates more or less up and down (translational motion) i.e., a motion inefficiently driven in the horizontal, the main bowing direction (but efficiently driven by the vertical bowing component). It was found that the bridge feet move out-of-phase for all the ODSs in the studied frequency range, except the ODS at 515 Hz (P1). At 515 Hz there is a nodal line under the treble bridge foot and now the bridge is only partly rocking; rather, it is limping. This limping motion is quite efficiently excited by the main (horizontal) bowing force, but also by the vertical bowing force component.

Note that the ODSs at 515 Hz (P1) and 591 Hz (P2) both are breathing ones, i.e., they are acting mainly as a breathing “spheres,” with some nonradiation motion superimposed.

The second question of this investigation was the following. With the improved understanding of the vibrations, is it possible to determine the role of the A1 resonance in supporting radiation of the violin?

Electrodynamical measurements show that A1 is at least weakly driven via bridge excitation in the main bowing direction, which is in agreement with the other findings of this study and those of a previous study.<sup>23</sup> The sound pressure in the cavity between the lower bouts and the upper bouts are in-phase 260–320 Hz (A0 motion) but out-of-phase 460–520 Hz (A1 motion). In the previous study it was shown that the air resonance A1 forced vibrations of the cavity walls. The wall vibrations did not have the same phase distribution as the air resonance itself. Rather, the ODS of the violin body looks like the structural modes of the nearby frequency range.

In the present study the air resonance A1 is located close to the P1 peak and it also shows up as a minor ODS maximum (Fig. 7). Since P1 emits sound, and the frequency responses (Figs. 1 and 2) show a peak at 479 Hz in the flank of the dominating P1 peak, it is plausible that the air resonance A1 constitutes a contribution to sound radiation by interaction with the structure. Reciprocal measurements of the radiation, with a probe microphone inside the violin and a loudspeaker outside, also support that the A1 mode is a non-negligible part of the sound-radiating system, but still not a major part. The present investigation of a complete varnished violin also is a challenge for the TV holography system itself.

The violin used in this study was not prepared in any

way. It was complete with strings, and a shiny and fairly dark varnish, i.e., a difficult investigation object. Still, the TV holography system recorded high-quality interferograms. Together with the phase-modulating technique, the relative phase of the vibration in the front plate, in the back plate, and in the ribs could be determined. The TV holography system answered well to the challenge of recording ODSs of a difficult investigation object, a complete violin without any preparations.

## ACKNOWLEDGMENTS

The Swedish National Board for Industrial and Technical Development (NUTEK) supported this work. We also thank the referees for valuable comments.

- <sup>1</sup>E. Arnold and G. Weinreich, "Acoustical spectroscopy of violins," *J. Acoust. Soc. Am.* **72**, 1739–1746 (1982).
- <sup>2</sup>L. Cremer, *The Physics of the Violin*, translated by J. Allen (MIT Press, Cambridge, MA, 1984).
- <sup>3</sup>F. A. Saunders, "Recent work on violins," *J. Acoust. Soc. Am.* **25**, 491–498 (1953).
- <sup>4</sup>C. M. Hutchins, "The Physics of Violins," in *Scientific American* (Nov. 1962), pp. 78–93.
- <sup>5</sup>E. Jansson, N.-E. Molin, and H. Sundin, "Resonances of a violin body studied by hologram interferometry and acoustical methods," *Phys. Scr.* **2**, 243–256 (1970).
- <sup>6</sup>E. Jansson, "On higher air modes in the violin," *Catgut Acoustical Society Newsletter* No. 19 (May 1973), pp. 13–15.
- <sup>7</sup>C. M. Hutchins, "A measurable controlling factor in the tone and playing qualities of violins," *J. Catgut Acoust. Soc.* **1**(4), 10–15 (1989).
- <sup>8</sup>J. Alonso Moral and E. V. Jansson, "Eigenmodes, input admittance, and the function of the violin," *Acustica* **50**, 329–337 (1982).
- <sup>9</sup>K. D. Marshall, "Modal analysis of a violin," *J. Acoust. Soc. Am.* **77**, 695–709 (1985).
- <sup>10</sup>G. W. Roberts, "Finite element analysis of the violin" and "Modelling air-cavity modes of stringed instruments," extract from "Vibrations of shells and their relevance to musical instruments," Doctoral thesis, University College, Cardiff, Wales, UK (1986).
- <sup>11</sup>H. Dünwald, "Ein erwarteteres Verfahren zur objektiven Bestimmung der Klangqualität von Violinen," *Acustica* **71**, 269–276 (1990).
- <sup>12</sup>E. V. Jansson, "Admittance measurements of 25 high quality violins," *Acust. Acta Acust.* **83**, 337–341 (1997).
- <sup>13</sup>A. Runnemalm, N.-E. Molin, and E. V. Jansson, "Operating deflection shapes and the function of the violin," *TMH-QPSR* 3/1998, Dept. of Speech, Music, and Hearing, Royal Institute of Technology (KTH), Stockholm (1998), pp. 5–17.
- <sup>14</sup>N.-E. Molin, H. O. Saldner, and E. V. Jansson, "Eigenmodes and electronic holography on three violins," *STL-QPSR* 4/1993, Speech Transmission Laboratory, Royal Institute of Technology (KTH), Stockholm (1993), pp. 1–6.
- <sup>15</sup>H. O. Saldner, N.-E. Molin, and E. V. Jansson, "Vibration modes of the violin forced via the bridge and action of the soundpost," *J. Acoust. Soc. Am.* **100**, 1168–1177 (1996).
- <sup>16</sup>H. O. Saldner, N.-E. Molin, and E. V. Jansson, "Sound distribution from forced vibration modes of a violin measured by reciprocity and TV holography," *CAS J.* **3**(4), 10–16 (1997).
- <sup>17</sup>J. Bretos Linaza, "Propiedades acusticas de interes musical en violines y xylofonos: Medidas experimentales y modelizacion por elementos finitos de su mecanica vibrational" (in Spanish). Department of Applied Physics II, University of Pais Vasco (1996).
- <sup>18</sup>J. Bretos Linaza, C. Santamaria, and J. Alonso Moral, "Vibrational patterns and frequency responses of the free plates and box of a violin obtained by finite element analysis," *J. Acoust. Soc. Am.* **105**, 1942–1950 (1999).
- <sup>19</sup>J. Bretos Linaza, C. Santamaria, and J. Alonso Moral, "Vibrational patterns of a violin-shaped air cavity obtained by finite element modeling," *Acust. Acta Acust.* **85**(4), 584–586 (1999).
- <sup>20</sup>K. A. Stetson, W. R. Brohinsky, J. Wahid, and T. Bushman, "An electro-optic system with real-time arithmetic processing," *J. Nondestruct. Eval.* **8**(2), 69–76 (1989).
- <sup>21</sup>K. A. Stetson, "Theory and applications of electronic holography," in *Proceedings SEM Conference on Hologram Interferometry and Speckle Metrology*, Baltimore, MD, 1990, pp. 294–300 (1990).
- <sup>22</sup>T. K. Kreis, *Holographic Interferometry* (Akademie GmbH, Berlin, 1996).
- <sup>23</sup>G. Bissinger and C. M. Hutchins, "A1 cavity-mode-enhanced fundamental in bowed violin and viola sound," *J. Catgut Acoust. Soc.* **1**(2), 11–13 (1988).
- <sup>24</sup>U. J. Hansen, T. D. Rossing, and G. J. Jelatis, "Holographic studies of vibrational modes in noncircular cylinders," in *Proceedings International Symposium on Musical Acoustics*, 28 Aug.–1 Sept. 1992, Tokyo, Japan (1992).

# Vibration characteristics of pipe organ reed tongues and the effect of the shallot, resonator, and reed curvature

G. R. Plitnik<sup>a)</sup>

Department of Physics, Frostburg State University, Frostburg, Maryland 21545

(Received 12 April 1999; revised 13 January 2000; accepted 1 February 2000)

Pipe organ reed pipes sound when a fixed-free curved brass reed mounted on a shallot connected to a resonator is forced to vibrate by an impressed static air pressure. Five sets of experiments were performed in order to investigate the influence of the most important parameters which could affect the tone of a reed pipe. First, the phase difference between the pressure variation in the shallot and the boot, and its relationship to the motion of the reed tongue were analyzed to compare their phases and their spectra. Next, the frequency dependence of the reed on three basic parameters (reed thickness, its vibrating length, and the imposed static air pressure) was investigated in an attempt to determine an empirical equation for the frequency. For each trial, two of the variables were kept constant while the third was altered in order to construct an equation giving frequency as a function of the three variables. Third, experiments were conducted using three different types of shallots: the American (or English) style, the French style, and the German style. The results show that for each shallot, the frequency increases linearly with thickness and linearly with air pressure (over the normal operating range of the reed). For each of the shallots, frequency varies inversely with length when the other variables are held constant. The effect on the reed spectrum of using the three different types of shallot was also investigated, as was the effect of reducing the interior volume of each type. Progressively filling the shallot interior generally decreases the frequency of the vibrating reed. The effect of the resonators on frequency and spectrum was studied because the resonator is an integral part of the resulting tone; virtually every reed stop has some type of resonator. The resonator tends to raise the  $Q$  of the impedance peaks and reduce the fundamental frequency. Finally, the influence of the type and degree of curvature on reed vibration was briefly examined; increasing the reed curvature tends to decrease the vibration frequency and increase the sound intensity by creating a richer spectrum. © 2000 Acoustical Society of America.

[S0001-4966(00)01505-8]

PACS numbers: 43.75.Np [WJS]

## INTRODUCTION

During the 19th century, scientific advances in the rapidly growing field of acoustics began to unveil some of the mysteries of sound production in musical instruments. It became possible, with aid of resonators, to dissect a tonal structure and quantify sounds which previously could only be qualitatively assessed. The mood of the time was to assume that science would set all things in order; thus, when the searchlight of science was focused on pipe organs, it was with the intention of transforming organ building from art to science so that empirical rules could be supported by theoretically derived equations.

Arguably, organ reed pipes were first subjected to scientific analysis by the German experimental physicist Wilhelm Weber in 1830; this research guided him to an essentially correct theory for the interaction between a vibrating reed and the input impedance of the resonator. The results of Weber's investigations, as well as an analysis of a vibrating reed tongue, are summarized in Smith's book *Sound in the Organ and the Orchestra* (1911). During the 19th century very few experimentalists concerned themselves with organ reed pipes, and the limited theory developed was of scant use to

organ builders (see Audsley, 1905, Vol. I, pp. 401–402), as even a cursory examination of Tyndall and Helmholtz, two of the 19th century's greatest investigators of musical acoustics, reveals. In *The Science of Sound*, Tyndall (1875) devotes a mere two pages to organ reed pipes, and even this is marred by several misleading statements which reveal his lack of understanding of organ building practice (Tyndall, pp. 234–236). Helmholtz (1954) fares only slightly better; *On the Sensations of Tone* devotes only about four pages to organ reed pipes and most of this is devoted to free, rather than striking, reeds. (Helmholtz mistakenly believed that striking reeds were an historical curiosity which had passed from general use among organ builders; in fact, the reverse was more nearly true.) His major contribution to the science of beating reeds was his discussion of the reed-resonator interaction and the formulation of a general theory for determining the phase angle between the pressure maximum and the vibration for inward (as well as outward) striking reeds (Helmholtz, pp. 390–394).

When, in 1888, the Rev. Max Allihn revised and extended Topfer's 1833 tome on the art of organ building (Topfer, 1939), one entire book (out of seven) was devoted to reed pipes (Allihn, pp. 277–396), and no fewer than six pages to Weber's theory (Allihn, pp. 326–331). Despite this heroic attempt to understand the functioning of reed pipes,

<sup>a)</sup>Electronic mail: gplitnik@fre.fsu.umd.edu

the great organ expositor Audsley was compelled to conclude (by quoting Allihn) that despite all the mathematical formulas, “The subject of reed pipes is a question so intricate and obscure that it creates a sensation as of one trying to find a path through a forest with which one is unacquainted, and in which numberless stray paths lead from the right road” (Audsley, 1919, pp. 453–454). Audsley thus dismisses 70 years of organ reed research by wryly commenting “As we are not prepared ... to enter the forest, with its ‘numberless stray paths’ which lead one away to pitfalls of false conclusions and thickets of perplexity, we shall rest content in directing the ardent student to the pages of the works above quoted from, for the special information they afford” (Audsley, 1905, Vol. II, p. 589).

The most recent incarnation of Topfer, continuing his preoccupation with scaling rules for organ pipes, was the excellent book by Ellerhorst, first published in 1936 (Ellerhorst, 1966). Unfortunately, in this 850-page tome a scant 40 pages is devoted to reed pipes, and half of this is descriptive material. The other half attempts to give a scientific understanding for the scaling of reed tongues and resonators. Although replete with equations and formulas, no real advance in the understanding of the physics of organ reeds is presented.

The first real 20th century advance in the scientific understanding of vibrating reeds of all types was promulgated by Bouasse, who devoted a large part of *Instruments a Vent* (Bouasse, 1929) to a comprehensive review and an expansion of Weber’s seminal work on organ reeds, as well as extending it to include the cane reeds of woodwind instruments. Among other results, he reported that the frequency of a reed pipe increased as the blowing pressure was raised. However, since Bouasse was more interested in orchestral reeds than in organ pipes, most of the subtle, but extremely important parameters unique to organ reeds (reed curvature, type of shallot, etc.) are not even mentioned. Several years later, Bonavia-Hunt (1933) identified two dozen parameters which allegedly effect the power and tonality of reed pipes, but no attempt was made to explain the details of exactly how most of these factors influence speech and tone quality.

In 1905, Audsley, after explaining the vibrating mechanism of a reed pipe remarked, “The operation of the mechanism above described appears to be extremely simple, while the acoustical phenomena which attend it have never been satisfactorily explained” (Audsley, 1905, Vol. I, p. 397). Bouasse’s work notwithstanding, Audsley’s comment remained essentially true for more than four decades, until two researchers, Bonavia-Hunt and Homer, undertook the first truly comprehensive study of organ reed pipes. The results of their experiments, in addition to a wealth of information about the voicing of reed pipes, appeared as *The Organ Reed*, unfortunately now out of print (Bonavia-Hunt, 1950). The entire second half of this work, subtitled “The Mechanical Properties of Reed Pipes,” summarizes the results of an elaborate set of experiments performed to determine how the reed tongue vibrates, how wind pressure influences the frequency, the nature of the air pulses produced, and the equivalent sound waves emitted (synthesized by means of a mechanical integrator!). Based on experimental results, they

concluded: (1) the reed tongue vibrates in a regular and uniform manner with no high-frequency flexing or rippling; (2) the tongue does not strike the shallot when vibrating; (3) the waveform of the vibrating tongue is not simple harmonic; (4) the center of vibration is not the rest position of the reed, but a position which increasingly approaches the shallot face as the static pressure is increased; (5) the force exerted on the tongue by the impinging wind is uniformly variable over each cycle of vibration (i.e., from maximum when the reed is closed to minimum when fully open) and variable over the length of the reed; (6) the curvature of a well-voiced reed tongue is not parabolic; (7) the rate of change of air flow through the shallot is a function of both the instantaneous position of the tongue and the type of shallot; (8) as the tongue swings outward, the increased aperture tends to compensate for the pressure drop in the boot, thus yielding a constant rate of flow when the reed is more than half open; (9) the air in the boot also vibrates, which has a non-negligible effect on the sound; and (10) the addition of a resonator suppresses the high dissonant overtones and boosts the lower harmonics.

An elaborate set of experiments with a 4-ft conical reed, coordinated and reported upon by Jann and Rensch (1973), investigated the effect of varying the following parameters: resonator diameter, shallot shape, shallot diameter, width and shape of shallot opening, cross section of shallot, width and thickness of the reed, length of the boot, toe hole diameter, resonator material, and wind pressure. Voicing differences were excluded, insofar as possible, by using the same reed tongue whenever feasible and by giving different tongues the same curvature. Although a plethora of data was collected, differences resulting from varying one parameter at a time were described on a continuum of vague psychoacoustic terms such as “dark to bright,” “round to dry,” or “O-like to E-like.” Although sets of spectra were included in an attempt to correlate acoustical data to the above vague terms, the spectra, lacking frequency and amplitude scales, were impossible to interpret, rendering them even more useless than the psychoacoustic terms. Perhaps the kindest words which can be written about this paper is that although the concept was excellent, the realization left much to be desired.

In an excellent survey article on the acoustics of reed pipes, Lottermoser (1983) discusses the mechanical properties of the reed tongue, the pressure variation in the boot of a vibrating reed, and the mechanism by which oscillation is sustained. He also covers the important concept of the acoustic input impedance of the resonator and how the shallot and tongue assembly interact with the resonator as the tongue length is changed. Unfortunately, this is merely a survey article which does not pretend to advance the scientific understanding of the organ’s lingual pipes. In another survey article concerned mostly with wind systems and organ flue pipes, Angster *et al.* (1997) briefly examined the attack of each partial for two reed pipes, but the results were considerably less interesting than for flue pipes.

Hirschberg *et al.* (1990) attempted to correlate measurements on an organ reed pipe to a quasistationary flow model based on a two-dimensional Borda tube. Although the theory



was necessarily somewhat simplistic, it did give fairly good agreement to the experimental results as well as provide tentative explanations for much of the observed behavior. Of particular interest was their careful documentation of the approximately linear increase of frequency with blowing pressure for a reed without its attached resonator. Although discrete jumps were observed when the resonator was added, over most of the range examined the frequency dependence on pressure was unchanged by the presence of a resonator, thus indicating that the pressure-frequency dependence is not primarily controlled by the acoustics of the resonator. No attempt was made to examine the influence of the important parameters of shallot type, reed thickness, or reed curvature on frequency or tonal quality.

Although almost a century has passed since these words were written, Audsley's statement still rings true. "It is quite evident that both the dimensions and forms of the tubes or resonators have a strong modifying influence on the sounds produced by the associated reeds, imparting marked varieties of tonal-coloring and distinctive timbres to the resultant tones; but exactly in what manner, or by what peculiar modes of vibration of the air columns in the resonators, the different tones are produced remains a mystery to the most painstaking investigators" (Audsley, 1905, Vol. I, p. 404). Perhaps the present state of affairs regarding the scientific understanding of organ reeds and its potential use to organ builders was best summarized by a well-known organ expositor thus: "We know a great deal about the laws that govern vibrating metal tongues; we are also well-informed about resonance in tubes, etc. But concerning the interaction of all these details, vast numbers of additional factors assert themselves, and the theory becomes hopelessly inadequate for any practical application. ... 'With an experience coefficient, the formula might have been corrected and put to some use, if it was not already so complicated that it was unsuitable for any practical purpose'" (Andersen, 1969). The present work is an attempt to rectify this situation somewhat by methodically investigating the effect of varying the most important parameters determining lingual pipe organ tone. Hopefully, after the basic science of reed pipes is understood, practical applications will ensue.

### I. STRUCTURE AND FUNCTION

The structure of an organ reed (or lingual) pipe is shown in Fig. 1. The main parts are the boot, the block, the shallot, the brass reed (or tongue), the tuning wire, a short socket to hold the resonator, and the resonator (not shown in this figure). Organ reed pipes, like organ flute pipes, create a tone in a vibrating air column. However, reed pipes differ from organ flue pipes in several important ways: (1) a vibrating brass reed modulates the airstream, (2) the energy is applied at an effectively closed end of the tube, (3) all the air passes through the pipe, and (4) although the reed's vibration is almost independent of the resonator, standing waves in the resonator can exert a non-negligible influence on the vibration frequency of the reed (Strong and Plitnik, pp. 319–320); the loose coupling between reed and resonator allows the pipe to sound over a considerable range of frequencies.

The sound-producing portion of the pipe (see Fig. 2)

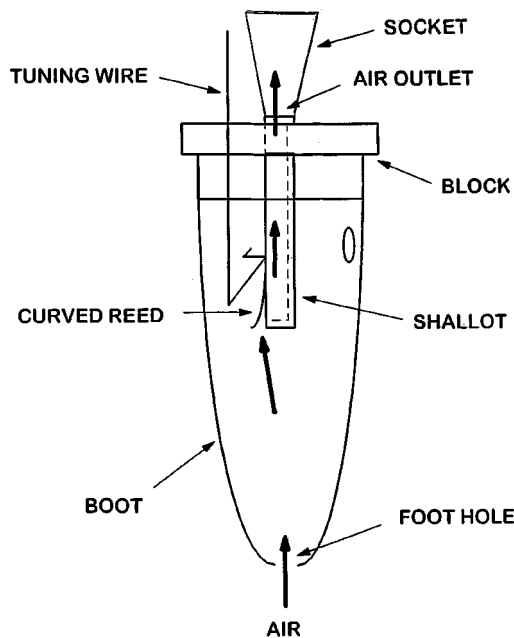


FIG. 1. Structure and components of the reed block assembly.

consists of a curved piece of brass, called the reed, which beats against the shallot at a frequency determined by the thickness of the reed, the length of its vibrating portion, the air pressure introduced into the boot, and to some extent, the reflections back from the resonator. The shallot is a flat-faced hollow tube, closed at the lower end, upon which the reed rests at a length determined by the tuning wire. When the pipe is activated, air rushes into the boot and the positive air pressure forces the reed to temporarily close against the shallot, thus reducing the air flow. The Bernoulli effect of the air flowing through the narrow opening also helps to bring the reed to closure. After the reed has been blown closed, elastic forces, due to its displacement from equilibrium, spring it back open. A returning pulse of air from the resonator may assist in this process, but it is not a necessary condition.<sup>1</sup> The opening and closing reed interrupts the air stream in a regular manner, thus producing the vibration which is the sounding frequency of the reed pipe. The puffs of air emerging

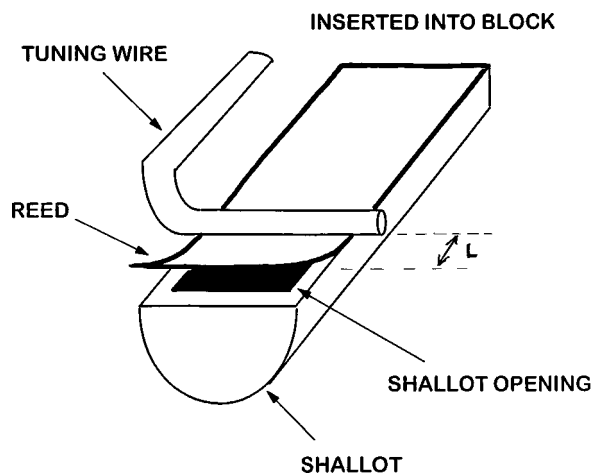


FIG. 2. Detailed view of the shallot and reed.  $L$  is the vibrating length of the reed determined by adjusting the tuning wire.

from the shallot into the resonator set up a standing wave in the resonator. Cylindrical resonators support mostly odd harmonics, while conical resonators tend to support all harmonics.

Although the natural frequencies of the vibrating reed tongue are generally set to agree with a resonant frequency of the resonator, there is considerable latitude. Also, as the tongue is lengthened, it not only vibrates at a lower frequency but with a greater amplitude, thus changing the harmonic structure of the sound. Part of the art of reed voicing is to adjust every note of a set of reed pipes to have a very similar, or only slowly varying over the compass, tone color. When the reed tongue is lengthened, the resonator must be shortened in order to adjust the frequency if the new tone color is that which is desired. If one wishes to change the loudness without modifying the tone color, one must curve the reed differently, a job which requires considerable skill.

Reed pipes are very sensitive to a variety of factors which affect the final tonal quality. Some factors under the control of the voicer which effect tonal quality are: the thickness of the reed tongue, whether there is a weight at the vibrating end, the type and position of the shallot opening, whether the shallot is filled with wax or solder, the diameter of the shallot base, and the type of curve on the tongue. A clarinet stop requires a long gradual curve on the reed, while a trumpet uses a somewhat thinner reed with a curve accelerated toward the end. A posthorn (or state trumpet), usually the loudest stop found on large organs, uses a thin reed curved very high, with a bit of extra curve at the very end to give the sound "crack and power."

It may be useful to classify organ reeds into categories because the properties of the resonators, as well as the manner in which the reeds are voiced, are quite different. One can think of reed pipes as falling into two general families termed *classic* and *romantic*, each of which may each be divided into two subgroups. The classic reeds include *chorus reeds*, used for ensemble playing or as a bright solo voice, and the *regals*. The chorus reeds serve for pipe organs the same purpose contributed by brass instruments to the orchestra. Resonators are usually a single or a double cone and are typically full length or half length, but occasionally double-length resonators are used in the upper octaves to strengthen the tone. The *regal* family of reeds, typically found only on larger pipe organs as a supplement to the chorus reeds, are characterized by having fractional length resonators (1/8 or 1/16 the fundamental wavelength) yielding a tone considerably softer than the chorus reeds, lacking fundamental, and having many overtones. The short resonators cannot support wavelengths of the lower harmonics; they serve rather to impose a formant-like envelope to the rich spectrum.

The *romantic reeds*, developed in the late 19th and early 20th centuries when organ transcriptions of orchestral music were in fashion, attempted to imitate the orchestral reed voices. Stops developed during this time (such as the French horn, orchestral oboe, bassoon, clarinet, and saxophone) actually endeavored to imitate the orchestral brass and woodwind instruments. The romantic reeds may also be divided into two categories, the *solo reeds* (orchestral imitators) and *special-effect* reeds. The special-effect reeds are to the ro-

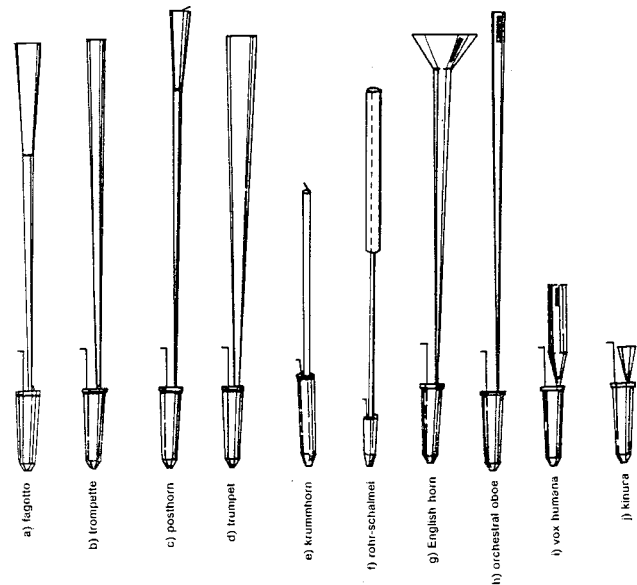


FIG. 3. An assortment of reed resonators (from Strong and Plitnik, p. 323) where a, b, c, and d are classical chorus reeds; e and f are regals; g and h romantic reeds; and i and j special-effect reeds.

mantic organ what the regals are to the baroque organ. Just as the regals may be used to add tonal color to baroque flue stops, the special effect reeds are employed to add unique and interesting timbres to romantic flues by employing fractional length resonators. The two best-known examples are the vox humana (as voiced for romantic organs) and the kinura. Figure 3 presents a diagram of the resonator shapes typical of the various tonal families.

Although many shapes and sizes of shallots have been employed in lingual pipes, there are only three major shallot styles typically employed for chorus reeds, the most common reed stops utilized on classical pipe organs. The three different types of shallots, the only types investigated in this paper, termed the American (or English), the French, and the German, are diagramed in Fig. 4. The American shallot is tapered with a triangular opening, the French shallot is narrow with parallel sides and a parallel opening, while the German style is wider with parallel sides and a large triangular opening.

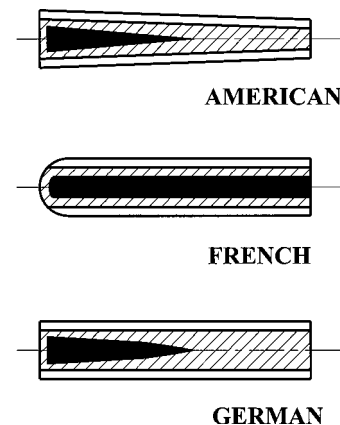


FIG. 4. The faces of the three types of shallots studied.

## II. MEASUREMENTS

Of the two dozen parameters which could affect the tone of a reed pipe, five areas (considered to be the most influential) were investigated. First, the phase relationships among the pressure variation in the shallot and the boot, and their relationship to the motion of the reed tongue was examined. These pressure variations are due to the motion of the reed, but are found on opposite sides of the air flow path. The pressure variations and the reed vibration were also analyzed to compare their phase and their spectra. Next, experiments were performed to determine an empirical equation for the frequency of the reed as a function of three basic parameters: the reed thickness, the reed length, and the static air pressure in the boot. An attempt was made to find an empirical equation giving frequency as a function of these three variables, i.e.,  $F=f(T,L,P)$ . Third, the effect on the reed spectrum of using the three different types of shallot was investigated, as was the effect of reducing the interior volume of each type of shallot by progressively filling their interiors with wax. Fourth, the effect on the spectra of adding different types of resonators, both full length and fractional length, to the reed block was studied. Finally, the effect of different reed curvatures was briefly examined.

Two Radio Shack wide-range response electret condenser microphones (270-092B) were used to obtain the pressure variations within the boot and the shallot, respectively. To obtain readings with a minimal disturbance to the signals, the microphones were placed firmly into one end of a small section of rubber tubing, while a tapered cork was placed in the other end such that the cork was against the microphone face. A piece of hollow wire insulation was then pushed into a small hole drilled through the cork. The resonant frequency of this tube is about 3000 Hz, well above the highest frequency of interest for this study. Reducing the input area to the microphone also helped to prevent overloading or clipping of the signals. One microphone (mic 1) was placed in a small hole drilled in the boot, while the other (mic 2) had the insulation tube extend through a hole in the boot through a second hole in the back of the shallot opposite the reed, as shown in Fig. 5. The signals from each microphone were simultaneously recorded on a Tektronix 2232 100-MHz digital storage oscilloscope, so that phase relationships could be analyzed. To record the effects of the shallot, reed curvature, and resonator, the microphone was carefully positioned at the sound output from the block.

The accelerometer measurements were performed on an Endevco model 122 Picomin piezoresistive accelerometer attached to the vibrating end of the reed with plastic cement. The lead wires were fed through the boot's foot hole to an Endevco model 2775A signal conditioner for processing the signal before it was sent to the oscilloscope.

The static pressure in the boot was measured through a small access hole with a water manometer. This device could measure only the static or average pressure (in inches of water<sup>2</sup>) during reed vibration as the pressure variations detected by the microphone were too small and too rapid. To operate the reed pipe, tubing from an air valve was fed into the foot hole of the boot and the desired operating pressure was set on the manometer.

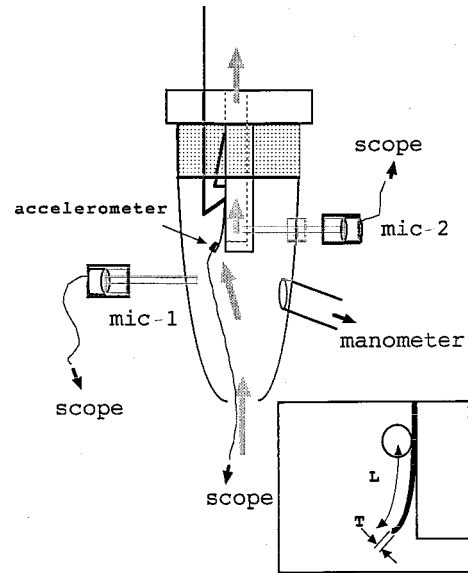


FIG. 5. The experimental setup. Mic-1 records the signal in the boot, mic-2 records the signal in the shallot, and the accelerometer the vibration of the reed tongue.

## III. RESULTS

### A. Phase relations

Phase differences between the accelerometer and the two microphones (one inside the shallot and one in the boot outside the shallot) were first investigated. For any one trial, either an accelerometer reading and the output from one microphone, or the output from both microphones were recorded simultaneously on the dual trace oscilloscope. Only qualitative measurements of phase were recorded.

The frequency of the air vibration measured in the boot was identical with that measured in the shallot, but with the expected phase shift of 90 deg (this was also confirmed by creating a Lissajous figure with the two signals). Accelerometer measurements also yielded the same frequency, and the acceleration showed no phase shift relative to the air vibration in the boot recorded by microphone mic-1 for lower static pressures. However, at the higher operating frequencies caused by increasing the driving pressures, the acceleration lagged the boot pressure by approximately 90 deg (see Fig. 6).

Fourier analysis of the waveforms from mic-1, mic-2, and the accelerometer verified that there were no substantive differences in fundamental frequency. Therefore, all further signal analyses were made using the input of microphone mic-1 only.

### B. Effect of reed thickness, pressure, and reed length

To investigate the frequency dependence of each of these three variables, two variables were held constant while the third was varied. The resulting waveform was recorded so that frequency and spectrum could be recorded. To study the relationship between reed thickness and frequency, a set of eight identically voiced reeds having the same length and width, but varying in thickness, was employed. Only the American shallot was used in tests of dependence of fre-

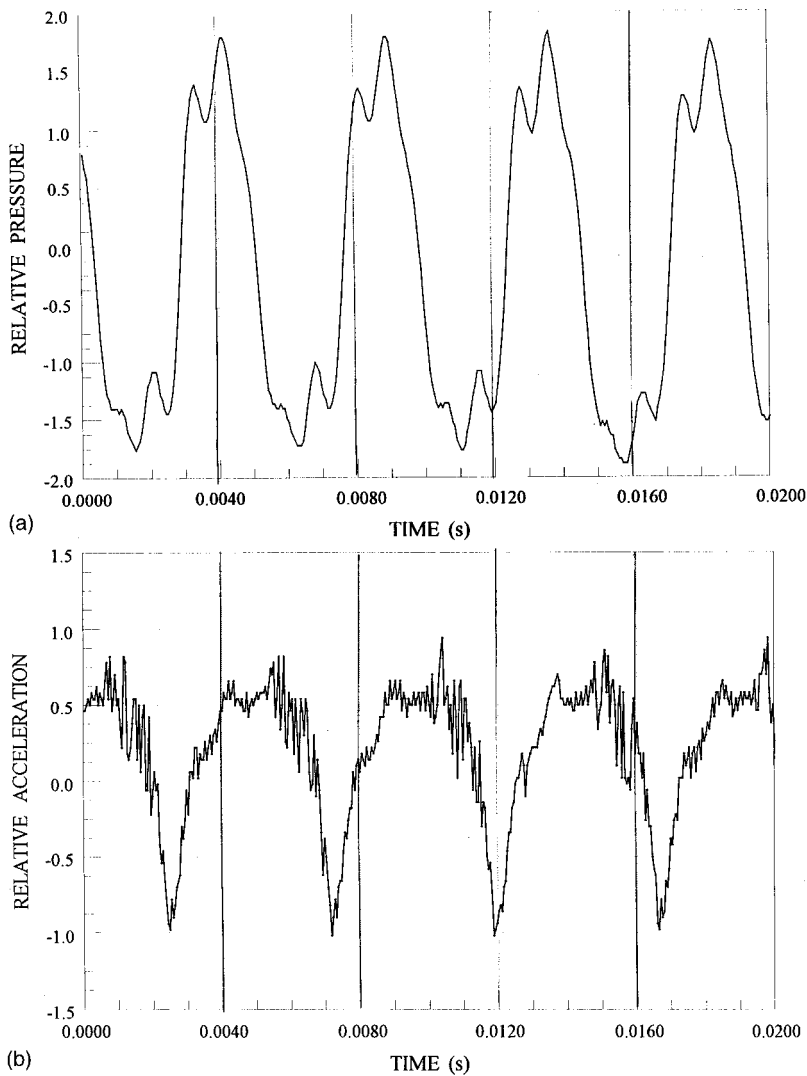


FIG. 6. (a) Pressure waveform in boot. (b) Acceleration of reed tip.

quency on reed thickness; for each trial, the tuning wire was set at a fixed position and the pressure was adjusted to identical values. After the trials were complete, the thickness was measured at several points by a micrometer and averaged.

Next, the dependence of frequency on pressure and length for each of the three different types of shallots was examined. To determine the pressure variation, one particular reed was used and the tuning wire was set at a standard position for an average operating pressure (which yielded a sound of good quality). The pressure was then varied from the lowest pressure at which the reed would sound to the maximum pressure (when the pressure is too high, the reed is permanently blown against the shallot face and no vibration ensues).

To study the frequency dependence on length, the tuning wire was used to change the length of a particular reed at a standard operating pressure. The length was changed from a position where the tuning wire is against the wooden wedge (which sounds approximately one octave below the operating frequency) to the shortest length which would still support vibration.

Figure 7 presents the results of determining the frequency dependence on reed thickness. The graph combines results of six different trials; for each trial the length of vi-

brating reed and the boot pressure were held constant. Data were collected at two different lengths and at three different pressures for each length. The lower group of data is for  $L = 2.96$  cm and the upper for  $L = 2.11$  cm. The three different points in each group correspond to three different pressures

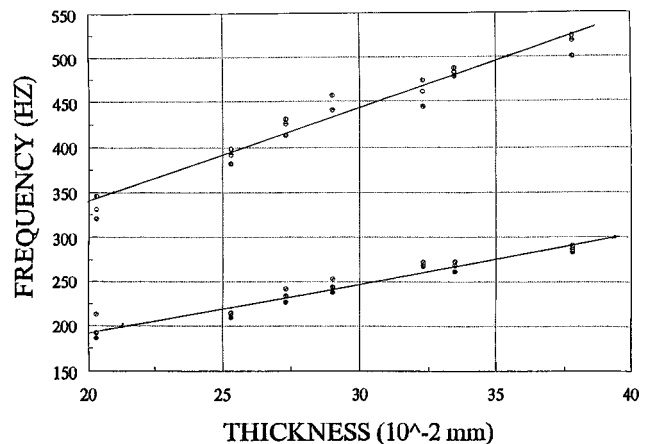


FIG. 7. Frequency versus reed thickness for three different pressures (3, 4, 5 in.) with two lengths of reed. The lower set of data points is for  $L = 2.96$  cm and the upper for  $L = 2.11$  cm.

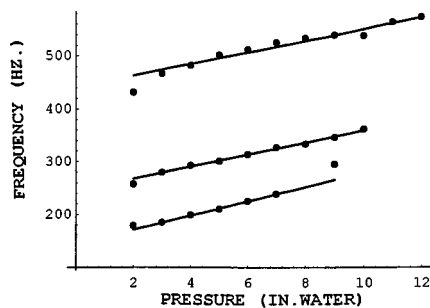


FIG. 8. Frequency versus pressure for the American shallot. Length #1 = 3.50 cm (lowest curve), length #2 = 2.60 cm (middle curve), and length #3 = 1.98 cm (top curve).

used,  $P = 3, 4,$  and  $5$  in., respectively. Since the three different pressures did not change the frequency very much, the three points corresponding to each thickness were averaged to obtain one representative frequency. These data were then fit by least-squares analysis to a linear curve which is superimposed on the data of Fig. 7. The linear graph fit the average data quite well, indicating that the frequency is directly proportional to thickness, as is also true for a freely vibrating cantilevered beam.

In determining the frequency dependence of the reed on varying static pressure, three separate types of shallot, the American, the German, and the French, were used. For each shallot, the same reed was employed, and measurements were made with the tuning wire set at the same three lengths, which were 3.50, 2.60, and 1.98 cm, respectively. For each length, the pressure was varied from the lowest possible pressure at which the reed sounded (about 2 in. of water pressure) to the maximum achievable (about 10 in.). This is shown for the American-style shallot in Fig. 8 for the three lengths of reed. In order to fit a linear curve to this data, it was necessary to delete the lowest and highest pressure points. This is not unreasonable because these points are far removed from the normal operating range of this reed. The area of confinement, 3 to 7 in. of water pressure, is the normal operating range of this organ reed; outside this range, it is not even certain that the reed is functioning in the same mode of operation. Thus, linear fits were used over the normal operating range of the reed, where it appears that the sounding frequency is directly proportional to the static pressure. Figure 9 shows that similar results are obtained for the French and German shallots, both reeds having been set to the same length of 2.60 cm.

For a freely vibrating cantilevered beam, the fundamental frequency of oscillation is inversely proportional to the inverse length squared. The data from this experiment indicate that this is not the case for the driven organ reed, as can be seen in Fig. 10 where frequency versus inverse length is plotted for each of the three shallot types. A linear curve fits the data nicely for these cases (and for a range of other pressures), indicating that the frequency of an organ reed is inversely proportional to the vibrating length for each shallot over the range of lengths and pressures tested.

Combining the above relationships and assuming that for organ reeds the frequency  $F$  is given by

$$F = C^*(T*P)/L, \quad (1)$$

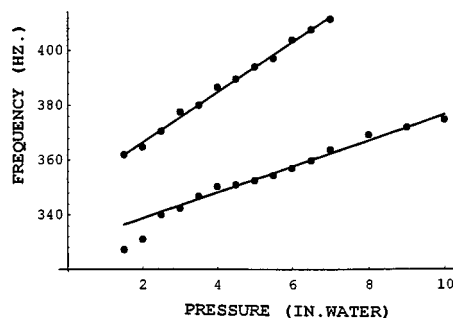


FIG. 9. Frequency versus pressure for the French shallot (upper curve) and the German shallot (lower curve).  $L = 2.60$  cm in both cases.

where  $C$  is a dimensional constant,  $T$  is the thickness of the reed, and  $L$  its length, and using data derived from Figs. 7–10,  $C$  may be determined by any one of the following relationships, written as an empirical equation of the form:

$$C = (F/T)/(P/L) = (F/P)/(T/L) = (F/(1/L))/(T*P). \quad (2)$$

Note that the numerator of each ratio is simply one of the slopes from the linear curve of best fit and that the denominator is a product or quotient of the appropriate constant parameters for the trial from which the corresponding slope was obtained. In order to reduce the possible number of variables (such as shallot shape and design),  $C$  was calculated using only trials involving the American shallot.

Unfortunately, the attempt to obtain a simple empirical equation for the operating frequency as a function of static pressure, thickness of the reed, and vibrating length of the reed was not entirely successful. Using Eq. (2), the values obtained for the “constant”  $C$  varied by as much as a factor of 10. Although shallot variations were avoided, there must be other important variables inadvertently ignored, such as the exact degree of curvature of the reed or the elastic properties of the brass used. An exact equation could not be obtained, but the attempt has led to a number of other interesting conclusions about the vibrating reed. First, the frequency dependence on static pressure is independent of, or at least not strongly affected by, the vibrating length of reed chosen to perform the measurements. This may be easily seen by considering the slopes of the three lines fitting the  $F$  vs  $P$

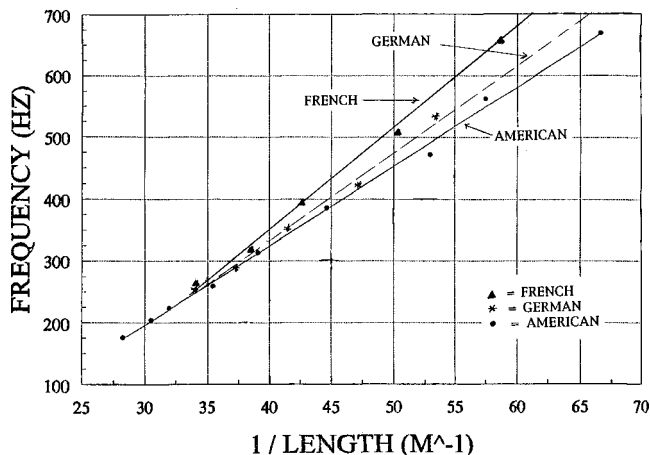


FIG. 10. Frequency versus inverse length; American, French, and German shallots.

data of Fig. 8. The numerical values of the slopes are 12.6, 11.7, and 11.5 Hz/in., respectively. While the pressure dependence of frequency appears relatively unaffected by length, it does seem to be strongly influenced by shallot type. The slopes of the  $F$  vs  $P$  plots for the French and German shallots were 9.03 and 4.64 Hz/in., respectively. This indicates that the type of shallot (which influences the flow dynamics of the air) has a substantial effect on the manner in which pressure influences frequency.

Next, it was shown that the operating frequency of a reed is linearly proportional to the inverse vibrating length for all types of shallots and within the normal operating pressures of the reed. This property has long been known to reed voicers as the rule of thumb that the pitch produced by a reed pipe when the tuning wire is completely extended (against the wedge) is approximately one octave below the desired pitch. To bring the pipe to its correct pitch the tuning wire must be pushed about halfway down the exposed length of reed, thus halving the length of reed doubles the frequency. An examination of the slopes of  $F$  versus  $1/L$  indicates that the nature of this dependence is not strongly affected by the shallot type. Slopes for the three types of shallot varied only by a factor of about 1.3.

The dependence of frequency on thickness was the relationship having the greatest degree of uncertainty, due to the fact that in order to vary thickness different reed tongues had to be inserted. Although the professional reed voicer who curved the reeds had been instructed to voice these to give as identical a sound as possible, achieved by varying the curvature (i.e., thinner reeds had to be curved more so they would operate correctly at the chosen pressure), different curvatures introduce another possible parameter (which will be considered later).

### C. Effect of shallot's interior geometry on spectrum

The shallot's interior geometry was investigated because, although not widely known outside the world of organ reed voicers, it has a profound effect upon the reed's tone (and spectrum). Changing the interior structure by filling in the shallot with solder or hard wax is a trade secret which has long been used by reed voicers to brighten the tone of a rank of pipes. Because changes in the interior geometry change the magnitudes of the forces acting on the reed, both fundamental frequency and the spectrum change.

One reed length on each of the three shallots was chosen for this investigation. Each shallot was investigated by measuring the fundamental frequency and the spectrum as it was progressively filled with wax. The static boot pressure was maintained at 4 in. and for each measurement the length of the reed was adjusted to exactly the same set of lengths as progressively larger portions of wax were added to the interior of the shallot. To insure that the thickness, curvature, and shape were identical, only one reed was used for all measurements in this investigation; its length was varied over five different values, the shortest of which was between 1.5 to 2.5 cm from the reed tip (depending on the shallot type). Reed-length adjustments were made consistent by measuring the longest reed length possible inside the boot, and then measuring its corresponding value on the tuning

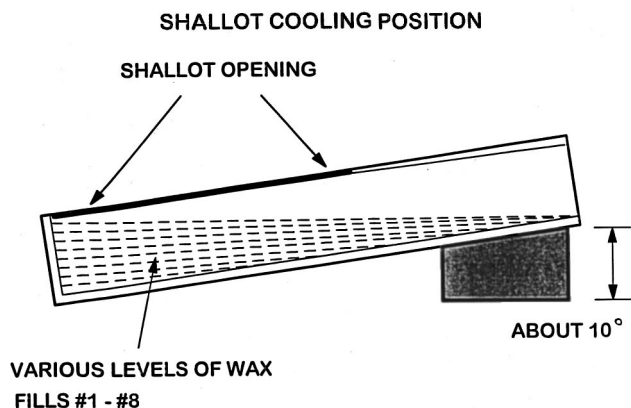


FIG. 11. Wax fills #1 through #8 for shallots. Fill #1 is the least wax added; the amount of wax increases by a constant mass until fill #8 (the greatest amount of wax) is achieved.

wire outside the boot. This enabled the exact determination of the vibrating length of reed inside the boot by measuring the protruding portion of the tuning wire with a Vernier caliper.

For this set of measurements the sound output from the socket was picked up by a microphone and input to the digital storage oscilloscope. The stored data was then Fourier transformed to produce a spectrum. The spectrum was used to compare the effects of changes to the interiors of each of the three types of shallot.

After obtaining spectra for five different reed lengths for each unfilled shallot, the shallots were filled with approximately 0.065 grams of wax (having a density of 0.878 g/cc). After a small ball of wax was inserted into each shallot, the shallots were heated until the wax melted. The shallots were cooled at a slight tilt of  $10^\circ$  so that the wax was concentrated at the base end of the shallot (where the face opening is largest; see Fig. 11). The same five lengths of reed were used to obtain spectra. Then, another 0.065 g of wax was added to each shallot and the process was repeated. This was done for a total of eight fills, ending with about 0.52 g of wax in each shallot. Table I shows the volume changes, due to the addition of the wax, for the American, French, and German shallots, respectively.

Table II presents the fundamental frequencies for five different reed lengths for the unfilled shallot and fills #1 through #8 for each type of shallot. The lengths in a region where the pipe would typically be tuned are 2.86 cm for the German shallot, 2.92 cm for the French shallot, and 2.39 cm for the American shallot. (These lengths, marked with stars in Table II, are those used for the following discussion.) Homer Lewis, a reed voicer with over 50 years of experience, has stated that, for most types of shallot, "If a small amount of wax is added to the shallot, the pitch will go sharp, but if a larger amount of wax is added, so as to restrict the air flow, then the pitch will go flat" (Lewis, 1996a). The present experiments partly confirm this intuition, at least for the American shallot; for the French and German shallots the frequency decreases for all fills. The German shallot exhibits a sudden frequency drop for fill #1; the frequency then increases somewhat to fill #4, after which there is a general decline with successive fills. Table II shows that this trend is typical of all the reed lengths studied for the German shallot.

TABLE I. Shallot volume changes.

Fill	Wax (gm)	V wax (cc)	American		French		German	
			V air (cc)	% Wax	V air (cc)	% Wax	V air (cc)	% Wax
Unfill	0	0	1.58	0	1.22	0	1.36	0
#1	0.065	0.074	1.51	4.7	1.15	6.1	1.29	5.4
#2	0.130	0.148	1.43	9.4	1.07	12.1	1.21	10.9
#3	0.195	0.222	1.36	14.1	0.998	18.2	1.14	16.3
#4	0.260	0.296	1.28	18.7	0.924	24.3	1.06	21.8
#5	0.325	0.370	1.21	23.4	0.850	30.3	0.990	27.2
#6	0.390	0.444	1.14	28.1	0.776	36.4	0.916	32.6
#7	0.455	0.518	1.07	32.8	0.702	42.5	0.842	38.1
#8	0.520	0.592	0.988	37.5	0.628	48.5	0.768	43.5

Figure 12 presents the spectral changes with decreasing shallot volume; by plotting the amplitudes of the first five harmonics for the unfilled case, fill #3, fill #5, and fill #8. This figure clearly demonstrates that higher harmonics increase in amplitude relative to the fundamental as the shallot is filled, thus verifying Mr. Lewis' statement that wax is added to the shallots in order to obtain a brighter tone.

As the French shallot was progressively filled with wax it exhibited the same general trend as the German shallot, except that the frequency dropped with the first fill and remained relatively constant for the next several fills, followed by a frequency decrease. Due to the open characteristic of the shallot, successive fills after the first do not seem to affect the fundamental frequency appreciably. Figure 13 shows that the effect on the spectrum of filling the shallot is, again, to increase the upper harmonics relative to the fundamental. By the eighth fill, however, the shallot passageway is so constricted that all the harmonics are somewhat suppressed, giving a spectral envelope similar to that of the unfilled shallot.

The American shallot behaves somewhat differently from the German and French shallots as wax is added. The

frequency first drops, then progressively increases to a value higher than the original frequency, then drops again and levels off at a frequency approximately that of the original unfilled case. This confirms Lewis' (1996a) prediction, based on his extensive experience with American shallots, that a moderate amount of wax causes the frequency to go sharp (fill #4), while larger amounts will cause a frequency decrease. Figure 14 shows that the amplitude of the fourth harmonic increases relative to the fundamental up through fill #3; further increasing the wax in the shallot progressively lowers the relative amplitude of the fourth harmonic until fill #8, where the relative amplitude of harmonic four is about the same as for the unfilled case.

The observed trends exhibited by these experiments can perhaps be explained by considering the traction forces on the reed. It is hypothesized that as the air flows through the shallot, the "pocket" formed at its base induces viscous losses due to the formation of eddy currents, reducing the energy available to the higher harmonics. As this pocket is progressively filled with wax, the flow is guided more smoothly through the channel, there are fewer viscous losses, and more energy is distributed to the higher harmonics, resulting in a brighter tone. The greater the boot pressure relative to the shallot pressure, the more rapidly the air will travel through the shallot opening, which increases this effect and causes the reed to vibrate more rapidly. With the first addition of wax the boot pressure, when the reed is at its maximum extension above the shallot, decreases because of the increased air flow through the shallot. The lower pressure

TABLE II. Comparison of results for the filled shallots. Fills consist of approximately 0.065 g of wax added each time to the unfilled shallot. (\* indicates lengths discussed.)

Reed length (cm)	Fill number—fundamental frequency (Hz)								
	Unfilled	#1	#2	#3	#4	#5	#6	#7	#8
<b>American shallot</b>									
L = 3.19	219	180	200	179	179	169	180	169	179
L = 2.79	339	258	284	269	309	280	279	269	289
*L = 2.39	419	339	399	419	458	409	423	419	449
L = 1.99	558	457	578	588	678	608	607	618	653
L = 1.59	757	696	817	787	897	787	849	877	877
<b>French shallot</b>									
L = 3.52	200	159	159	150	155	159	159	159	129
L = 3.22	229	190	189	189	170	190	180	169	179
*L = 2.92	299	259	260	260	240	269	250	249	229
L = 2.62	359	319	329	418	318	349	329	349	329
L = 2.32	419	397	389	518	428	468	468	488	489
<b>German shallot</b>									
L = 3.76	189	150	159	139	150	159	150	149	149
L = 3.46	199	169	170	160	160	160	159	159	150
L = 3.16	259	210	229	200	195	189	179	219	179
*L = 2.86	299	275	289	289	290	279	269	269	219
L = 2.56	359	336	357	403	406	379	369	398	349

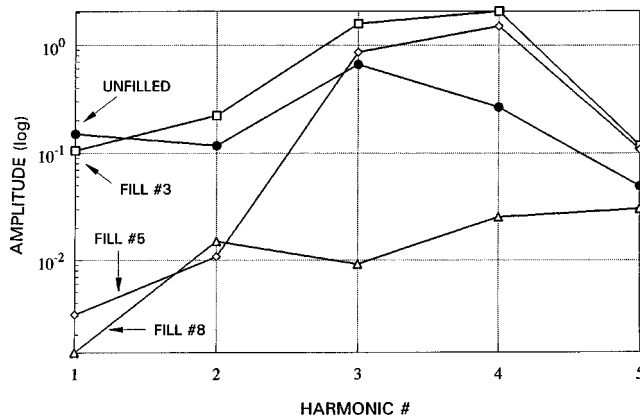


FIG. 12. Harmonic amplitude changes for different fills, German shallot.

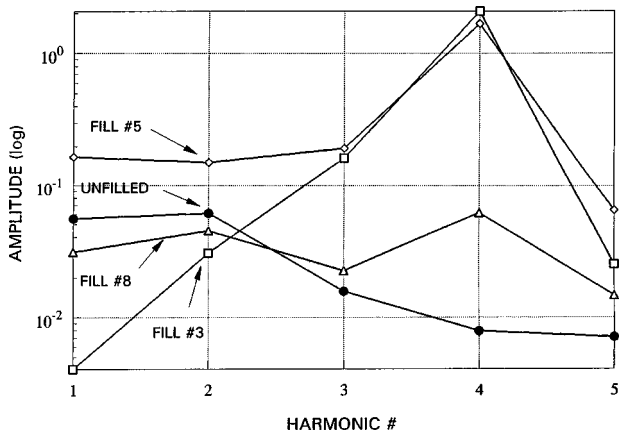


FIG. 13. Harmonic amplitude changes for different fills, French shallot.

in the boot means there is less force on the reed, causing the contact point of the reed on the shallot to move toward the tuning wire, thus lowering the vibration frequency. As the wax progressively fills the shallot base, reduced viscous losses allow a smoother and more rapid air flow which reduces the shallot pressure, thus effectively increasing the force on the reed which results in progressively higher frequencies of vibration. As more wax is added, the effect continues until some maximum point is reached (before the shallot is half filled with wax). Further filling of the shallot constricts the air flow, which causes the frequency to decrease, with an attendant reduction in the intensity of the sound. The larger base of the American shallot (0.3 cm larger where the wax is added) suggests that these effects will be less for the American than for the French or German shallots.

#### D. Effect of the resonator

To investigate the effect of the resonator on reed vibration, seven different resonators were successively attached to the socket. The shallot chosen was the American; the reed was the same one used previously. The reed pipe was tuned to middle C (263 Hz) and the output from each resonator was recorded at the open end. The same procedure was also used to record the waveform from each resonator after the shallot had been filled twice with wax (only amounts of wax corre-

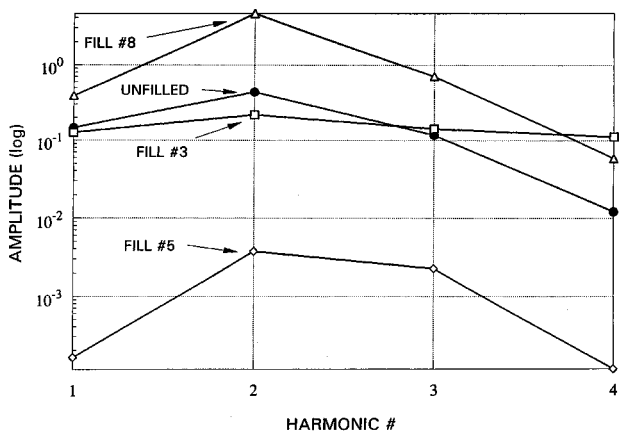


FIG. 14. Harmonic amplitude changes for different fills, American shallot.

TABLE III. Comparison of results for the resonators.

American shallot Reed length	Fundamental frequency (Hz)		
	$L = 2.64$ cm Unfilled	$L = 2.79$ cm Fill #2—0.13 g	$L = 2.66$ cm #8—0.52 g
Apparatus (middle C—263 Hz)	259	259	269
Resonator #1—trompette $L = 30.5$ cm	259	239	249
Resonator #2—oboe $L = 49.5$ cm	239	259	249
Resonator #4—brass trompette $L = 41.5$ cm	258	250	249
Resonator #5—orchestral oboe $L = 50.0$ cm	219	230	229
Resonator #7—holz regal $L = 18.0$ cm	258	259	259
Resonator #8—vox humana $L = 21.5$ cm	258	269	259
Resonator #11—fagot $L = 46.3$ cm	239	249	249

sponding to fill #2 and fill #8 were applied). The resonators used were an assortment of typical resonators for middle C found on pipe organ reeds.

The results of this experiment are summarized in two tables. Table III lists the fundamental frequency for three different shallot fills (unfilled, the #2 typical fill, and the extreme fill #8) of the American shallot without resonator (row 1) as well as the respective fundamentals when seven different resonators are attached. Table IV tabulates the relative log amplitudes of the first seven harmonics for each fill combined with each resonator. As the resonators were varied, the shallot fills and reed lengths remain unchanged. Adding a resonator to the boot and socket has two effects on the resulting spectrum. First, the signal-to-noise ratio is substantially increased, yielding well-defined harmonics with relatively high  $Q$  as compared to the spectra obtained when no resonator is attached. Second, the fundamental frequency is decreased by the resonator. This effect, which can be easily understood as adding an effective acoustic mass to the system, is well-known to reed voicers (Lewis, 1996b). An examination of the spectra show that the fourth harmonic (about 1000 Hz) is almost always the largest in amplitude. This can, perhaps, be explained by the resonance of the boot. It has long been recognized by reed voicers (Lewis, 1996b) that boot resonance can, especially for short-length resonators, cause problems with tonal stability. If the boot is considered as a closed-closed tube, its resonant frequency would be about 900 Hz, in the same general region of the fourth harmonic of middle C.

Each resonator affects the spectrum differently, depending on the resonator's length, shape, and the nature of the terminal end (open, partially capped, or closed with a small opening). In order to reinforce the reed vibration, the resonator must be either a "full length" (i.e., equal to one-half the desired wavelength for conical type resonators, or one-



TABLE IV. Log amplitudes of harmonics for various fills and resonators.

Har.#	boot	res #1	res #2	res #4	res #5	res #7	res #8	res #11
Unfilled shallot								
1	6.40	2.20	6.80	2.70	3.40	6.20	6.80	3.50
2	4.40	4.95	6.20	3.90	3.10	6.55	6.60	4.05
3	7.20	4.15	6.90	5.00	3.70	6.30	6.00	5.30
4	7.45	6.65	9.20	5.90	5.10	7.00	8.20	6.50
5	5.90	4.70	7.40	4.70	4.10	6.50	6.00	5.25
6	6.00	3.20	4.85	3.40	2.60	4.90	4.85	3.70
7	5.25	1.60	4.00	3.40	1.80	4.90	4.30	...
Fill #2								
1	6.40	4.45	5.10	3.30	4.35	5.30	6.10	5.90
2	4.80	5.15	5.20	4.00	4.15	5.35	5.90	5.55
3	5.90	4.60	6.45	5.60	4.60	5.25	5.40	6.85
4	7.20	6.30	7.30	6.30	6.55	5.90	7.85	7.80
5	5.90	5.30	6.00	4.90	6.55	6.10	7.00	6.70
6	6.30	3.45	5.85	4.10	3.20	3.35	6.25	6.15
7	4.80	2.80	3.70	3.75	3.20	3.80	3.20	4.70
Fill #8								
1	5.15	2.05	4.35	2.50	3.00	5.70	5.90	5.60
2	4.30	4.45	4.05	3.40	4.05	6.10	7.25	5.10
3	6.15	4.40	4.80	5.20	4.85	5.55	6.05	6.30
4	7.10	5.60	6.20	5.90	6.60	6.95	7.85	7.00
5	5.90	4.85	5.00	4.90	6.30	6.90	8.00	5.40
6	5.90	3.20	4.65	3.70	4.50	5.35	6.30	5.55
7	5.60	2.70	3.70	3.33	3.20	4.80	4.10	4.10

quarter the wavelength for cylindrical) or a “half length” (one-quarter the desired wavelength for conical type resonators). Many other types and resonators have been employed for regals and special effect reeds; these resonators are typically a relatively small fraction of the wavelength (1/8 or 1/16 are not uncommon). For the full or half-length resonator, its length is adjusted, depending on geometry, to match its first and/or second resonance to the desired frequency. The very short resonators act more like a high-pass filter, reinforcing the higher frequencies and attenuating the lower frequencies; feedback from the resonator has little effect upon the reed for these short resonators (Strong and Plitnik, 1992, p. 327). Most of the resonators tested in this portion of the experiment were designed to be used on a pipe sounding middle C (263 Hz).

The trompette resonator [an inverted cone as in Fig. 3(b)] was tested with the American shallot in three conditions (unfilled, fill #2, and fill #8). As already discussed, filling in the shallot lowered the frequency and increased the upper harmonics of the spectrum. For this configuration, fill #2 seemed to be optimum; it brought the lower and upper partials more in line with the strong fourth harmonic, thus giving a much more uniform spectral envelope and increasing the  $Q$  of most harmonics. Fill #8, being somewhat excessive, i.e., more than would typically be used for this stop, caused a deterioration of the uniform spectrum by reducing the upper and lower partials relative to the fourth, although their  $Q$  values were sharpened somewhat.

The effect on the spectra for resonator #2, a typical oboe [double cone as in Fig. 3(a)] indicates that, in a similar manner, progressively filling the shallot boosts the upper partials, evens out the spectral envelope, and increases the  $Q$ .

The effect on resonator #4, a brass trumpet (a large cone

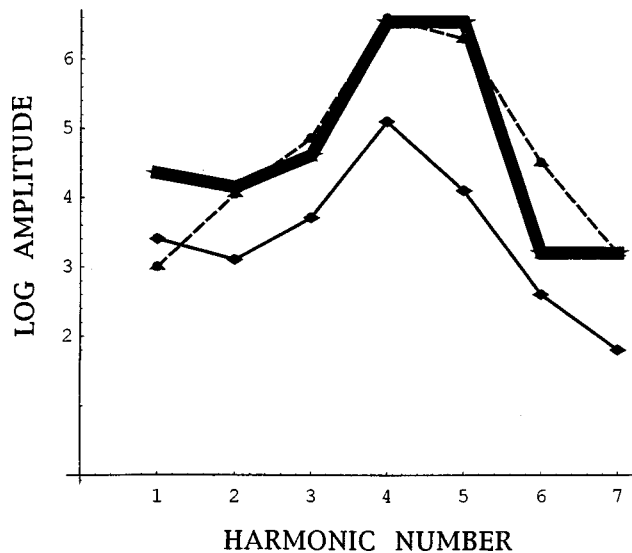


FIG. 15. Changes in the spectral envelope for the orchestral oboe with three different shallot fills. The solid line is the unfilled case, the heavy line is for fill #2, and the dashed line is for fill #8.

with an exponentially flared bell), of filling the shallots is that there are only small changes in the fundamental frequency, but considerable change in the spectral envelopes. The lower harmonics are attenuated and the  $Q$  becomes higher as the volume of the shallot decreases. The optimum fill seems to be fill #2; after this the fundamental becomes increasing reduced until it is almost missing for fill #8.

The effect of resonator #5, orchestral oboe [a narrow capped cone as in Fig. 3(h)], is for the filled shallots to lower the vibration frequency, to increase the  $Q$ s, attenuate the lower partials, and increase the upper. This resonator probably has more effect on the spectrum because it is quite narrow, the only opening being a small slot near the capped top. Adding the resonator decreased the frequency by about 40 Hz, independent of the amount of shallot fill. Figure 15 shows the spectral envelopes for this resonator with three different shallot fills. It can be seen that fill #2 smoothes the spectrum, raises the amplitude of all higher partials, but with a greater relative increase to the higher frequencies. The overall higher amplitude created with fill #2 is most likely due to the reduced shallot interior volume being a better match to the small resonator employed for this pipe. Typically, a smaller diameter shallot would be used. Figure 15 also shows that there is almost no difference between the spectral envelope of fill #2 and fill #8, probably because all flow constrictions in the shallot produce a fairly good match to the small diameter resonator.

The effect of resonator #7, holz regal (a fractional length wooden resonator of square cross section), was that adding the resonator did not lower the vibration frequency, except for fill #8. For this resonator, the sharpest  $Q$ s occurred for the unfilled shallot; additional filling did not much affect the spectral envelope, but it did reduce the  $Q$  values. Although this resonator maintains a constant cross-sectional area over most of its length, all harmonics are present in the spectrum, indicating that the resonator is not controlling the reed so much as acting as a filter with minimal feedback. This is also

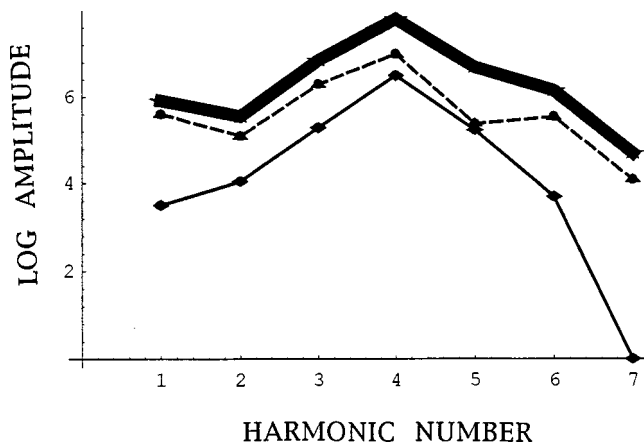


FIG. 16. Changes in the spectral envelope for the fagotto with three different shallot fills. The solid line is the unfilled case, the heavy line is for fill #2, and the dashed line is for fill #8.

supported by the observation that adding the resonator did not particularly affect the vibration frequency for most of the filled shallots.

The effect of resonator #8, vox humana [a fractional length capped cylinder as shown in Fig. 3(i)], is to increase the amplitude of the upper harmonics and to increase the  $Q$ s as the shallot is progressively filled. The fundamental frequency does not change when the resonator is added (except for fill #8), indicating that this fractional length resonator acts as a filter which exerts very little control on the vibrating reed.

The effect on resonator #11, fagotto [a narrow double cone as in Fig. 3(a)] is exactly that expected for a full-length, conical-type resonator. Adding the resonator lowers the frequency of the fundamental, while progressively filling the shallot sharpens the  $Q$  values. In this case, the optimum fill seems to be fill #2 since it has the most uniform spectral envelope, the best developed harmonics, and harmonics with the greatest amplitude, as can be seen by inspecting Fig. 16. Fill #8 gives a fairly uniform, but somewhat attenuated, spectral envelope. The tone produced with this resonator is quite bright due to the prevalence of strong, well-developed upper harmonics.

#### E. Effect of reed curvature on the spectrum

To determine the effect of reed curvature, a typical reed tongue was fitted to the German shallot. The experiment began with the reed having a low initial curvature. After each successive measurement (to obtain a spectrum) a burnishing tool was used to progressively increase the reed's curvature. For purposes of this experiment, the reed curvature was defined as the height of the reed tip above the shallot, as measured by a micrometer.

The results of this experiment are presented in Table V. A professional reed voicer informed us that if the curvature is increased near the tuning wire the pitch will decrease, while increasing the curvature near the tip will increase the pitch (Lewis, 1996b). For this experiment the reed was curved over its entire length; the results, shown in Fig. 17, confirm the voicer's prediction. The pitch decreases for the first two increases of curvature, then remains constant; in-

TABLE V. Analysis of reed curvatures. German shallot—reed length 2.5 cm.

Curvature number	Amount of curvature (mm)	Frequency of fundamental (Hz)
#1	1.75	389
#2	2.54	369
#3	2.82	319
#4	3.11	319
#5	3.73	318

creasing the curvature no longer lowers the frequency. This can be explained by considering that as the overall curvature is increased, the effective vibrating length is also increased, i.e., the point of contact between the reed and the shallot moves closer to the tuning wire when the air enters the boot. Likewise, as the curvature is reduced, the point of contact moves away from the tuning wire, thus increasing the vibration frequency. When a large portion of the reed is in contact with the shallot (low curvature), the pressure in the boot will prevent the section of the reed touching the shallot from vibrating, even though the tuning wire does not prevent it.

A summary of the change in spectra as the curvature is increased is shown in the spectral envelope of Fig. 18; changing the curvature seems to have only a small effect on the spectrum; the first two harmonics become progressively more attenuated relative to the upper harmonics which are only slightly affected. This agrees with the general consensus that a greater curvature, as would be used for a thinner reed, tends to produce a tone having greater brightness (see Williams and Owen, 1988)

#### IV. CONCLUSIONS

Although the vibration of organ reeds seems to be a much more complicated phenomenon than one may have been led to assume, all overtones were integer multiples of the fundamental and no nonharmonic partials were observed.

Concerning the phase difference on opposite sides of the reed, there are two main conclusions. First, there is an approximately  $90^\circ$  phase shift in pressure on opposite sides of the reed, and second, the acceleration of the reed tongue lags

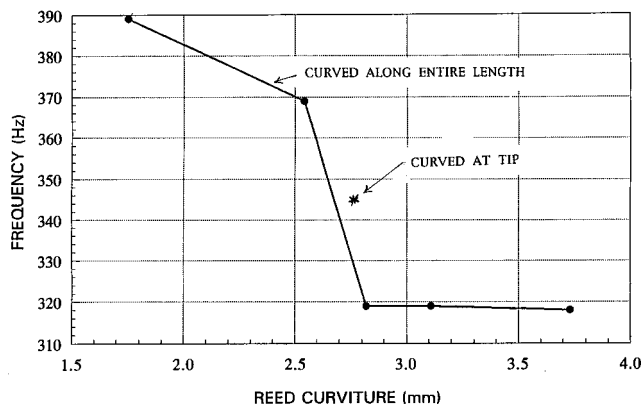


FIG. 17. The effect of reed curvature on fundamental frequency. As the curvature increases (as measured at the tip) the frequency decreases.

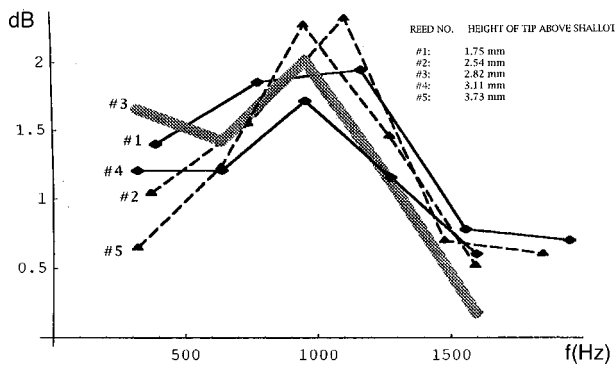


FIG. 18. Changes in spectral envelope due to different reed curvatures.

the boot pressure by approximately  $90^\circ$  when the system is vibrating at the higher frequencies induced by the higher boot pressures.

Fourier analyses of the signals indicated that each signal was composed of one strong fundamental and a number of higher partials. A consideration of the accelerometer signals from different vibrating lengths of reed also indicates that when the length is shorter (higher frequency) the vibration is considerably simpler (i.e., less energy is distributed to the overtones) than for the lower frequencies obtained when the vibrating length of reed is longer. It is hypothesized that this results from the more intense beating of the reed against the shallot, which “clips” the waveform to produce more overtones of greater amplitude.

Although the attempt to obtain a simple empirical equation for the operating frequency as a function of static pressure, thickness of the reed, and vibrating length of the reed was not entirely successful, the attempt to find it has revealed several interesting, if not unexpected, relationships. First, the frequency dependence on static pressure is essentially independent of the vibrating length of reed, but strongly influenced by shallot type. Next, it was shown that the operating frequency of a reed is linearly proportional to the inverse vibrating length for all types of shallots and within the normal operating pressures of the reed. Then, it was shown that the frequency of a sounding reed is linearly proportional to the inverse vibrating length for all types of shallots operating within the normal pressure range of the reed. Furthermore, this dependence is not strongly affected by the shallot type. Finally, it was established that frequency is linearly dependent on reed thickness, even though the varying tongue curvature necessitated by this experiment lends a degree of uncertainty to the conclusion.

Concerning the filled shallots, there are three main conclusions: (1) filling the shallot with wax causes the frequency to decrease, (2) as more wax is added the frequency increases to values approaching the “unfilled” case, and (3) wax tends to brighten the tone by raising the amplitudes of the higher harmonics relative to the fundamental.

Several conclusions can be drawn from the investigation of the resonators. First, the addition of a resonator almost always lowers the fundamental frequency, increases the spectral amplitudes, raises their  $Q$  value, and increases the signal to noise ratio of the spectrum. Second, the amplitude of the fourth harmonic of the spectrum is generally the larg-

est in magnitude, perhaps due to the resonance of the boot reinforcing its amplitude. Third, all the frequencies of vibration decrease, relative to the resonatorless case, when fill #8 (the maximum amount of wax) is applied to the American shallot. This may be attributed to the increased effect of the returning air pulses “focused” by the decreasing cross-sectional area of the resonator where it enters the block. Fourth, the longer the resonator, the larger the frequency drop as compared to the waveform produced by the reed with no resonator in place. Finally, the fractional length resonators (holz regal and vox humana) have little effect on the reed frequency, but the holtz regal had a profound effect on increasing the  $Q$  of the harmonics as well as the signal-to-noise ratio.

The preliminary conclusions drawn from the analysis of changes to the reed curvature were that as curvature increases, the frequency decreases, up to a point; beyond this there is no further change. Changes to the spectrum are minimal for reeds curved over their entire length, but according to a professional reed voicer (Clipp, 1998), additional curvature applied near the tip of a reed tongue brightens the tone considerably. The all-important aspect of reed curvature is now being investigated in considerably more detail; the results will be reported in a future paper.

## ACKNOWLEDGMENTS

The experimental work of my students, Ron Knox and Jeff Ritchie, constitute the foundation upon which this work was constructed. A debt of gratitude is owed to the master reed voicers of Trivo, Inc. (Hagerstown, MD), Homer Lewis and Joe Clipp, for their willingness and eagerness in unselfishly donating their time to discuss reed voicing and to provide reed parts as needed. Finally, the important contributions of Dr. Chandra Thamire, Dr. Paul LaChance, and Dr. William J. Strong who read through several proofs of this paper and offered invaluable suggestions, are gratefully acknowledged.

<sup>1</sup>The single known exception is the Zajic regal, developed by the late Adolph Zajic who was a master reed voicer at the M. P. Moller Pipe Organ Co. of Hagerstown, MD. This regal had absolutely no resonator; the opening for sound egress was the top of the block.

<sup>2</sup>Pressure is measured in inches of water as this has always been, and continues to be, the standard in the American pipe organ industry.

Allihn, M. (1888). *Die Theorie und Praxis des Orgelbaues* (B. F. Voigt, Weimar).

Andersen, P. G. (1969). *Organ Building and Design* (Oxford University Press, New York), pp. 55–56.

Angster, J., Angster, J., and Miklos, A. (1997). “Akustische Messungen und Untersuchungen an Orgelpfeifen,” *Acta Organologica*, Band 25, Kassel (Merseburger).

Audsley, G. A. (1905). *The Art of Organ-Building, Vols. I and II* (Dover, New York), 1965 reprint of the 1905 edition.

Audsley, G. A. (1919). *The Organ of the Twentieth Century* (Dover, New York), 1970 reprint of 1919 edition, pp. 453–454.

Bonavia-Hunt, N. A. (1933). *Modern Studies in Organ Tone* (Office of Musical Opinion, London), p. 91.

Bonavia-Hunt, N., and Homer, H. W. (1950). *The Organ Reed* (J. Fisher, Glen Rock, NJ).

Bouasse, H. (1929). *Instruments a Vent* (Librairie Delagrave, Paris), Vol. I, pp 68–70, 103–117.

Clipp, Joe. (1998). Personal interview, Trivo Organ Reed Co., Hagerstown, MD, 28 June 1998.

- Ellerhorst, W. (1966). *Handbuch der Orgelkunde* (Frits Knuf, Hilversum), pp. 282–323.
- Helmholtz, H. (1954). *On the Sensations of Tone* (Dover, New York), reprint of the 1885 edition.
- Hirschberg, A., van de Laar, R., Marrou-Maurieres, J., Wijnands, A., Dane, H., Kruijswijk, S., and Houtsma, A. (1990). “A quasi-stationary model of air flow in the reed channel of single-reed woodwind instruments,” *Acustica* **70**, 146–154.
- Jann, G., and Rensch, R. (1973). “Experiments with Measurements in Reed Pipes,” ISO Information #9 (Feb. 1973), pp. 633–646.
- Lewis, Homer (1996a). Personal interview, Trivo Organ Reed Co., Hagerstown, MD, 9 April 1996.
- Lewis, Homer (1996b). Personal interview, Trivo Organ Reed Co., Hagerstown, MD, 23 April 1996.
- Lottermoser, W. (1983). “Akustik der Zungenpfeifen,” in *Die akustischen Grundlagen der Orgel*, Vol. I of *Orgeln, Kirchen and Akustik* (Bochinsky Frankfurt am Main).
- Smith, Hermann (1911). *The Making of Sound in the Organ and in the Orchestra* (Charles Scribner’s Sons, New York), pp. 230–245, 249–250.
- Strong, W., and Plitnik, G. (1992). *Music, Speech, Audio* (Soundprint, Provo, UT).
- Topfer, J. G. (1833). *Die Orgelbaukunst nach Einer Neuen Theorie* (Hoffman, Weimar) (edited and revised by Paul Smets, Rheingold, Mainz, 1939).
- Tyndall, J. (1875). *The Science of Sound* (Citadel, New York), 1964 reprint of the 1875 edition.
- Williams, P., and Owen, B. (1988). *The Organ* (W. W. Norton, New York), p. 36.

# Anisotropy of ultrasonic backscatter and attenuation from human calcaneus: Implications for relative roles of absorption and scattering in determining attenuation

Keith A. Wear<sup>a)</sup>

U. S. Food and Drug Administration, Center for Devices and Radiological Health, HFZ-142,  
12720 Twinbrook Parkway, Rockville, Maryland 20852

(Received 8 December 1999; revised 9 February 2000; accepted 2 March 2000)

Although bone sonometry has been demonstrated to be useful in the diagnosis of osteoporosis, much remains to be learned about the processes governing the interactions between ultrasound and bone. In order to investigate these processes, ultrasonic attenuation and backscatter in two orientations were measured in 43 human calcaneal specimens *in vitro* at 500 kHz. In the mediolateral (ML) orientation, the ultrasound propagation direction is approximately perpendicular to the trabecular axes. In the anteroposterior (AP) orientation, a wide range of angles between the ultrasound propagation direction and trabecular axes is encountered. Average attenuation slope was 18% greater while average backscatter coefficient was 50% lower in the AP orientation compared with the ML orientation. Backscatter coefficient in both orientations approximately conformed to a cubic dependence on frequency, consistent with a previously reported model. These results support the idea that absorption is a greater component of attenuation than scattering in human calcaneal trabecular bone. © 2000 Acoustical Society of America. [S0001-4966(00)03806-6]

PACS numbers: 43.80.Cs, 43.80.Jz [FD]

## INTRODUCTION

Bone sonometry has been established as a useful modality for prediction of osteoporotic fracture risk.<sup>1-12</sup> Most commercial bone sonometers measure attenuation and/or sound speed in the calcaneus. Despite the documented clinical utility, the physical mechanisms underlying the interaction between ultrasound and bone are not completely understood yet.

The calcaneus is predominantly composed of trabecular bone and is surrounded by a thin cortical shell. The trabecular interior consists of a three-dimensional lattice of branching spicules and platelike structures. The spaces between the trabeculae are filled with marrow which consists of fat and cellular components of blood constituents. The interfaces between mineralized bone trabeculae and marrow (which have substantially different acoustic impedances) are likely candidates for the sources of ultrasonic scattering.

In a previous study conducted in this laboratory, backscattering from calcaneus under mediolateral (ML) insonification (which is the orientation employed by commercial calcaneus-based bone sonometers) was found to be consistent with a proposed theory in which calcaneal trabeculae were modeled as long thin cylinders that are narrow relative to the wavelength and long relative to the beam cross section.<sup>13</sup> Scattering by a cylinder of an incident plane wave propagating perpendicular to the cylinder axis has been studied in detail.<sup>14,15</sup> (The problem of a plane wave propagating in a direction not perpendicular to the cylinder axis is far more complicated and has received less attention.) In the low-frequency limit, it may be shown that the intensity of an inelastically scattered wave is given by<sup>14</sup>

$$I_s \cong \frac{\pi \omega^3 b^4}{8rc^3} I(1 - 2 \cos \phi)^2, \quad (1)$$

where  $I_s$  is the intensity of the scattered wave,  $I$  is the intensity of the incident plane wave,  $b$  is the radius of the cylinder,  $\omega = 2\pi f$  where  $f$  is the frequency of the wave,  $c$  is the speed of sound in the medium outside the cylinder,  $r$  is the distance from the scatterer to the observation point, and  $\phi$  is the angle between the incident plane wave propagation direction and the observation direction. Thus, at low frequencies, inelastic scattering is proportional to the cube of the ultrasonic frequency at all angles. This is also at least approximately true for the elastic case if the material properties of hydroxyapatite are used to characterize the cylinder.<sup>13,15</sup> (Submicroscopic deposits of calcium phosphate similar to hydroxyapatite are the major inorganic constituents of bone.) The cubic frequency dependence of scattering also would hold for an ensemble of unresolvable cylinders provided that the cylinders (trabeculae) are positioned sufficiently irregularly that their scattered waveforms may be assumed to add incoherently,<sup>13</sup> which would appear to be the case for calcaneus (see Figs. 1 and 2). The cubic dependence of scattering in the ML orientation, in combination with the widely reported finding that attenuation in calcaneus varies approximately as frequency to the first power,<sup>1-12</sup> provides evidence that in this orientation, absorption is a greater component of attenuation than scattering. Some diagnostic promise for backscattering measurements has been demonstrated.<sup>16-19</sup>

Further insight into the nature of the interaction between tissue and ultrasound may be obtained by investigation of anisotropy of acoustic properties, as has been done in heart<sup>20,21</sup> and kidney.<sup>22</sup> Glüer and co-workers investigated anisotropy of broadband ultrasonic attenuation in ten cubical specimens of purely trabecular bovine bone.<sup>23</sup> They found

<sup>a)</sup>Electronic mail: kaw@cdrh.fda.gov

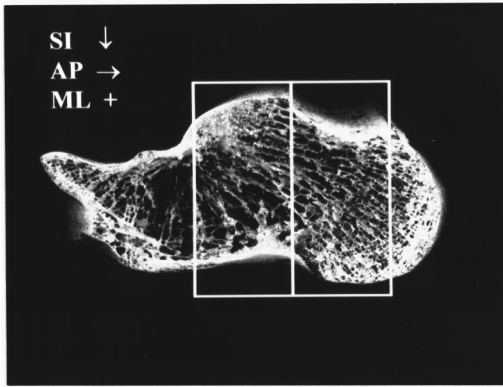


FIG. 1. Calcaneus with lateral cortical sides removed. Two adjacent boxes corresponding to the sections cut from each calcaneus for ultrasonic interrogation are shown. Long thin trabeculae running throughout the calcaneus are apparent. For mediolateral (ML) insonification, the trabecular orientations are approximately perpendicular to the ultrasound propagation direction (perpendicular to the figure). The width of the image is 10.5 cm. Directions of superoinferior (SI) and anteroposterior (AP) propagation are also indicated.

attenuation to be 44%–54% larger along the trabecular axis as compared with the two perpendicular axes.

The objective of this paper is to investigate anisotropy of both attenuation and backscatter in human calcaneus *in vitro*. The parameters are measured for mediolateral (ML) and anteroposterior (AP) orientations. Experimental methods are first described, followed by results. Finally, these results are discussed in the closing section.

## I. EXPERIMENTAL METHODS

### A. Biological methods

Twenty-two human (gender and age unknown) calcaneus samples were obtained from cadavers. They were defatted using a trichloro-ethylene solution. Defatting was presumed not to significantly affect measurements since attenuation<sup>6,24</sup> and speed of sound<sup>24,25</sup> of defatted trabecular bone have been measured to be only slightly different from their counterparts with marrow left intact. In addition, defatting may not have a profound effect on backscatter measurements provided that the acoustic properties of water (which fills spaces between trabeculae during the *in vitro* interroga-

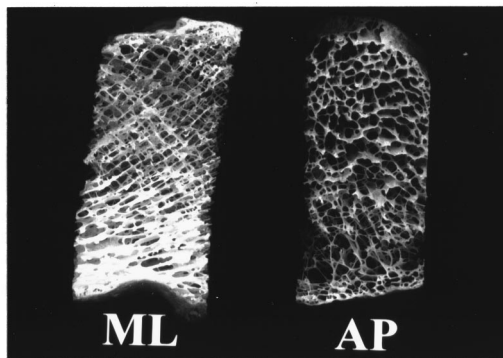


FIG. 2. Two calcaneal sections. The one on the left is shown in the ML orientation. The one on the right is shown in the AP orientation. The width of each section is approximately 15 mm. The height is approximately 35 mm.

tion of defatted bone) and marrow (which normally fills the spaces between trabeculae) are much closer to each other than either is compared with the acoustic properties of mineralized bone trabeculae. The cortical lateral sides were sliced off leaving two parallel surfaces with direct access to trabecular bone. Two portions were cut from each calcaneus sample, one from the central region and one from the posterior region, as shown in Fig. 1. This yielded a total of 43 (22 calcaneae  $\times$  2 samples/calcaneum minus 1 sample which fragmented in the machining process and was unusable) specimens each approximately 15 mm (ML)  $\times$  15 mm (AP)  $\times$  35 mm (craniocaudal) in size. Samples were interrogated in both the ML and AP orientations. Although the trabeculae exhibit a wide range of orientations, they are all oriented approximately perpendicular to the ultrasound propagation direction under conditions of ML insonification (perpendicular to the plane of Fig. 1). In the AP view, on the other hand, a variety of trabecular orientations ranging from perpendicular to parallel is encountered by the ultrasound beam. See Figs. 1 and 2.

In order to remove air bubbles, the samples were vacuum degassed underwater in a desiccator. After vacuum, samples were allowed to thermally equilibrate to room temperature prior to ultrasonic interrogation. Ultrasonic measurements were performed in distilled water at room temperature. The temperature for the experiment ranged from 21.9 °C to 23.4 °C.

Apparent density, the ratio of the dehydrated, defatted tissue mass to the total specimen volume,<sup>6</sup> was measured for each sample. Mass was measured using a balance. Volume was assessed from separate measurements of thickness in the three dimensions using calipers.

### B. Ultrasonic methods

For some experiments, a Panametrics (Waltham, MA) 5800 pulser/receiver was used. For the remaining experiments, a MetroTek MP 215 pulser and a MetroTek MR 101 receiver were used. Samples were interrogated in a water tank using Panametrics 1.00-in.-diam circular, focused (focal length=1.5 in.), broadband transducers with nominal center frequency of 500 kHz. Received ultrasound signals were digitized (8 bit, 10 MHz) using a LeCroy (Chestnut Ridge, NY) 9310C Dual 400 MHz oscilloscope and stored on computer (via GPIB) for off-line analysis.

Backscatter coefficients were measured using a reference phantom method.<sup>26,27</sup> With this method, the dependences of measurements on machine-dependent factors (e.g., transducer aperture, distance from transducer to sample, transducer electromechanical response, gain settings, etc.) are suppressed. A reference phantom, consisting of glass beads in agar, with known frequency-dependent backscatter coefficient and attenuation coefficient, placed in the water tank at the same distance as for the bone samples was used. The backscatter coefficient versus frequency data,  $\eta(f)$ , were least-squares fit to a power law,  $\eta(f) = Af^n$  over the range from 400 to 800 kHz. Both the midband value,  $\eta(500 \text{ kHz})$ , and the exponent,  $n$ , were used to characterize backscatter.

The broadband beam may be thought of as a superposi-

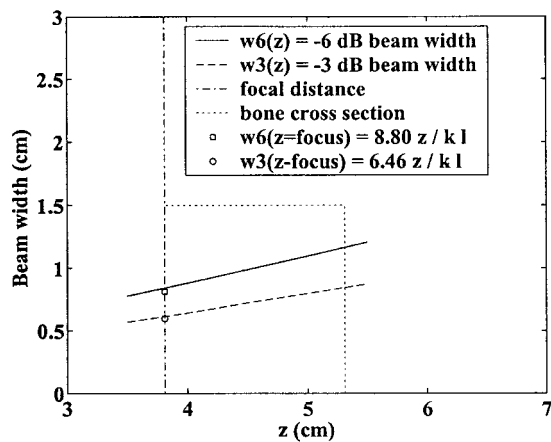


FIG. 3. Beam widths ( $-3$  and  $-6$  dB) as functions of depth for a 400-kHz ultrasound beam emanating from circular piston transducer with a diameter of 25.4 mm and a focal length of 38.1 mm. The  $-6$ -dB beam width is completely contained within the calcaneal section ( $15 \times 15 \text{ mm}^2$ ). The  $-3$ - and  $-6$ -dB beam widths at the focal plane predicted by Fourier optics are also shown.

tion of components at the various contributing frequencies. The lower limit of the range of frequencies used for analysis was set by the minimum frequency for which the beam cross section at that frequency was such that the bulk of the acoustic energy (that portion contained within the  $-6$ -dB points of the beam cross section) could be reliably assumed to be included within the bone specimen. This criterion is somewhat conservative in that it does not make use of the fact that the directivity pattern of the receiving transducer will also serve to suppress signals that are sufficiently laterally displaced from the axis of the beam. The frequency-dependent beam profile was established using numerical simulations employing the Huygens–Fresnel principle without making the Fresnel approximation (which can break down as source dimensions become appreciable relative to the propagation distance).<sup>28</sup> The simulated beam properties at 400 kHz are illustrated in Fig. 3. Since the beam width decreases with frequency (beam width at the focal plane is proportional to  $cz/lf$  where  $c$  is the sound speed,  $f$  is the frequency,  $z$  is the focal distance, and  $l$  is the transducer diameter<sup>28</sup>), frequencies above 400 kHz were considered acceptable as far as diffraction considerations were concerned. The upper limit of the usable band of frequencies was determined by signal-to-noise ratio considerations. See Fig. 4.

Attenuation was measured using a standard substitution method. Using two opposing coaxially aligned transducers (one transmitter and one receiver), transmitted signals were recorded both with and without the bone sample in the acoustic path. Each power spectrum was estimated from the average squared magnitude of fast Fourier transforms of four digitized radio frequency signals. The attenuation coefficient was then estimated using a log spectral difference technique.<sup>29</sup> If the power spectra recorded with and without the bone sample in the acoustic path are denoted by  $Y(f)$  and  $X(f)$ , respectively, then

$$Y(f) = T^4 e^{-2d\alpha(f)} X(f), \quad (2)$$

where  $T$  is the amplitude transmission coefficient at each

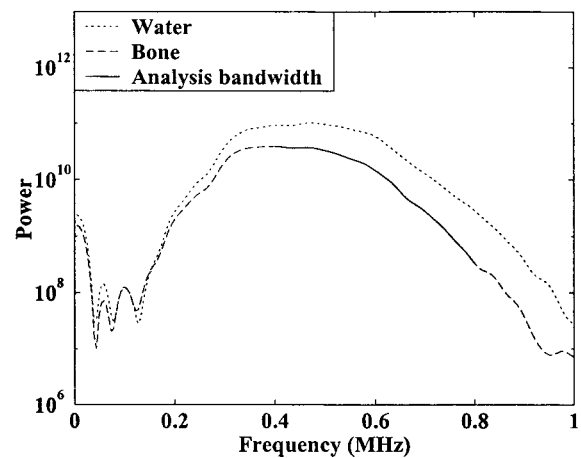


FIG. 4. Power spectra from through transmission measurements through water only (dotted line) and with bone in the acoustic path (dashed line). The range of frequencies used for analysis is shown by the solid line. The lower limit of the analysis band was determined by beam width not by signal-to-noise ratio considerations.

bone/water interface,  $d$  is the thickness of the bone sample, and  $\alpha(f)$  is the frequency-dependent amplitude attenuation coefficient. Taking the log spectral difference,

$$\alpha'(f) = \alpha(f) - \frac{2}{d} \log T = \frac{1}{2d} [\log X(f) - \log Y(f)], \quad (3)$$

where  $\alpha'(f)$  combines the effects of attenuation and transmission losses and may be referred to as the apparent attenuation coefficient or signal loss. If transmission losses are negligible, then the apparent attenuation coefficient equals the true attenuation coefficient.

Apparent attenuation versus frequency was least-squares fit to a linear function over the range from 400 to 800 kHz. The function was then characterized by its value at 500 kHz and the slope of the resulting line. This latter parameter is often referred to as “normalized broadband ultrasonic attenuation” (nBUA) in the bone densitometry field<sup>6</sup> and “attenuation slope” in the most of the rest of the biomedical ultrasonics field.<sup>29</sup> The speed of sound in the calcaneus, approximately 1475–1650 m/s,<sup>30</sup> is sufficiently close to that of distilled water at room temperature, 1487 m/s,<sup>31</sup> that potential diffraction-related errors in this substitution method due to the disparity in sound speeds between the two media<sup>32,33</sup> may be ignored.<sup>30</sup> Due to the high content of marrow in trabecular bone, the speed of sound is relatively close to that of water or soft tissue and much less than that of cortical bone. Note that the slope of apparent attenuation versus frequency is relatively unaffected by transmission losses provided that  $T$  is independent of frequency which is true for dispersionless media.<sup>34</sup> Dispersion in trabecular bone is minimal.<sup>30,35–38</sup>

## II. RESULTS

The densities of the 43 specimens ranged from 0.16 to 0.43 g/ml. The mean density was 0.27 g/ml. The standard deviation was 0.07 g/ml. The spectral properties of the broadband measurement system are illustrated in Fig. 4.

TABLE I. Experimental measurements for the 43 human calcaneus samples. The asterisk (\*) and dollar (\$) symbols denote that the null hypothesis that the mean difference equals zero can be rejected at the 95% and 90% confidence levels, respectively. The pound sign (#) denotes that the null hypothesis that the means from the AP and ML distributions are equal can be rejected at the 95% confidence level.

	Anteroposterior (AP) Mean $\pm$ SD(SE)	Mediolateral (ML) Mean $\pm$ SD(SE)	Difference AP-ML values Mean $\pm$ SD(SE)
Apparent attenuation coefficient at 500 kHz (dB/cm)	7.59 $\pm$ 4.51 (0.69)#	4.34 $\pm$ 2.83 (0.43)#	3.25 $\pm$ 2.59 (0.40)*
Attenuation slope or nBUA (dB/cmMHz)	11.80 $\pm$ 7.44 (1.13)	10.01 $\pm$ 5.56 (0.85)	1.79 $\pm$ 6.29 (0.96)\$
Backscatter coefficient at 500 kHz (cm <sup>-1</sup> Sr <sup>-1</sup> )	0.018 $\pm$ 0.016 (0.002)#	0.036 $\pm$ 0.042 (0.006)#	-0.019 $\pm$ 0.039 (0.006)*
Exponent ( <i>n</i> ) from power law fit to backscatter coefficient versus frequency, $BC=Af^n$	2.7 $\pm$ 1.4 (0.2)	3.2 $\pm$ 1.4 (0.2)	...

Apparent attenuation at 500 kHz was significantly greater for AP insonification than ML insonification. See Table I. The mean  $\pm$  standard deviation (standard error) for apparent attenuation coefficient at 500 kHz in the AP orientation was 7.59 $\pm$ 4.51(0.69) dB/cm compared with 4.34 $\pm$ 2.83(0.43) dB/cm in the ML orientation. Apparent attenuation coefficients, as functions of frequency, are plotted for the two orientations in Fig. 5.

Attenuation slope in the AP orientation was 11.80 $\pm$ 7.44(1.13) dB/cm MHz compared with 10.01 $\pm$ 5.56(0.85) dB/cm MHz in the ML direction. The difference between AP and ML attenuation slopes for individual samples had a mean of 1.79 dB/cmMHz and a standard error of 0.96 dB/cmMHz. A *t*-test revealed that this difference was significantly different from zero. (See Table I.) These results are consistent with the finding by Glüer *et al.* that the attenuation slope is larger along the axis of compressive trabeculae compared with the two perpendicular axes in bovine trabecular bone.<sup>23</sup>

In contrast to attenuation, the backscatter coefficient at

500 kHz was significantly lower for AP than ML insonification. The backscatter coefficient at 500 kHz in the AP orientation was 0.018 $\pm$ 0.016(0.002) cm<sup>-1</sup> Sr<sup>-1</sup> compared with 0.036 $\pm$ 0.042(0.006) cm<sup>-1</sup> Sr<sup>-1</sup> in the ML direction. The difference between AP and ML backscatter coefficients for individual samples had a mean of -0.019 cm<sup>-1</sup> Sr<sup>-1</sup> and a standard error of 0.006 cm<sup>-1</sup> Sr<sup>-1</sup>. Backscatter coefficients, as functions of frequency, are plotted for the two orientations in Fig. 6.

The exponents for the power law fits for backscatter coefficient versus frequency were close to the cubic model for both orientations. For AP insonification, the exponent was 2.7 $\pm$ 1.4 with a standard error of 0.2. For ML insonification, the exponent was 3.2 $\pm$ 1.4 with a standard error of 0.2. This approximate frequency-cubed variation is consistent with previous findings.<sup>13</sup> Cubic fits for experimentally measured backscatter coefficients for the two orientations are shown in Fig. 6.

Experimental results are summarized in Table I.

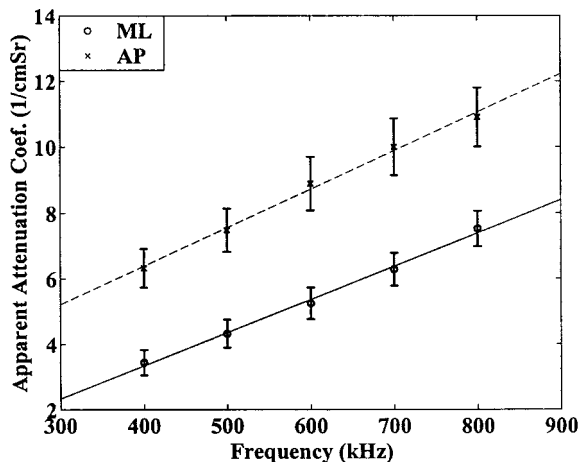


FIG. 5. Attenuation coefficients, as functions of frequency, for the two orientations. Linear fits to the data (solid line: ML, dashed line: AP) are also shown.

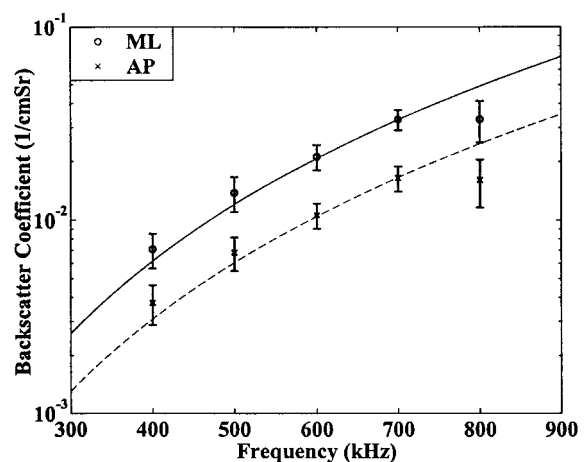


FIG. 6. Backscatter coefficients, as functions of frequency, for the two orientations. Cubic fits to the data (solid line: ML, dashed line: AP) are also shown.



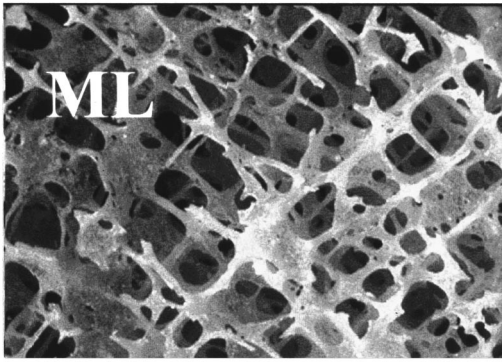


FIG. 7. Magnified view of calcaneal trabecular bone from ML orientation. The scale of the image is 11.8 mm×8.7 mm.

### III. DISCUSSION

The anisotropy of attenuation slope measured in the current study is similar to that reported by Glüer and co-workers based on their study in ten cubical specimens of bovine trabecular bone.<sup>13</sup> Both studies indicate minimal attenuation for insonification perpendicular to the predominant trabecular orientation. Glüer *et al.* found attenuation slope (BUA) to be 44%–54% larger along the axis of the compressive trabeculae as compared with the two perpendicular axes. The difference between the AP measurements and the ML measurements reported in the current study was 18% on the average. This diminished difference may be in part attributable to the substantial differences between human calcaneus and bovine proximal radius. In addition, while ML insonification in calcaneus fairly closely approximates perpendicular incidence to the trabecular axes, AP insonification in calcaneus is associated with a complex variety of angles relative to trabecular orientation. Thus the study by Glüer *et al.* may have been a more controlled comparison of parallel versus perpendicular incidence than the present investigation and hence revealed a greater difference. Nevertheless, both studies show the same basic trend.

The measured approximate cubic dependence of backscatter at diagnostic frequencies in the ML orientation reinforces an earlier investigation which indicated that not only backscatter but total scattering in all directions varies as frequency to the third power.<sup>13</sup> Attenuation (the combined result of absorption and scattering) in calcaneus is widely re-

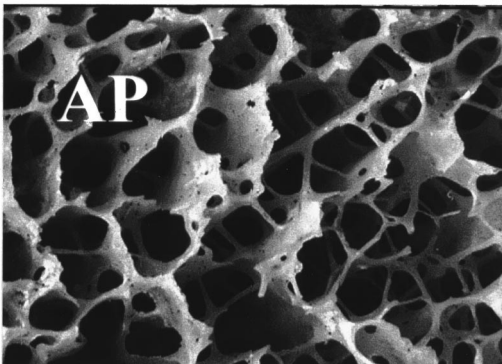


FIG. 8. Apparent magnified view of calcaneal trabecular bone from AP orientation. The scale of the image is 11.8 mm×8.7 mm.

ported to vary approximately as frequency to the first power.<sup>1–12</sup> These two findings could be consistent if and only if absorption is a greater component of attenuation than scattering.<sup>13</sup> (If, for example, scattering were the dominant component, then attenuation would also vary as frequency cubed, contrary to empirical observations.<sup>1–13</sup>) The measured approximate cubic dependence of backscatter in the AP direction is a previously unreported result. One possible explanation for this is that scattering in the AP orientation is dominated by that fraction of cylinders which happen to be oriented approximately perpendicular to the ultrasound propagation direction. This would also provide some explanation for the diminished magnitude of backscatter in the AP orientation (since the effective density of cylinders encountered in the AP direction is less than that for the ML direction—see Figs. 2, 7, and 8).

While attenuation was greater, backscatter at 500 kHz was much lower in the AP orientation compared with the ML orientation (see Table I). This is reasonable since, in the perpendicular (ML) orientation, the trabeculae offer the greatest possible cross-sectional area available to intercept the incident beam to produce scattering. The reduction of scattering in the AP orientation may extend to total scattering in all directions (not just backscatter). This would be the case if, as stated above, the cylindrical model (see the Introduction) is applicable to the AP orientation as well as the ML orientation (as is suggested by the approximate cubic frequency dependence of backscatter<sup>13</sup>) and the reduction in scattering in the AP direction is merely due to a decrease in the effective number of cylinders encountered by the incident ultrasound beam. This would imply that the excess attenuation exhibited in the AP orientation is not attributable to excess scattering. By default it must be due to excess absorption. This would further support the notion that absorption is a greater component of attenuation than scattering. Examination of Figs. 7 and 8 offers some support for the cylindrical model for both orientations, though in the AP case the cylinders appear somewhat more jagged.

These results were obtained *in vitro* from cadaveric specimens with relatively low bone mineral density. A certain amount of uncertainty with regard to the extent to which the conclusions of this study may be extended to higher density calcanea *in vivo* should be acknowledged.

### ACKNOWLEDGMENTS

The author is grateful for funding provided by the U.S. Food and Drug Administration Office of Women's Health. The author also appreciates the loan of a pulser and a receiver from Dr. Gerald Harris, FDA. The reference phantom used in this study was provided by Dr. Timothy J. Hall, University of Kansas.

<sup>1</sup>C. M. Langton, S. B. Palmer, and R. W. Porter, "The measurement of broadband ultrasonic attenuation in cancellous bone," *Eng. Med.* **13**, 89–91 (1984).

<sup>2</sup>P. Rossmann, J. Zagzebski, C. Mesina, J. Sorenson, and R. Mazess, "Comparison of Speed of Sound and Ultrasound Attenuation in the Os Calcis to Bone Density of the Radius, Femur and Lumbar Spine," *Clin. Phys. Physiol. Meas* **10**, 353–360 (1989).

<sup>3</sup>M. B. Tavakoli and J. A. Evans, "Dependence of the velocity and attenu-

- ation of ultrasound in bone on the mineral content," *Phys. Med. Biol.* **36**, 1529–1537 (1991).
- <sup>4</sup>M. Schott *et al.*, "Ultrasound discriminates patients with hip fracture equally well as dual energy X-ray absorptiometry and independently of bone mineral density," *J. Bone Miner. Res.* **10**, 243–249 (1995).
  - <sup>5</sup>C. H. Turner *et al.*, "Calcaneal ultrasonic measurements discriminate hip fracture independently of bone mass," *Osteoporosis Int.* **5**, 130–135 (1995).
  - <sup>6</sup>C. M. Langton *et al.*, "Prediction of Mechanical Properties of the Human Calcaneus by Broadband Ultrasonic Attenuation," *Bone (N.Y.)* **18**, 495–503 (1996).
  - <sup>7</sup>D. Hans *et al.*, "Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study," *Lancet* **348**, 511–514 (1996).
  - <sup>8</sup>C. C. Glüer *et al.*, "Osteoporosis: Association of recent fractures with quantitative US findings," *Radiology* **199**, 725–732 (1996).
  - <sup>9</sup>D. C. Bauer *et al.*, "Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women," *Arch. Intern Med.* **157**, 629–634 (1997).
  - <sup>10</sup>C. C. Glüer, "Quantitative ultrasound techniques for the assessment of osteoporosis: expert agreement on current status," *J. Bone Miner. Res.* **12**, 1280–1288 (1997).
  - <sup>11</sup>P. Thompson, J. Taylor, A. Fisher, and R. Oliver, "Quantitative heel ultrasound in 3180 women between 45 and 75 years of age: compliance, normal ranges and relationship to fracture history," *Osteoporosis Int.* **8**, 211–214 (1998).
  - <sup>12</sup>C. C. Glüer and D. Hans, "How to use ultrasound for risk assessment: a need for defining strategies," *Osteoporosis Int.* **9**, 193–195 (1999).
  - <sup>13</sup>K. A. Wear, "Frequency dependence of ultrasonic backscatter from human trabecular bone: Theory and experiment," *J. Acoust. Soc. Am.* **106**, 3659–3664 (1999).
  - <sup>14</sup>P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton U.P., Princeton, NJ, 1986).
  - <sup>15</sup>J. J. Faran, "Sound scattering by solid cylinders and spheres," *J. Acoust. Soc. Am.* **23**, 405–418 (1951).
  - <sup>16</sup>V. Roberjot, P. Laugier, P. Droin, P. Giat, and G. Berger, "Measurement of integrated backscatter coefficient of trabecular bone," in *Proc. 1996 IEEE Ultrason. Symp. Vol. 2*, pp. 1123–1126, (1996).
  - <sup>17</sup>K. A. Wear and B. S. Garra, "Assessment of bone density using broadband ultrasonic backscatter," *Proc. 22nd Int. Symp. Ultrason. Imag. and Tissue Char.*, Washington, DC (1997), p. 14 (abstract).
  - <sup>18</sup>P. Giat, C. Chappard, C. Roux, P. Laugier, and G. Berger, "Preliminary clinical assessment of the backscatter coefficient in osteoporosis," in *Proc. 22nd Int. Symp. Ultrason. Imag. and Tissue Char.*, Washington, DC (1997) p. 16 (Abstract).
  - <sup>19</sup>K. A. Wear and B. S. Garra, "Assessment of bone density using ultrasonic backscatter," *Ultrasound Med. Biol.* **24**, 689–695 (1998).
  - <sup>20</sup>J. G. Mottley and J. G. Miller, "Anisotropy of the ultrasonic backscatter of myocardial tissue: I. Theory and measurements *in vitro*," *J. Acoust. Soc. Am.* **83**, 755–761 (1988).
  - <sup>21</sup>E. I. Madaras, J. Perez, B. E. Sobel, J. G. Mottley, and J. G. Miller, "Anisotropy of the ultrasonic backscatter of myocardial tissue: II. Measurements *in vivo*," *J. Acoust. Soc. Am.* **83**, 762–769 (1988).
  - <sup>22</sup>M. F. Insana, T. J. Hall, and J. L. Fishback, "Identifying acoustic scattering sources in normal renal parenchyma from the anisotropy in acoustic properties," *Ultrasound Med. Biol.* **17**, 613–626 (1991).
  - <sup>23</sup>C. C. Glüer, C. Y. Wu, and H. K. Genant, "Broadband ultrasound attenuation signals depend on trabecular orientation: an *in vitro* study," *Osteoporosis Int.* **3**, 185–191 (1993).
  - <sup>24</sup>J. M. Alves, J. T. Ryaby, J. J. Kaufman, P. P. Magee, and R. S. Siffert, "Influence of marrow on ultrasonic velocity and attenuation in bovine trabecular bone," *Calcif. Tissue Int.* **58**, 362–367 (1996).
  - <sup>25</sup>C. F. Njeh and C. M. Langton, "The effect of cortical endplates on ultrasound velocity through the calcaneus: an *in vitro* study," *Br. J. Radiol.* **70**, 504–510 (1997).
  - <sup>26</sup>J. A. Zagzebski, L. X. Yao, E. J. Boote, and Z. F. Lu, "Quantitative Backscatter Imaging," in *Ultrasonic Scattering in Biological Tissues*, edited by K. K. Shung and G. A. Thieme (CRC, Boca Raton, FL, 1993).
  - <sup>27</sup>K. A. Wear, B. S. Garra, and T. J. Hall, "Measurements of ultrasonic backscatter coefficients in human liver and kidney *in vivo*," *J. Acoust. Soc. Am.* **98**, 1852–1857 (1995).
  - <sup>28</sup>J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).
  - <sup>29</sup>R. Kuc and M. Schwartz, "Estimating the acoustic attenuation coefficient slope for liver from reflected ultrasound signals," *IEEE Trans. Sonics Ultrason.* **SU-26**, 353–362 (1979).
  - <sup>30</sup>P. Droin, G. Berger, and P. Laugier, "Velocity dispersion of acoustic waves in cancellous bone," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 581–592 (1998).
  - <sup>31</sup>A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981), p. 31.
  - <sup>32</sup>W. A. Verhoef, M. J. T. M. Cloostermans, and J. M. Thijssen, "Diffraction and dispersion effects on the estimation of ultrasound attenuation and velocity in biological tissues," *IEEE Trans. Biomed. Eng.* **BME-32**, 521–529 (1985).
  - <sup>33</sup>W. Xu and J. J. Kaufman, "Diffraction correction methods for insertion ultrasound attenuation estimation," *IEEE Trans. Biomed. Eng.* **40**, 563–570 (1993).
  - <sup>34</sup>L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics* (Wiley, New York, 1982), Chap. 6.
  - <sup>35</sup>P. H. F. Nicholson, C. G. Lowet, C. M. Langton, J. Dequeker, and G. Van der Perre, "Comparison of time-domain and frequency-domain approaches to ultrasonic velocity measurements in trabecular bone," *Phys. Med. Biol.* **41**, 2421–2435 (1996).
  - <sup>36</sup>R. Strelitzki and J. A. Evans, "On the measurement of the velocity of ultrasound in the os calcis using short pulses," *Eur. J. Ultrasound* **4**, 205–213 (1996).
  - <sup>37</sup>K. A. Wear, "The effects of frequency-dependent attenuation and dispersion on group velocity measurements: applications in human trabecular bone," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **47**, 265–273 (2000).
  - <sup>38</sup>K. A. Wear, "Measurements of phase velocity and group velocity in human calcaneus," *Ultrason. Med. Biol.* (in press).

# Cavitation nucleation agents for nonthermal ultrasound therapy

Douglas L. Miller, Oliver D. Kripfgans, J. Brian Fowlkes, and Paul L. Carson  
Department of Radiology, University of Michigan Medical Center, Ann Arbor, Michigan 48109

(Received 30 October 1999; accepted for publication 26 February 2000)

The use of a nucleation-promoting agent can greatly enhance therapeutically useful nonthermal bioeffects. A blank agent (saline), Optison™ ultrasound contrast agent, a stabilized perfluoropentane droplet suspension (SDS), and retained air space were compared as nucleation agents in whole blood. Fresh canine whole blood with added agent was exposed in 1.3-ml disposable pipette bulbs to lithotripter shock waves (2-Hz rate; +24.4, -5.2 MPa peak pressure amplitudes). Cavitation activity was assessed by measuring hemolysis. The droplet suspension performed nearly as well as retained air when added at a concentration sufficient to provide a roughly equal volume of gas after vaporization. Optison™ also yielded nucleation, but a concentration of 10%–20% was needed for large enhancement of hemolysis comparable to 5% SDS. Exposure at room temperature, which was less than the 29 °C boiling point of perfluoropentane, eliminated the enhancement of the hemolysis effect relative to the blank. Application of 100-kPa excess pressure during exposure reduced but did not eliminate the nucleation ability of Optison™, SDS, or retained air. However, this small pressure (relative to the peak positive pressure of the shock waves) eliminated the hemolysis induced with the blank agent. The stabilized perfluoropentane droplet suspension appears to be a good nucleation agent for nonthermal ultrasound therapy applications. © 2000 Acoustical Society of America. [S0001-4966(00)01606-4]

PACS numbers: 43.80.Gx, 43.80.Sh, 43.35.Ei, 43.35.Wa [FD]

## INTRODUCTION

Nonthermal bioeffects of high-pressure-amplitude ultrasound can be quite dramatic when bubbles or cavitation nuclei are available to initiate vigorous cavitation activity. The pulsation of the cavities transfers ultrasonic energy into mechanical perturbation of the surrounding medium or tissue. Cavitation can produce a range of effects from sonoporation of cells, which survive the treatment, to cell lysis and homogenization of tissue. Some of the cavitation processes are unique, and likely impossible to produce by any other noninvasive modality. Although highly localized near individual cavities, these phenomena hold the potential for development of novel applications of the nonthermal bioeffects for ultrasound therapy.

Under conditions of natural nucleation, cavitation is difficult to produce in animals (except gas-bearing tissues, such as lung and intestine). Nuclei are particularly sparse in whole blood flowing within the heart and arteries, and cavitation is difficult to produce there even with lithotripter shock waves (Williams *et al.*, 1989). Nevertheless, sufficiently high-amplitude ultrasound pulses can reliably cavitate whole blood (Ivey *et al.*, 1995). Other tissues are more susceptible to cavitation, which can produce unwanted side effects during lithotripsy [e.g., hemorrhage (Delius, 1994)], or ultrasound hyperthermia treatment (Fry *et al.*, 1970; Vykhodtseva *et al.*, 1994). For lithotripter treatment, the cavitation threshold has been found to be in the range 1.5–3.5 MPa in human tissue (Coleman *et al.*, 1995). However, the occurrence of cavitation and concomitant effects *in vivo* tend to be sparse, random events which depend on the presence of cavitation nuclei (e.g., microscopic gas bodies or other inhomogeneities) in the medium.

For nonthermal therapy, a key to control and application of *in vivo* cavitation therefore appears to be the augmentation and deliberate control of the nucleation process. One simple method is to place a macroscopic gas bubble within an exposure chamber or tissue exposed to nonthermal-mode treatment, such as lithotripter shock waves (Miller and Thomas, 1996). In a study of sonoporation of erythrocytes by shock waves *in vitro*, this method was very effective in regulating the nucleation and resulting effects (Miller *et al.*, 1998). A 0.2-ml air space was retained when 1.3-ml pipette-bulb exposure chambers were filled with 45% hematocrit red-cell suspension. Hemolysis was enhanced by about a factor of 10 relative to a full chamber. The reproducible, well-regulated results could be described by a simple theory in terms of exponential functions of the number of shock waves. Williams *et al.* (1999) have also noted the enhancement of lysis by macroscopic gas bubbles for *in vitro* lithotripter exposures. This method has been used *in vivo*. Prat *et al.* (1993) utilized bubbles to enhance cavitation tissue destruction for cancer therapy with lithotripters. Air bubbles mixed into gelatin were infused into the abdominal cavity of rats in combination with chemotherapy, which improved survival. In an attempt to combine cell killing with sonoporation for gene therapy of mouse melanoma tumors, nucleation was assured by injecting free air at 10% of the tumor volume together with a DNA plasmid solution (Bao *et al.*, 1998). Tumor cell survival decreased, and DNA transfer increased with increasing number of shock waves. Air injection plus 800 shock waves gave approximately a sevenfold increase in gene transfection relative to injection of DNA alone. This method is simple and reliable, but has the obvious disadvantage of the risk of air embolism.

An alternative nucleation method is to inject an ultrasound contrast agent, such as Albunex<sup>®</sup> or Optison<sup>™</sup> (Mallinckrodt, Inc.), which consist of suspensions of stabilized air or perfluorocarbon bubbles, respectively. The gas bodies are a few microns in diameter and pass through the circulation without producing significant embolization. Addition of up to 10% of Albunex<sup>®</sup> to whole blood improved results of hemolysis by lithotripter shock waves, but the enhancement was not statistically significant relative to the erratic result without added agent (Miller and Thomas, 1996). The hemolysis effect with lithotripter shock waves without added agents still appears to depend on cavitation phenomena; for example, Delius (1997) reported a significant reduction in hemolysis when a small (100 kPa) hydrostatic overpressure was applied. This agent did significantly enhance the hemolysis observed relative to exposure without added gas bodies after 1.28-MHz focused ultrasound treatment with 100- $\mu$ s bursts repeated at 10 ms (Miller and Thomas, 1996), and for 1.1-MHz high-intensity focused ultrasound (Polichik *et al.*, 1999). Under suitable conditions (e.g., high concentration, or special rotating chambers), hemolysis can be produced by moderate-amplitude ultrasound with contrast agent (Brayman *et al.*, 1996; Miller *et al.*, 1997; Miller and Gies, 1998a), and the cavitation nuclei appear to survive the destruction of the gas bodies (Brayman and Miller, 1997). *In vivo*, several studies have been conducted to determine the potential for bioeffects to result from the interaction of ultrasound with contrast-agent gas bodies. The addition of contrast agents to the circulation enhances the vascular damage induced by lithotripter shock waves in intestine and other organs (Dalecki *et al.*, 1997a). The contrast-agent-related bioeffects can be significant: mice exposed to 800 lithotripter shock waves with 1 ml/kg Albunex<sup>®</sup> had a significantly reduced survival rate relative to the same exposure without contrast agent (Miller and Gies, 1999). At lower amplitudes of pulsed ultrasound, the addition of Albunex<sup>®</sup> to the circulation of mice results in hemolysis (Dalecki *et al.*, 1997b) and in enhanced induction of petechiae in the intestinal wall approximately in proportion to the dose of contrast agent (Miller and Gies, 1998b). Although ultrasound contrast agents are clearly capable of nucleating cavitation activity, they too have disadvantages, including instability (e.g., they are destroyed by the ultrasound; Walker *et al.*, 1997) and rapid removal from the circulation.

Recently, stabilized perfluorocarbon droplets have been produced and examined in regard to bubble placement *in vivo* for diagnostic or therapeutic purposes (Kripfgans *et al.*, 1999; Kripfgans *et al.*, in press). Perfluoropentane droplets of about 1.5- to 4- $\mu$ m diameter are stabilized by an albumin coating. If the droplets are placed in a medium above the boiling point of 29 °C, they do not vaporize, but rather remain as superheated droplets. However, the superheated droplets can be vaporized by ultrasound into relatively large gas bubbles, which can readily be seen on a diagnostic ultrasound image. It should be noted that these droplets are much larger than those utilized in the ultrasound contrast agent Echogen (Correas *et al.*, 1997).

In regard to the cavitation nucleation problem for non-thermal therapy, such droplets appear to combine the small

size and relative safety of the contrast-agent method with a potential for site-targeted production of large bubbles with long-term stability, similar to injected air. In this study, this idea was tested by comparing the cavitation nucleating ability of stabilized droplet suspension (SDS) to that of a blank, Optison<sup>™</sup>, and retained air space in an *in vitro* exposure chamber. The nucleation and cavitation effectiveness was detected and quantified by measuring the hemolysis for lithotripter shock wave exposure of whole blood. Since whole blood simulates *in vivo* conditions, this system provided a simple means to evaluate the nucleating ability of various agents. The results were promising for the development of nucleation agents for nonthermal ultrasound therapy.

## I. METHODS

The lithotripter system used for this study has been described previously (Miller and Thomas, 1995). Briefly, the system was similar to a Dornier HM-3 lithotripter with an elliptical brass reflector fitted with a standard spark gap (Dornier Medical Systems, Kennesaw, GA). The system was charged at 20 kV and discharged at a 2-Hz rate. The field at the focus, 12 cm from the mouth of the reflector, was measured with a bilaminar shielded hydrophone with a 0.5-mm sensitive spot (Marconi type Y-34-3598, National Physical Laboratory, Middlesex, UK) coupled to a 40-MHz-bandwidth digital oscilloscope. For nine discharges with a medium-age gap (i.e., about 2000 discharges), the spatial peak pressure amplitude averaged 24.4 MPa (5.2-MPa standard deviation) peak positive (40-MHz bandwidth) and 5.2 MPa (1.3 MPa s.d.) peak negative, with a mean pressure amplitude [i.e.,  $(p_+ + p_-)/2$ ] of 14.8 MPa. These pressures were somewhat less than those of commercial lithotripters, but this allowed extended use of the gaps, which were changed regularly after a maximum of 4200 discharges. The -6-dB beamwidth was approximately 20 mm. The shock wave was also characterized by the pulse-integration method (AIUM 1994). At the spatial peak, the pulse average intensity was 3300 W cm<sup>-2</sup> with a 5.3- $\mu$ s duration. The pulse energy is also a useful exposure metric for lithotripsy (Delius *et al.*, 1995). The energy delivered at the spatial peak was 174 J m<sup>-2</sup>, and totaled about 52 mJ per pulse when integrated over the cross section of the exposure chamber (described below).

A vial of stabilized droplet suspension was prepared for each day's experiments, as described previously (Kripfgans *et al.*, in press). Briefly, vials containing saline with albumin and 10% by volume of perfluoropentane (PFP) were vigorously agitated in a vial shaker to emulsify the PFP into droplets. The suspension products have been characterized by digital photomicroscopy (DRMP, Leica, Inc., Deerfield, IL with Spot SP104 digital camera, Diagnostic Instruments, Inc., Sterling Heights, MI) on a hemacytometer with subsequent image analysis (using MATLAB 5.1, Mathworks, Inc., Natick, MA). For this study, an albumin concentration of 0.085% was used, which yields about 480 · 10<sup>6</sup> droplets ml<sup>-1</sup> with a mean diameter of about 3  $\mu$ m (Kripfgans *et al.*, in press). The variable perfluorocarbon concentration of each initial droplet suspension was determined by centrifuging the droplets in a hematocrit tube, which gave an average volume

fraction of 5.2%. To normalize the suspensions, each initial suspension was diluted (by an average factor of 8) to contain 0.65% perfluoropentane liquid in the final stabilized droplet suspension (SDS). At this volume fraction, the gas produced by vaporization would equal the initial volume of liquid, assuming ideal gas conditions and complete vaporization (i.e., 154 times volume). With this suspension, 1% SDS added to blood would yield about 1% free PFP gas by volume if all the droplets vaporized upon treatment. The SDS contained about 60 000 000 droplets per ml. Vaporization of the average 3- $\mu\text{m}$  diameter droplet would yield a roughly 16- $\mu\text{m}$  diameter bubble (i.e., 5.3 times radius), neglecting the effects of surface tension and diffusion of gases into and out of the bubble, which can be important under some conditions (Kabalnov *et al.*, 1998). The other nucleating agents were a blank, contrast agent, and air. The blank agent consisted of phosphate-buffered saline with 1% by weight of bovine serum albumin and was centrifuged after mixing to remove all visible air bubbles. Optison™ ultrasound contrast agent was purchased from Mallinckrodt, Inc. (St. Louis, MO) and carefully withdrawn from the sealed vial using an Optispikes™ to minimize loss of the perfluoropropane gas in the head space of the vial. Optison™ contains 500–800 · 10<sup>6</sup> perfluoropropane gas bodies per ml, with a mean diameter of 2 to 4.5  $\mu\text{m}$  (Mallinckrodt, 1998). For comparison to the SDS, 10% Optison™ contains only about 0.3% gas, under similar assumptions.

Canine whole blood was collected by sterile venipuncture into blood collection tubes with EDTA anticoagulant (Vacutainer no. 6457, Becton-Dickinson and Co., Franklin Lakes, NJ). All animal procedures were performed in accordance with the guidance and approval of the University Committee on Use and Care of Animals in Research, and the Unit for Laboratory Animal Medicine. Several tubes were pooled for each day's experimentation and used without dilution (hematocrits were in the range 48% to 51%) within about 4 h of being drawn. For exposure, the desired mixture was drawn into disposable transfer pipette bulbs (Cryo-Pet no. 241, Saint-Amand Mfg. Co., San Fernando, CA), which were 0.9 cm in diameter and about 2.6 cm long and acoustically transparent (Miller and Thomas, 1996). The SDS, contrast and blank agents were mixed with whole blood in separate microfuge tubes, from which the exposure bulb was filled. For the air bubble, a measured amount of air was retained in the pipette-bulb exposure chambers by filling the bulb completely with blood containing blank agent and then removing some liquid with a calibrated pipette tip (leaving air to refill the bulb). The suspension always contained 80% whole blood, plus blank and active agent mixed in the desired concentration to make up the other 20%. The order of sample exposure was changed for each repetition to avoid any systematic trend with time (e.g., gap wear or blood age). The bulb was mounted in a 37 °C water bath at the focus of the lithotripter with the bulb oriented vertically upward, which allowed the retained air bubble to rise to the rounded end of the bulb. The upper portion of the bulb was exposed from the side, with the beam axis horizontal, which reduced exposure of the bulb neck and mounting fixture. The pipette neck was attached to a small plastic tube, which was filled

with saline. The other end of the tube was routed out of the bath to provide for venting any excess pressure created by droplet vaporization. For some experiments, the tube was attached to a syringe with pressure gauge, which allowed the bulb to be pressurized to an excess 100 kPa (i.e., 200 kPa total). The water in the exposure bath was degassed by vacuum for 1 h to about 10%–20% of saturation at 25 °C before use, and was continuously filtered through a 0.2- $\mu\text{m}$  filter to minimize the occurrence of cavitation in the water. After exposure, the pipette contents were removed and sampled for determination of free haemoglobin. The sample was diluted by a factor of 10 with PBS, centrifuged, and the absorbance of the supernatant measured using a spectrophotometer set to 414 nm. This instrument was zeroed with PBS, and calibrated using a blood sample hemolyzed in water. Results are presented as the mean percentage of hemolysis of four independent repetitions with tests for significant differences between means by the two-sided *t*-test at the  $P < 0.05$  level.

## II. RESULTS

Each agent behaved in a characteristic manner upon exposure to shock waves (SW). The appearance of each agent in blood serum (rather than whole blood) before and after exposure to 200 SWs was observed to visualize the bubbles present. Photographs of the 9-mm-diameter bulbs lying hori-

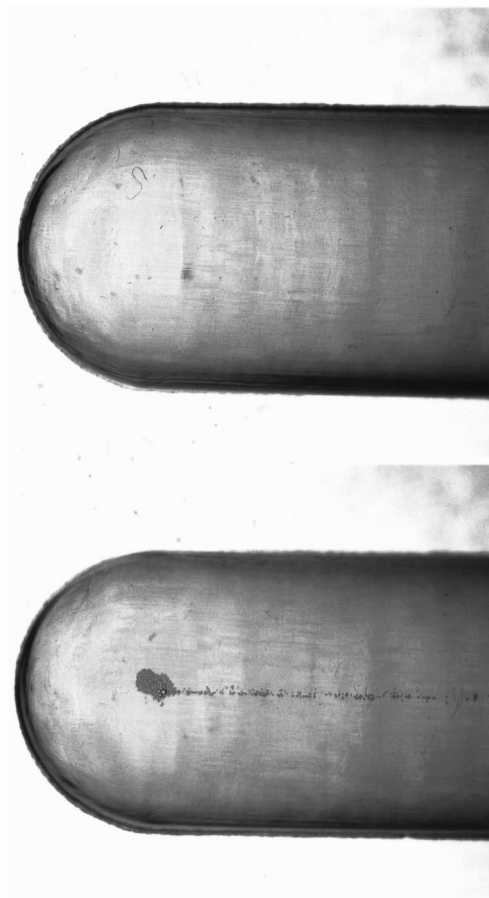


FIG. 1. Photograph of the bulb with the blank agent in serum before (top) and after (bottom) exposure to 200 SWs. A few bubbles were produced by the treatment.

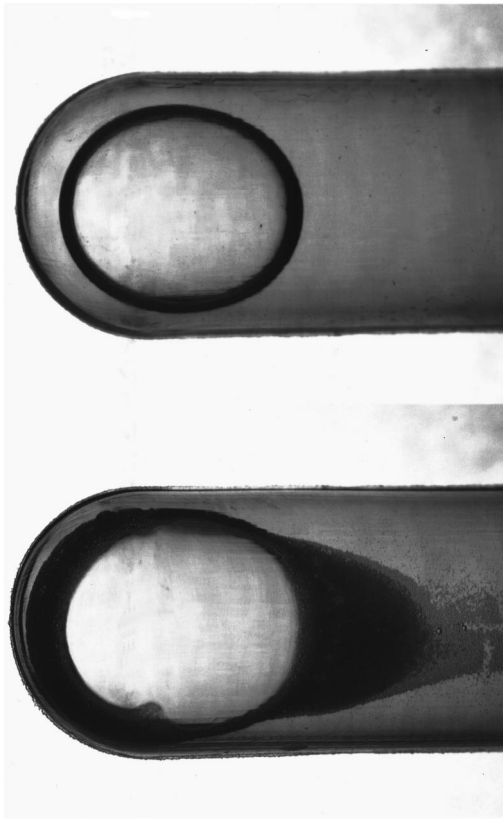


FIG. 2. Photograph of the bulb with 10% retained air in serum before (top) and after (bottom) exposure to 200 SWs. Many small bubbles were produced by the treatment, which appear dark in this backlit photograph.

zontally to show the interior bubbles are presented in Figs. 1–4 for the blank agent, retained air, 10% SDS, and 10% contrast agent, respectively. Using back lighting, small bubbles appear dark and large bubbles refract the light to give a bright center in the photographs. The blank agent often exhibited a few bubbles after exposure, presumably resulting from “natural” cavitation. The retained air, which was initially one large bubble, was partly broken into many small bubbles, possibly in a process similar to the production of contrast agents by sonication. The essentially clear SDS was transformed into a mass of bubbles by the exposure. The contrast agent was less dramatically changed by the exposure, with an apparent loss of many small bubbles and the appearance of a few bubbles much larger than the initial size range.

Each of the four agents was used in various concentrations for exposure with 200 SWs. The amount of retained air was difficult to measure, and was limited to 5%, 10%, and 20%. Optison™ was also used at these concentrations. The SDS was tested over a 0.5% to 20% range, which approximately matched the gas volume fraction of air (a volume of SDS would yield approximately an equal volume of gas if vaporized, see Methods), and of Optison™ (the 20%-Optison™ mixture would contain roughly 0.6% gas). Results are shown in Fig. 5. The mean result for the blank agent is also shown for its normal 20% fraction, which was nearly ( $P \sim 0.06$ ) significantly different from a sham exposure. Selected data are also presented in Table I. All the other agents and concentrations produced significantly more hemolysis

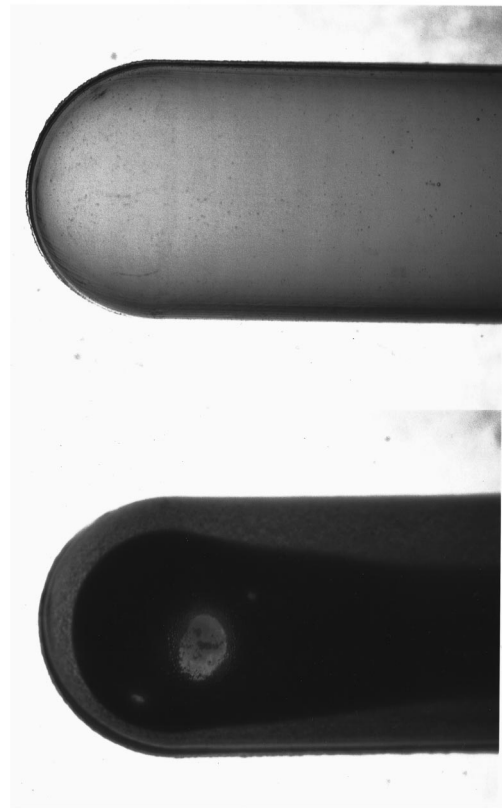


FIG. 3. Photograph of the bulb with stabilized droplet suspension in serum before (top) and after (bottom) exposure to 200 SWs. Many bubbles were produced by the treatment, which made the bulb nearly opaque.

(i.e.,  $P < 0.05$ ) than sham exposure and also more than the blank agent exposure, except 5% contrast agent addition (since 5% Optison™ did not improve the result over the blank, lower concentrations of this agent were not tested). One-percent droplet suspension was about as effective as 10% contrast agent. Ten-percent droplet suspension was about as effective as 10% air. The hemolysis results tended to increase with concentrations, but this trend seemed to be relatively weak at the lower concentrations of SDS.

The hemolysis found for increasing numbers of shock waves is shown in Fig. 6. For these exposures, 10% of retained air, SDS, and contrast agent were used. The response increased monotonically with increasing exposure, reaching nearly 100% hemolysis for 10% retained air after 1000 shock waves. The result for exposure with 20% blank agent is shown for comparison at 200 SWs. From previous work (Miller and Thomas, 1996), the hemolysis with the blank would be expected to increase to about 5% hemolysis for 1000 SWs.

In order for the droplets to instigate cavitation activity, the perfluorocarbon must vaporize to form gas bubbles. The ambient temperature is critical for this process, because the boiling point of perfluoropentane is 29 °C. An experiment was performed with the blank, 10% droplets, and 10% air at room temperature (i.e., about 22 °C). The hemolysis obtained with the droplets was dramatically reduced at room temperature compared to 37 °C, as shown in Fig. 7, to a value which was not significantly different from the blank agent. The hemolysis obtained with the retained air remained significant

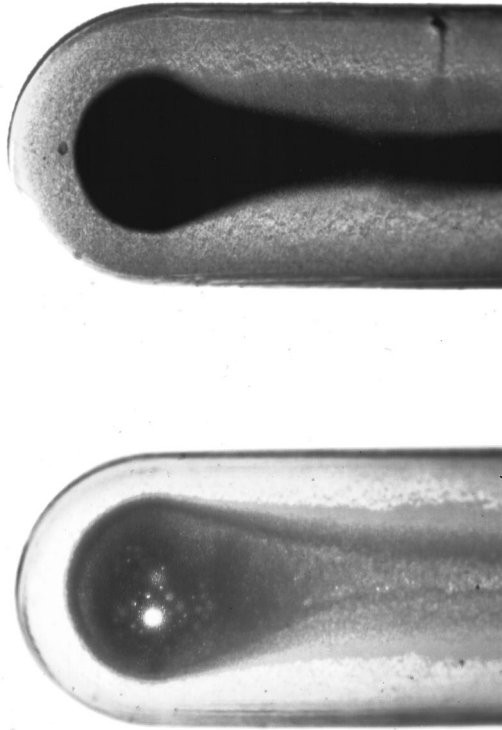


FIG. 4. Photograph of the bulb with Optison™ in serum before (top) and after (bottom) exposure to 200 SWs. The smaller bubbles present initially appear to have been destroyed, and a few relatively large bubbles produced by the treatment.

at room temperature, although the effect was reduced somewhat (the reduction was not statistically significant at  $P \sim 0.08$ ). Another important parameter for cavitation activity is ambient pressure. All four agents were tested with 100-kPa excess pressure applied immediately before and during exposure with 200 SWs. The effect obtained with the blank agent was reduced by a factor of 10, to a level which was not significantly different from the result of sham exposure. This elimination of the effect for the blank agent did not occur if the excess pressure was applied for 100 s, then released immediately before exposure (statistically significant relative to sham, and to the exposed blank), as shown in Fig. 8. The hemolysis obtained with the other agents were all reduced by factors of 2 to 4 by pressurization during exposure, but remained significant relative to sham exposure.

### III. DISCUSSION

Due to a paucity of cavitation nuclei *in vivo*, the use of a nucleation-enhancing agent can greatly improve production

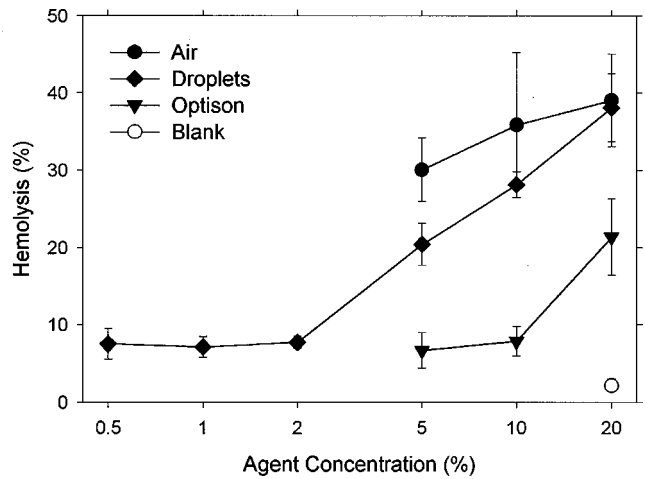


FIG. 5. Hemolysis results presented as the mean of four repeated measurements with standard error bars. The concentrations of the agents were varied for exposure to 200 SWs. The result for 20% blank agent is shown for comparison (standard error bars are smaller than the symbol).

of therapeutically useful nonthermal bioeffects. In this study, a blank (saline), Optison™ ultrasound contrast agent, a stabilized perfluoropentane droplet suspension, and retained air space were compared as nucleation agents in whole blood. The agents were mixed with freshly drawn whole blood and exposed to lithotripter shock waves in pipette-bulb exposure chambers. The droplet suspension promoted hemolysis nearly as well as retained air when added at a concentration with the potential to provide a roughly equal volume of gas after vaporization. Contrast agent also yielded nucleation, but equivalent effects did not seem to require an equal volume of gas. For example, 5% SDS, expected to provide roughly 5% free gas, gave 20% hemolysis after 200 SWs, while 20% Optison™, expected to provide only about 0.6% gas, produced 21% hemolysis. One explanation of this phenomenon might be a more populous distribution of the gaseous nuclei throughout the chamber volume. The 20% Optison™ would contain about  $100\text{--}160 \cdot 10^6$  gas bodies per ml, while the 5% SDS would contain only about  $3 \cdot 10^6$  droplets per ml. The Optison™ appeared to perform about the same as Alunex®, a related air-based contrast agent tested previously (Miller and Thomas, 1996). The agents all gave similar dose-response trends with increasing shock wave numbers. The response for SDS appears to be divided into two trends (i.e., a constant effect at low concentrations then a steep increase in effect with concentration) in Fig. 5, but it is un-

TABLE I. Numerical hemolysis results for several points displayed in Fig. 5. The data are the means of four repetitions with standard error in parentheses. The estimated percentage of free gas after vaporization was zero for the blank, 0.3% for 10% Optison™, and the same as the concentration for SDS and retained air. A sham exposure was conducted with each agent, and the very similar results were simply averaged. Exposures were to 200 SWs.  $P$  values are given for student  $t$ -tests of the hemolysis obtained with the agents with respect to (wrt) sham exposure or exposure with the gas-body-free blank.

	Sham		Exposure with nucleation agent			
	Mixed	Blank	Optison™ 10%	SDS 1%	SDS 10%	Air 10%
Hemolysis%	0.3 (0.05)	2.2(0.8)	7.9(1.9)	7.1(1.4)	28.2 (1.7)	30.1(4.1)
Significance wrt sham		$\sim 0.06$	$<0.01$	$<0.005$	$<10^{-5}$	$<0.001$
Significance wrt exposure with blank			$<0.05$	$<0.05$	$<10^{-4}$	$<0.02$

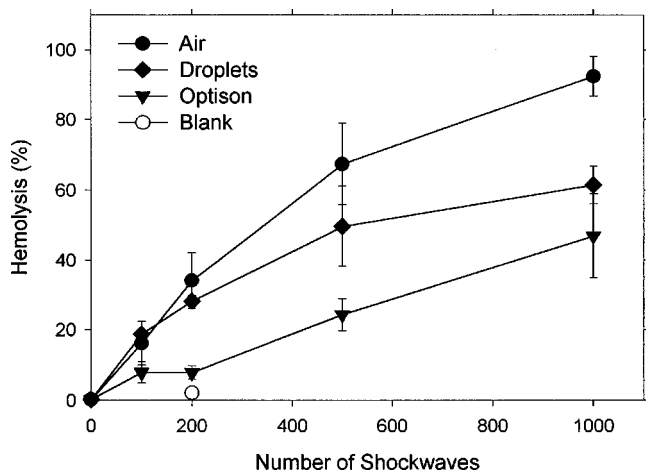


FIG. 6. Increasing hemolysis with increasing numbers of SWs for 10% additions of each agent. All the agents provided nucleation, with hemolysis reaching nearly 100% for the retained air.

certain whether this was real or not (due to the small sample sizes). The retained air was the most effective agent, possibly because of the the macroscopic mixing action obtained with the free air space (Miller *et al.*, 1998).

The droplets are superheated at normal body temperature of 37 °C. The ultrasound exposure apparently initiates vaporization of these droplets as noted previously (Kripfgans *et al.*, in press). This superheated state is critical to the nucleation process. Exposure at room temperature, which was less than the 29 °C boiling point of perfluoropentane, eliminated the enhancement of the hemolysis effect relative to the blank. Application of 100-kPa excess hydrostatic pressure reduced but did not eliminate the nucleation ability of contrast agent, SDS, or retained air. However, this small pressure (relative to the 24.4-MPa peak positive pressure of the shock waves) eliminated the hemolysis produced with the blank agent. This confirms the earlier observation of this phenomenon by Delius (1997). The fact that a pretreatment

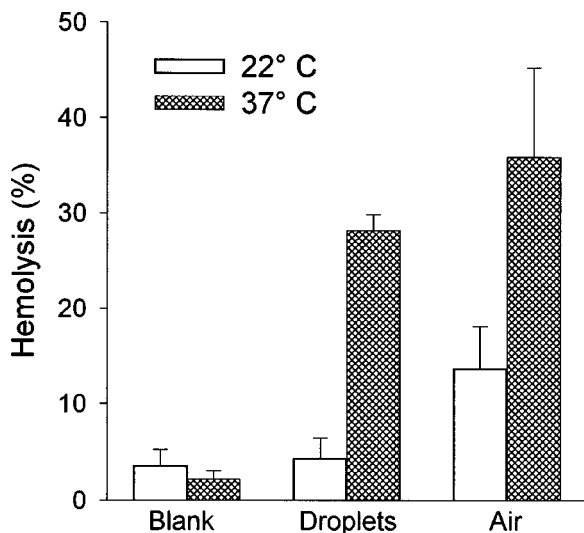


FIG. 7. The influence of reduced bath temperature on the hemolysis results for 200 SWs. The enhancement provided by the droplet suspension relative to the blank agent was eliminated by the temperature below the boiling point of perfluoropentane.

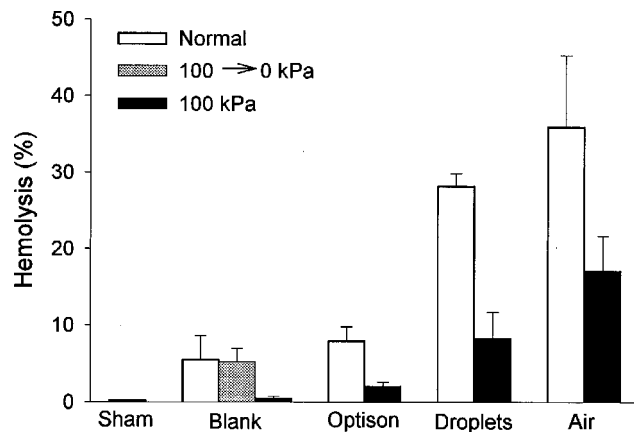


FIG. 8. The influence of increasing the ambient pressure on the hemolysis results. The 100-kPa excess pressure reduced but did not eliminate the enhancement of hemolysis from 200 SWs by 10% added agent. However, the added pressure eliminated the hemolysis induced by shock wave exposure with the blank agent.

with excess pressure did not reduce the effects of exposure at atmospheric pressure indicates that the excess pressure did not destroy pre-existing nuclei. The excess pressure seems to suppress cavitation activity during exposure, perhaps by helping to shrink unstable gaseous cavities which might otherwise persist during the period between pulses.

#### IV. CONCLUSION

These results are encouraging for the development of *in vivo* uses of the perfluoropentane droplet suspension. Perfluorochemicals have an extensive record of medical usage, and are well tolerated (see, e.g., Mattrey, 1994; Correias *et al.*, 1997). After local injection and vaporization, the stabilized droplet preparation might provide nearly the effectiveness of an equal volume of injected air, but would largely avoid the risk of embolism. For lithotripter shock waves, a small volume of SDS appeared to provide comparable results to a larger-volume ultrasound contrast agent, for example, 1% SDS compared to 10% contrast agent. If the droplet suspension was concentrated, rather than diluted to obtain the SDS (see Methods), effective concentrations might be achievable even for I. V. injection. The stabilized droplet suspension therefore appears to be a good candidate for a nucleation agent in nonthermal ultrasound therapy applications, such as tissue ablation and enhancement of tumor gene therapy.

#### ACKNOWLEDGMENTS

We thank Robert Greco, Unit for Laboratory Animal Medicine, for drawing the canine blood samples. Research supported by PHS Grant No. HL54201 awarded by the National Institutes of Health, DHHS.

AIUM (1992). *Acoustic Output Measurement and Labeling Standard for Diagnostic Ultrasound Equipment* (American Institute of Ultrasound in Medicine, Rockville, MD), pp. 1–122.  
 Bao, S., Thrall, B. D., Gies, R. A., and Miller, D. L. (1998). "In vivo transfection of melanoma cells by lithotripter shock waves," *Cancer Res.* **58**, 219–221.



- Brayman, A. A., and Miller, M. W. (1997). "Acoustic cavitation nuclei survive the apparent ultrasonic destruction of Alunex microspheres," *Ultrasound Med. Biol.* **23**, 793–796.
- Brayman, A. A., Azadniv, M., Cox, C., and Miller, M. W. (1996). "Hemolysis of Alunex-supplemented 40% hematocrit human erythrocytes *in vitro* by 1 MHz pulsed ultrasound: Acoustic pressure and pulse length dependence," *Ultrasound Med. Biol.* **22**, 927–938.
- Coleman, A. J., Kodama, T., Choi, M. J., Adams, T., and Saunders, J. E. (1995). "The cavitation threshold of human tissue exposed to 0.2 MHz pulsed ultrasound: Preliminary measurements based on a study of clinical lithotripsy," *Ultrasound Med. Biol.* **21**, 405–417.
- Correas, J., Kessler, D., Worah, D., and Quay, S. C. (1997). "The first phase shift ultrasound contrast agent: EchoGen," in *Ultrasound Contrast Agents*, edited by B. B. Goldberg (Dunitz, London), pp. 101–120.
- Dalecki, D., Raeman, C. H., Child, S. Z., Penny, D. P., Mayer, R., and Carstensen, E. L. (1997a). "The influence of contrast agents on hemorrhage produced by lithotripter fields," *Ultrasound Med. Biol.* **23**, 1435–1439.
- Dalecki, D., Raeman, C. H., Child, S. Z., Cox, C., Francis, C. W., Meltzer, R. S., and Carstensen, E. L. (1997b). "Hemolysis *in vivo* from exposure to pulsed ultrasound," *Ultrasound Med. Biol.* **23**, 307–313.
- Delius, M. (1994). "Medical applications and bioeffects of extracorporeal shock waves," *Shock Waves* **4**, 55–72.
- Delius, M. (1997). "Minimal static excess pressure minimizes the effect of extracorporeal shock waves on cells and reduces it on gallstones," *Ultrasound Med. Biol.* **23**, 611–617.
- Delius, M., Ueberle, F., and Gambihler, S. (1995). "Acoustic energy determines haemoglobin release from erythrocytes by extracorporeal shock waves *in vitro*," *Ultrasound Med. Biol.* **21**, 707–710.
- Fry, F. J., Kossoff, G., Eggleton, R. C., and Dunn, F. (1970). "Threshold ultrasonic dosages for structural changes in the mammalian brain," *J. Acoust. Soc. Am.* **48**, 1413–1417.
- Ivey, J. A., Gardner, E. A., Fowlkes, J. B., Rubin, J. M., and Carson, P. L. (1995). "Acoustic generation of intra-arterial contrast boluses," *Ultrasound Med. Biol.* **21**, 757–767.
- Kabalnov, A., Klein, D., Pelura, T., Schutt, E., and Weers, J. (1998). "Dissolution of multicomponent microbubbles in the bloodstream. I Theory," *Ultrasound Med. Biol.* **24**, 739–749.
- Kripfgans, O. D., Fowlkes, J. B., Miller, D. L., Eldevik, O. P., Frey, K. A., and Carson, P. L. (1999). "Ultrasonic-induced phase transitions of micrometer-size droplets," in *Collected Papers from the Joint Meeting "Berlin 99"* (Deutsche Gesellschaft für Akustik e. V., Oldenburg, Germany), pp. 1–4.
- Kripfgans, O. D., Fowlkes, J. B., Miller, D. L., Eldevik, O. P., and Carson, P. L. "Production of micrometer sized stabilized droplets and their ultrasonic induced phase transitions," *Ultrasound Med. Biol.* (in press).
- Mallinckrodt, Inc. (1998). Optison™ (package insert). (Mallinckrodt, Inc., St. Louis, MO), pp. 1–24.
- Mattrey, T. F. (1994). "The potential role of perfluorochemicals (PFCs) in diagnostic imaging," *Artif. Cells Blood Substit, Immobil. Biotechnol.* **22**, 295–313.
- Miller, D. L., and Thomas, R. M. (1995). "Thresholds for hemorrhages in mouse skin and intestine induced by lithotripter shock waves," *Ultrasound Med. Biol.* **21**, 249–257.
- Miller, D. L., and Thomas, R. M. (1996). "Contrast-agent gas bodies enhance hemolysis induced by lithotripter shock waves and high-intensity focused ultrasound in whole blood," *Ultrasound Med. Biol.* **22**, 1089–1095.
- Miller, D. L., and Gies, R. A. (1998a). "Enhancement of ultrasonically induced hemolysis by perfluorocarbon-based compared to air-based echo-contrast agents," *Ultrasound Med. Biol.* **24**, 285–292.
- Miller, D. L., and Gies, R. A. (1998b). "Gas-body-based contrast agent enhances vascular bioeffects of 1.09 MHz ultrasound on mouse intestine," *Ultrasound Med. Biol.* **24**, 1201–1208.
- Miller, D. L., and Gies, R. A. (1999). "Consequences of lithotripter shock-wave interaction with gas body contrast agent in mouse intestine," *J. Urol.* **162**, 606–609.
- Miller, D. L., Gies, R. A., and Chrisler, W. B. (1997). "Ultrasonically induced hemolysis at high cell and gas body concentrations in a thin-disc exposure chamber," *Ultrasound Med. Biol.* **23**, 625–633.
- Miller, D. L., Williams, A. R., Morris, J. E., and Chrisler, W. B. (1998). "Sonoporation of erythrocytes by lithotripter shockwaves *in vitro*," *Ultrasonics* **36**, 947–952.
- Poliachik, S. L., Chandler, W. L., Mourad, P. D., Bailey, M. R., Bloch, S., Cleveland, R. O., Kaczkowski, P., Keilman, G., Porter, T., and Crum, L. A. (1999). "Effect of high-intensity focused ultrasound on whole blood with and without microbubble contrast agent," *Ultrasound Med. Biol.* **25**, 991–998.
- Prat, F., Chapelon, J. Y., El Fadil, F. A., Theillere, T., Ponchon, T., and Cathignol, D. (1993). "In vivo effects of cavitation alone or in combination with chemotherapy in a peritoneal carcinomatosis in the rat," *Br. J. Cancer* **68**, 13–17.
- Vykhodtseva, N. I., Hynynen, K., and Damianou, C. (1994). "Pulse duration and peak intensity during focused ultrasound surgery: Theoretical and experimental effects in rabbit brain *in vivo*," *Ultrasound Med. Biol.* **20**, 987–1000.
- Walker, K. W., Pantely, G. A., and Sahn, D. J. (1997). "Ultrasound-mediated destruction of contrast agents, effect of ultrasound intensity, exposure, and frequency," *Invest. Radiol.* **32**, 728–734.
- Williams, A. R., Delius, M., Miller, D. L., and Schwarze, W. (1989). "Investigation of cavitation in flowing media by lithotripter shock waves both *in vitro* and *in vivo*," *Ultrasound Med. Biol.* **15**, 53–60.
- Williams, J. C., Stonehill, M. A., Colmenares, K., Evan, A. P., Andreoli, S. P., Cleveland, R. O., Bailey, M. R., Crum, L. A., and McAteer, J. A. (1999). "Effect of macroscopic air bubbles on cell lysis by shock wave lithotripsy *in vitro*," *Ultrasound Med. Biol.* **25**, 473–479.

# Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985–1996

G. Pavan

Centro Interdisciplinare di Bioacustica e Ricerche Ambientali (CIBRA), Università degli Studi di Pavia, Via Taramelli 24, 27100 Pavia, Italy and Department of Urban Science, IUAV, Venice, Italy

T. J. Hayward

Naval Research Laboratory, Washington, DC 20375

J. F. Borsani,<sup>a)</sup> M. Priano, M. Manghi, and C. Fossati

Centro Interdisciplinare di Bioacustica e Ricerche Ambientali (CIBRA), Università degli Studi di Pavia, Via Taramelli 24, 27100 Pavia, Italy

J. Gordon

Wildlife Conservation Research Unit, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom

(Received 19 July 1999; accepted for publication 7 March 2000)

A distinctive vocalization of the sperm whale, *Physeter macrocephalus* (= *P. catodon*), is the coda: a short click sequence with a distinctive stereotyped time pattern [Watkins and Schevill, *J. Acoust. Soc. Am.* **62**, 1485–1490 (1977)]. Coda repertoires have been found to vary both geographically and with group affiliation [Weilgart and Whitehead, *Behav. Ecol. Sociobiol.* **40**, 277–285 (1997)]. In this work, the click timings and repetition patterns of sperm whale codas recorded in the Mediterranean Sea are characterized statistically, and the context in which the codas occurred are also taken into consideration. A total of 138 codas were recorded in the central Mediterranean in the years 1985–1996 by several research groups using a number of different detection instruments, including stationary and towed hydrophones, sonobuoys and passive sonars. Nearly all (134) of the recorded codas share the same “3+1” (/// /) click pattern. Coda durations ranged from 456 to 1280 ms, with an average duration of 908 ms and a standard deviation of 176 ms. Most of the codas (a total of 117) belonged to 20 coda series. Each series was produced by an individual, in most cases by a mature male in a small group, and consisted of between 2 and 16 codas, emitted in one or more “bursts” of 1 to 13 codas spaced fairly regularly in time. The mean number of codas in a burst was 3.46, and the standard deviation was 2.65. The time interval ratios within a coda are parameterized by the coda duration and by the first two interclick intervals normalized by coda duration. These three parameters remained highly stable within each coda series, with coefficients of variation within the series averaging less than 5%. The interval ratios varied somewhat across the data sets, but were highly stable over 8 of the 11 data sets, which span 11 years and widely dispersed geographic locations. Somewhat different interval ratios were observed in the other three data sets; in one of these data sets, the variant codas were produced by a young whale. Two sets of presumed sperm whale codas recorded in 1996 had 5- and 6-click patterns; the observation of these new patterns suggests that sperm whale codas in the Mediterranean may have more variations than previously believed. © 2000 Acoustical Society of America. [S0001-4966(00)02706-5]

PACS numbers: 43.80.Ka, 43.80.Gx [WA]

## INTRODUCTION

The typical vocalization patterns of sperm whales, *Physeter macrocephalus* (= *P. catodon*), engaged in long, deep feeding dives are extended sequences of loud clicks produced at regular rates of approximately 0.5–2 clicks per second. Feeding sperm whales rarely produce clicks at the surface, usually falling silent a few minutes before reaching the surface and resuming clicking some 30–90 s after fluking up to commence a dive (Gordon *et al.*, 1992; Mullins *et al.*, 1988). Such vocalizations were termed “usual” clicking by

Watkins and Schevill (1977); we will also refer to these vocalizations as “regular” clicking. Occasionally, sperm whales also produce short click sequences with distinctive stereotyped time patterns. Watkins and Schevill (1977) first described hearing these patterned clicks during long regular click sequences, and because they usually came at the end of sequences, they named them “codas.”

Typically, only a small number of codas are heard in such situations. For example, Mullins *et al.* (1988) heard no codas from two separate males recorded off Nova Scotia, while Gordon *et al.* (1992) heard codas on only 2 out of 40 days of nearly continuous monitoring of the vocalizations of feeding mature males of Kaikoura, New Zealand. Higher rates of coda production, including quite vigorous coda ex-

<sup>a)</sup>Present address: Istituto Centrale per la Ricerca Applicata al Mare (ICRAM), via di Casalotti 300, 00166 Roma, Italy.

TABLE I. Year, data source, total duration of analyzed recordings, presence of codas, minimum number of animals present and storage medium analyzed for data sets in which sperm whale sounds were recorded. The storage medium refers to the medium in which recordings were made available for the present work. Recordings from the same source have been grouped even if they were made in different cruises or at different locations.

Year	Source	Duration	Codas observed?	Minimum number of whales recorded	Storage medium analyzed
1985	WHOI	cuts	yes	undetermined	data files
1988	Tethys	1 h 20 min	no	2	open reel
1989	Tethys	2 h 30 min	no	5	open reel
1990	Tethys	30 min	no	1	open reel
1991	CIBRA-Tethys	12 h 45 min	yes	4	DAT
1992	CIBRA-Tethys	7 h	yes	2	DAT
1993	CIBRA-Tethys	5 h	no	3	DAT
1994	CIBRA	16 h 20 min	yes	2	DAT
1994	IFAW	cuts	yes	undetermined	DAT
1995	CIBRA-Italian Navy	7 h 25 min	no	1	DAT
1995	Italian Navy	1 h	yes	7	DAT
1996	CIBRA	27 h	yes	10	DAT
1996	Saclant Center	3 h	yes	3	data files
1996	Italian Navy	1 h	no	4	DAT
1997	CIBRA	19 h 30 min	no	13	DAT

changes, are often heard from tight social groups of females and immature males interacting at the surface. Much of the more recent research on codas has been based mainly on analysis of recordings made in such situations (e.g., Moore *et al.*, 1993; Weilgart and Whitehead, 1993; Weilgart and Whitehead, 1997). Major studies of codas from the Galapagos and the Caribbean (Weilgart and Whitehead, 1993 and Moore *et al.*, 1993, respectively) have reported several distinctive codas varying from each other in both click number and pattern. The total number of distinctive codas reported from each area was remarkably similar: 23 from the Galapagos and 28 from the Caribbean.

Weilgart and Whitehead (1997) compared coda repertoires at several different levels: between social units encountered on different occasions within the same area; between “places” (areas of a few thousand kilometers across); between large geographical areas; and finally between oceans. They found strong group-specific dialects overlaid on weaker geographical variation. There were significant differences in repertoires between oceans (Watkins *et al.*, 1985; Weilgart and Whitehead, 1988; Weilgart, 1990; Moore *et al.*, 1993; Weilgart *et al.*, 1993; Borsani *et al.*, 1997; Pavan and Borsani, 1997; Weilgart and Whitehead, 1997; Whitehead *et al.*, 1998).

Within a broad research project on the acoustic behaviors of cetaceans in the Mediterranean Sea, more than 100 hours of recordings of sperm whale sounds, now archived in a Cetacean Sound Library (Priano *et al.*, 1997), were collected from multiple sources (see Table I) and analyzed. This ongoing research project is aimed at adding new data to our knowledge base concerning sperm whales and their acoustic behavior in the Mediterranean Sea in order to improve understanding of the population size, structure and trend in time and to support conservation efforts. Bioacoustic characteristics of particular interest in this research include the temporal patterns and waveform features of sperm whale vocalizations.

In this work, we characterize statistically the time patterns of clicks within sperm whale codas, and the repetition patterns of the codas themselves, in material recorded over a period of 12 years from widely dispersed locations in the Mediterranean Sea. A total of 138 codas, collected by six research groups, are analyzed to describe a coda pattern that may be typical of the Mediterranean. The results include interclick interval (ICI) measurements and their statistical distributions as well as coda repetition patterns. Interpulse interval (IPI) measurements are also performed on the component pulses of the clicks within the codas to estimate whale sizes and to aid in associating codas with individual whales.

Section I describes the instruments and methods of data collection by several research groups and catalogs the codas contained in the data sets. Section II describes the analysis of the coda data. Section III presents the results, including the context of coda observations and statistical characterizations of coda repetition patterns, series lengths, and temporal patterns of clicks. Section IV summarizes the results, interprets them in the context of previous research on sperm whale vocalizations and discusses their implications.

## I. DATA COLLECTION

Underwater recordings of sperm whale sounds were made in the central Mediterranean Sea by six research groups in the years 1985–1997 using a number of different detection instruments, including stationary and towed hydrophones as well as sonobuoys and passive sonars (see Fig. 1 and Table I).

The earliest recordings we analyzed were made in 1985 near Malta by W. Watkins of the Woods Hole Oceanographic Institution (WHOI) and colleagues (W. Watkins, personal communication). They collected eight codas from an undetermined number of animals; these codas were made avail-

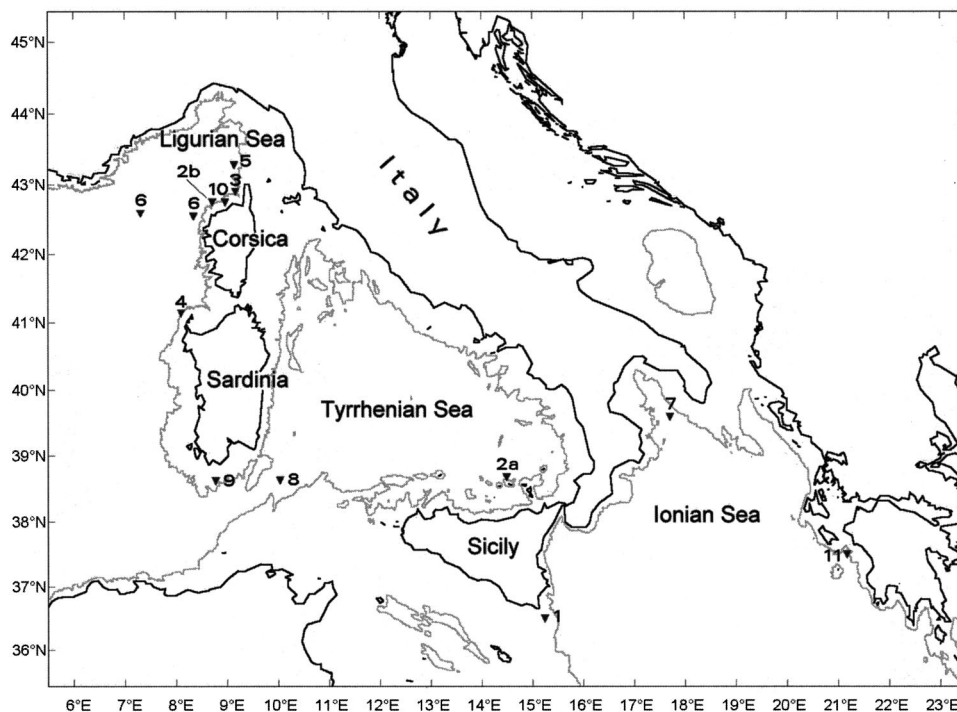


FIG. 1. Recording sites for 136 sperm whale codas (numbers according to Table II): (1) W. Watkins (WHOI), Malta 1985 (9 codas); (2a) CIBRA-Tethys, Aeolian Islands 1991 (9 codas); (2b) CIBRA-Tethys, Corsica 1991 (1 coda); (3) CIBRA-Tethys, Corsica 1992 (6 codas); (4) CIBRA, Sardinia 1994 (7 codas); (5) CIBRA, Corsica 1994 (7 codas); (6) J. Gordon (IFAW), Ligurian Sea 1994 (2 sites, 15 codas); (7) Italian Navy, Gulf of Taranto 1995 (2 codas); (8) Italian Navy-CIBRA, Sardinia 1995 (10 codas); (9) Italian Navy, Sardinia 1995 (8 codas); (10) CIBRA, Corsica 1996 (54 codas); and (11) Saclant Center, Greece (8 codas). The gray line is the 1000-m bathymetric contour.

able as data files belonging to the *SOUND* database of marine animal vocalizations developed at WHOI (Watkins *et al.*, 1992).

In July 1988 and August 1989, the Tethys Research Institute recorded sperm whale vocalizations each time animals were visually sighted; 1 h of recordings in 1988 and 3 h in 1989 were collected. Recordings were made using an ITC8073 preamplified hydrophone connected to an Uher 4400 Report Monitor tape recorder (tape speed  $19 \text{ cm s}^{-1}$ ).

In subsequent years (1991 to 1997), the research team of the Interdisciplinary Center for Bioacoustics and Environmental Research (CIBRA) of the University of Pavia, Italy, conducted auxiliary sailing vessel cruises in the Mediterranean Sea, gathering data on the distribution and the acoustic behavior of sperm whales and other cetacean species. The acoustic data were collected using wide-band towed arrays and digital audio tape (DAT) recorders.

Cruises in 1991, 1992 and 1993 were conducted within a joint research program between CIBRA and Tethys; cruises in 1994 were organized by CIBRA; cruises in 1995 were organized by CIBRA and the Italian Navy; cruises in 1996 and 1997 were organized by CIBRA and the association Ambiente Mare (Rome). In 1996, additional sperm whale codas were recorded by the NATO Saclant Center for Undersea Research (La Spezia, Italy).

As sperm whales are relatively difficult to find by visual search methods alone, acoustic techniques were used in the CIBRA cruises to detect, track and locate them. The procedure was to listen 24 h a day for at least 5 min every 30 min while towing an array of hydrophones at speeds of 7 to  $10 \text{ km h}^{-1}$ .

The towed array used to collect most of the recordings (those made during the years 1994–1996) consists of an oil-filled, acoustically transparent hose 12 m long and 6 cm in

diameter holding two groups of acoustic transducers spaced 8 m apart (Pavan and Borsani, 1997).

Upon detection, whales were acoustically tracked and approached by estimating their bearing through binaural audition of the hydrophone output signals and through computer-aided direction finding based on arrival-time differences of the clicks on the two transducer groups of the array. DAT recorders were started as soon as the signal-to-noise ratio was high enough to provide good-quality continuous recordings. In 600 h spent listening underwater, more than 100 h of recordings were made. A description of the instrumentation is provided by Pavan and Borsani (1997).

After the whales were approached, photographs of their flukes were taken whenever possible, resulting in the photo-identification of five individual whales on six occasions. One whale was sighted twice, in 1995 and 1996 (Pavan *et al.*, 1999).

In summer 1994, the research vessel *SONG OF THE WHALE* of the International Fund for Animal Welfare (IFAW) recorded 19 codas using similar equipment and procedures to those described above (Leaper *et al.*, 1992).

In 1995, several codas were recorded by the Italian Navy within a joint research project with the University of Pavia. Recordings were made with DIFAR (2.4-kHz bandwidth) and MISAR (10-kHz bandwidth) sonobuoys deployed by anti-submarine warfare (ASW) patrolling aircraft and by a submarine (Pavan, 1996; Pavan *et al.*, 1997a).

In 1996, the NATO Saclant Center for Undersea Research (La Spezia, Italy) recorded several sperm whales with the passive receiver of the LFAS (low frequency active sonar) system. Among the recordings, three sets of codas were found, two of them with variant click patterns (D'Amico, 1998).

TABLE II. Recordings in which sperm whale codas were found: Data set, research organization, location, month and year recorded, number of codas and number of coda series and coda bursts, number of whales emitting codas and whale IDs, estimated whale length, number of other whales detectable in the area, and acoustic receiver.

Data set	Source	Location	Date	Number of codas; series/bursts	Number of whales; ID/photo	Estimated whale length (m)	Number of other whales detectable	Receiver
1	WHOI	Malta	August 1985	9 n.d.	n.d. ...	n.d.	n.d.	Stationary hydrophones
2a	CIBRA-Tethys	Aeolian Islands	July 1991	9 3/3	1 SW0791	13.5	1	towed array
2b	CIBRA-Tethys	Corsica	September 1991	1 1/1	1 ...	n.d.	2	towed array
3	CIBRA-Tethys	Corsica	September 1992	6 2/2	1 ...	12	0	towed array
4	CIBRA	Sardinia	June 1994	7 1/1	1 SW0694	12.7	1	towed array ALENIA
5	CIBRA	Corsica	September 1994	7 2/2	1 SW0994	12.0	1	towed array ALENIA
6	IFAW	Ligurian Sea	Summer 1994	15 <sup>a</sup> 4/5	1 or 2	12.5 n.d.	n.d.	towed array
7	Italian Navy	Gulf of Taranto	May 1995	2 1/1	1 ...	13.2	1	wideband sonobuoy
8	Italian Navy	Sardinia	October 1995	10 1/2	1 ...	n.d.	2	DIFAR sonobuoy
9	Italian Navy	Sardinia	April 1995	8 1/1	1 ...	n.d.	n.d.	submarine
10	CIBRA	Corsica	September 1996	54 <sup>b</sup> 8/17	2 or more SW0996B SW0996C	9.3 12.3	>1	towed array ALENIA
11	Saclant Center	Greece	May 1996	8 n.d.	1 ...	n.d.	>3	LFAS passive

<sup>a</sup>Two or three more codas were partially masked by noise; coda onset times were estimated for two of these.

<sup>b</sup>An additional series of two codas was not recorded.

## II. DATA ANALYSIS

In all, more than 100 h of sperm whale recordings were analyzed to find codas and to determine, whenever possible, the context in which codas were produced. All of the data files and recordings were analyzed with the real-time Digital Signal Processing Workstation (DSPW) developed at CIBRA (Pavan, 1992, 1994; Pavan and Borsani, 1997). The browsing and analysis of this large amount of data were greatly facilitated by the real-time capabilities of the DSPW.

A total of 138 codas were identified; click timings were measured for 136 of them. Two codas in data set 6 were partially masked by incidental noise and therefore their click timings could not be measured; however, their coda onset times were estimated. Two other codas were heard in the 1996 cruise but were not recorded. Table II reports information on the recordings in which codas were found. Codas occurred in a wide geographic area (see Fig. 1).

For each coda, a spectrogram was computed and displayed. The time intervals between the first click and each subsequent click were measured on-screen with an accuracy of  $\pm 2$  ms by placing a cursor on the onset of each click [see Fig. 2(a)]. We also measured the intervals between consecutive codas emitted by the same whale by measuring the interval between the onset of the first click of each coda and the onset of the first click of the next coda.

Each click emitted by a sperm whale consists of multiple pulses, hypothetically associated with multiple internal re-

flections of the initial pulse within the spermaceti organ in the whale's head (Norris and Harvey, 1972; Clarke, 1978; Gordon, 1991; Goold, 1996) (see Fig. 3). The interpulse interval (IPI) is defined as the interval between the first and second pulses. Whenever possible, measurements of the IPIs of the clicks in the codas, and the IPIs of regular (noncoda) clicks preceding the codas, were made to aid in discriminating codas from different whales and to provide estimates of the length of the emitting animal (Clarke, 1978; Gordon, 1991; Goold, 1996; Pavan *et al.*, 1997b; Pavan *et al.*, 1999). The IPI measurements were based on cepstrum analysis of the time series and were made using the cepstrogram (cepstrum vs time) display available on the DSPW along with the spectrogram (spectrum vs time) display [see Fig. 2(b)]. The animal length estimates were then computed based on the empirical formula

$$L = 4.833 + 1.453\text{IPI} - 0.001\text{IPI}^2$$

from Gordon (1991) (see Table II). Although the correspondence between IPI and animal length has not been fully investigated, the IPI data provided information for discriminating clicks emitted by different animals when more than one animal was present.

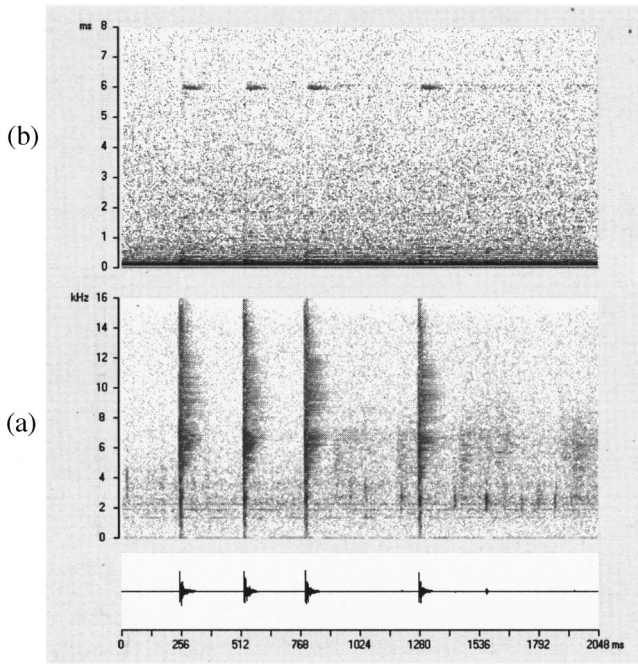


FIG. 2. (a) Spectrogram of a “3+1” (/// /) coda. (b) Corresponding cepstragram showing interpulse interval of approximately 6 ms (see Sec. III).

### III. RESULTS

#### A. Context of coda occurrences

Sperm whale acoustic contacts and sightings occurred at locations where the sea bottom depth ranged from 100 to 2600 m, though most of the sightings occurred where the bottom depth exceeded 1200 m. Codas were recorded at 12 sites over a wide area of the central Mediterranean (see Fig. 1).

Codas were heard and recorded only when two or more whales were acoustically detectable in the area in a period extending from 30 min before to 30 min after the codas occurred (see Table II). No codas were observed while recording isolated animals, i.e., while no other animals were detectable. During the CIBRA cruises, most of the codas observed were emitted by a whale at the end of its regular clicking bouts and just before it surfaced.

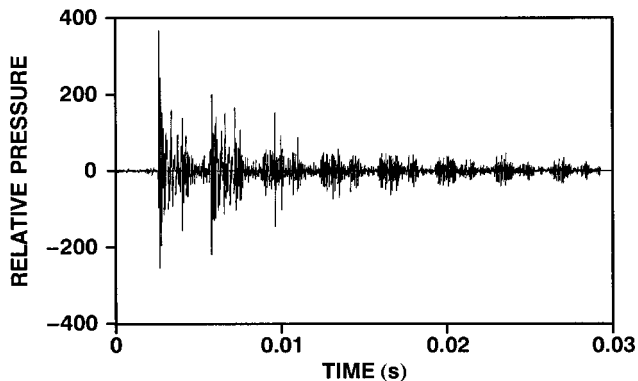


FIG. 3. Expanded time series plot of a single click from a sperm whale coda, showing the multipulse structure.

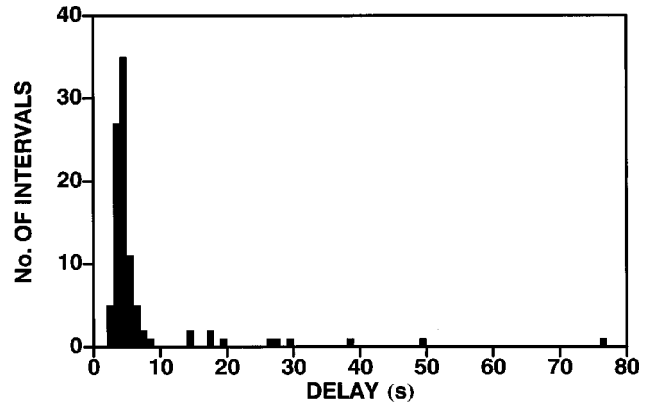


FIG. 4. Histogram of coda repetition intervals (CRIs) within the 20 series of two or more codas.

#### B. Coda repetition patterns and series lengths; coda bursts

Of the 138 codas, 17 were provided as individual data files; thus, no information about their patterns of production was available. Of the remaining 121 codas, 117 belonged to 20 series of 2 or more codas and 4 were isolated codas. In our statistical analyses, the isolated codas were treated as “series” of length 1. Each series consisted of up to 16 codas that were produced by the same whale. Within each series of length 2 or more, the coda repetition intervals (CRIs) were measured; for two consecutive codas, the CRI is defined as the interval between the onset of the first click of one coda and the onset of the first click of the next coda. In the 20 series of two or more codas, the 97 measured CRIs ranged from 2.76 to 77 s (mean 7.37 s, standard deviation 10.28 s). Figure 4 shows a histogram of the coda repetition intervals that occur within all of the coda series.

Within each of the 24 series, the codas occur in one or more “bursts” of fairly regularly spaced codas, separated by quiet intervals (see Fig. 5). A few of the bursts contain gaps

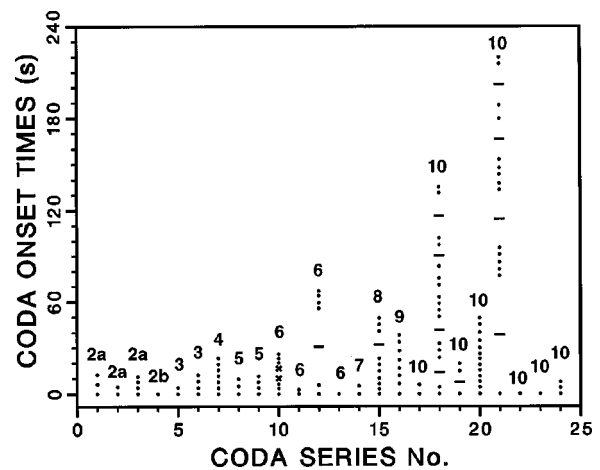


FIG. 5. Relative onset times (plotted as points) of the codas in each of 24 coda series (20 series of two or more codas and four isolated codas). In each series, codas occur in one or more “bursts” of fairly regularly spaced codas, separated by quiet intervals. The partition of the coda series into coda bursts is indicated by horizontal lines. The number above each series plot is the data set number from Table II. Estimated onset times of the two codas in data set 6 for which click intervals could not be measured are indicated by an “x.”

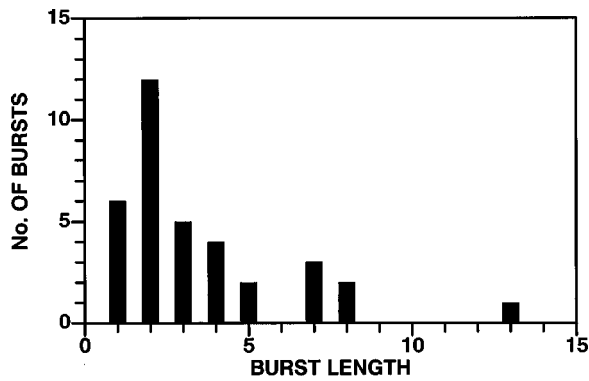


FIG. 6. Histogram of coda burst lengths (number of codas in each coda burst), taking 10 s as the threshold to separate bursts.

that suggest the omission of a coda followed by resumption of regularly spaced codas.

If we take 10 s of silence as the threshold to separate coda bursts, then there were 35 coda bursts, with between 1 and 13 codas in a burst. The partition of the coda series into coda bursts is indicated by horizontal lines in Fig. 5. The mean number of codas in a burst was 3.46, and the standard deviation was 2.65. Figure 6 shows a histogram of the number of codas in each coda burst.

Of the 35 coda bursts, 17 consisted of 3 or more codas, hence two or more coda repetition intervals. The average of the coefficients of variation of the CRIs within each of these 17 coda bursts was 0.114.

### C. Coda click patterns

Of the 136 codas that were measured, 134 match a “3+1” (/// /) click pattern. Two exceptional codas with embedded “3+1” patterns were found. In 1994, a coda in a series of seven had the 5-click pattern “4+1” (//// /). In 1995, a coda in a series of eight had the pattern “3+2” /// //. For each of these codas, the embedded pattern “3+1” was analyzed and included in the statistics.

Timings of the second, third, and fourth clicks (relative to the first click) were measured for each of the codas (see Fig. 7). The coda duration, defined as the time interval between the onsets of the first and last clicks, varied from 456 to 1280 ms. The mean duration was 908 ms, the standard

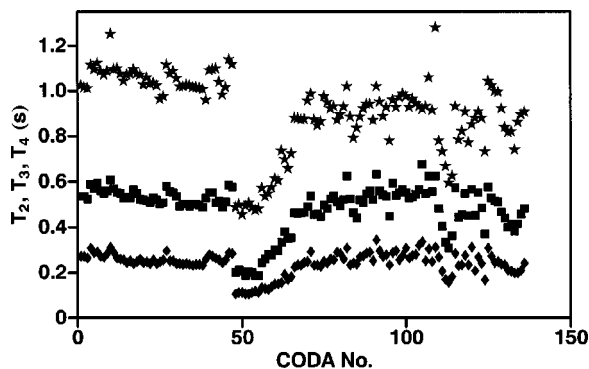


FIG. 7. Onset times (relative to first click onset) of the second ( $\blacklozenge$ ), third ( $\blacksquare$ ), and fourth ( $\star$ ) clicks in 136 “3+1” codas.

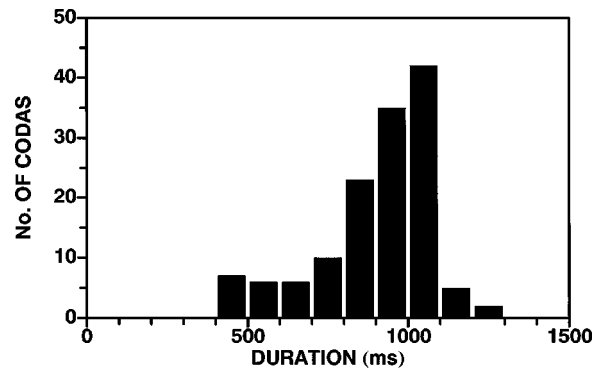


FIG. 8. Histogram of durations of 136 “3+1” codas.

deviation was 176 ms, and the coefficient of variation was 0.194. Figure 8 shows a histogram of the coda durations.

In a preliminary report (Pavan *et al.*, 1997a) based on a smaller sample (52 codas), it was shown that codas could be clearly divided into two different groups, “short codas” recorded in 1995, and “normal codas” recorded in previous years. With the increased data now available, the coda lengths have a statistical distribution that includes the intermediate values (see Figs. 7 and 8).

In some of the data files provided by the Saclant Center, short series of clicks, identified as probable sperm whale clicks, were found. The click patterns were “1+2+2” (see Fig. 9) and “1+2+3” (see Fig. 10). In the recordings, many animals appeared to be present and vocalizing at the same time at different bearings. This might be the first recording of a new coda pattern in the Mediterranean.

### D. Rhythmic patterns of the codas

The results in Sec. III C show a wide variation in the durations of the “3+1” codas (see Figs. 7 and 8). However, since the rhythmic patterns (interval ratios) of all the codas appear similar when perceived aurally, we analyzed these patterns by considering the interclick intervals (ICIs) normalized by coda duration.

The click intervals within a 4-click coda may be described by three parameters: the coda duration  $D$  and the normalized intervals  $\tau_1$  and  $\tau_2$ , which are the first and sec-

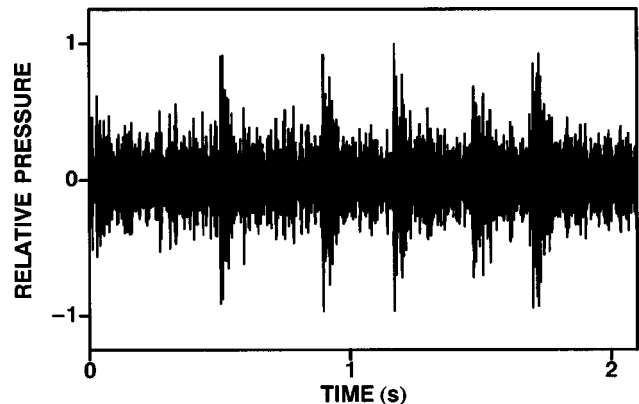


FIG. 9. Time series of a “1+2+2” click sequence.

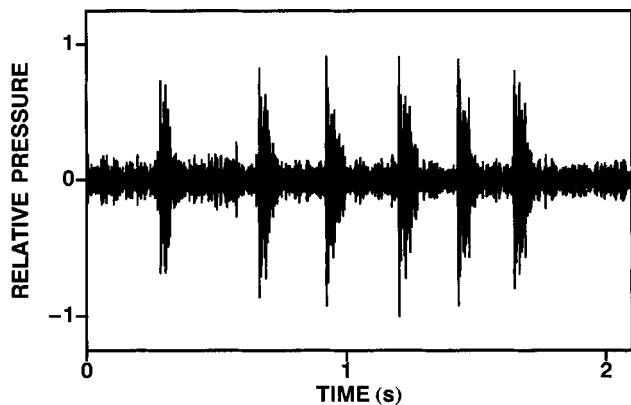


FIG. 10. Time series of a "1+2+3" click sequence.

ond interclick intervals divided by  $D$ . The parameters  $\tau_1$  and  $\tau_2$  completely determine the rhythmic pattern (interval ratios) of a coda, since the third normalized interclick interval is equal to  $1-\tau_1-\tau_2$ . Figure 11 shows a scatter plot of the pairs  $(\tau_1, \tau_2)$  for all 136 measured codas in data sets 1-11 and reveals a substantial variability in the coda rhythms.

Despite their overall variability, the coda rhythms are quite stable for data sets 1-8, which span a large variety of geographic locations and recording times (see Fig. 1 and Table II). Figure 12 shows a scatter plot of  $\tau_1, \tau_2$  for all codas in data sets 1-8 and displays the high degree of stability of the normalized intervals for these data sets. For these data sets, the mean values of  $\tau_1, \tau_2$ , and  $1-\tau_1-\tau_2$  were 0.248, 0.258 and 0.494, respectively; the ratios of these intervals are very close to the idealized 1:1:2 rhythmic pattern that is perceived aurally in the "3+1" coda. The coefficients of variation of  $\tau_1$  and  $\tau_2$  were 0.066 and 0.058, respectively, confirming the stability of the rhythmic pattern over these data sets.

Table III shows the means, standard deviations, and coefficients of variation of the three coda parameters  $\tau_1, \tau_2$ , and  $D$  for data sets 1-8 combined and for data sets 9, 10, and 11. The mean values of  $\tau_1$  and  $\tau_2$  for data set 10 differ from those for data sets 1-8 by 2.6 and 1.5 times the standard deviations of  $\tau_1$  and  $\tau_2$  for data sets 1-8, respectively. This

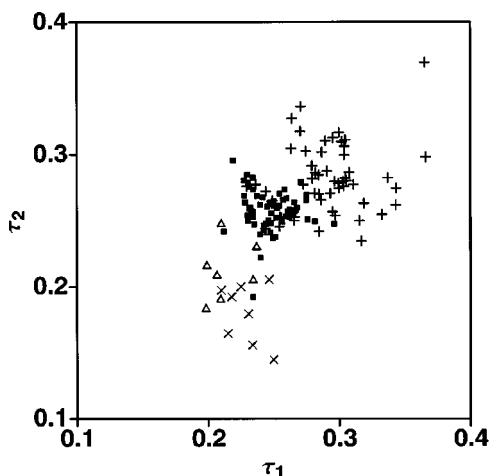


FIG. 11. Scatter plot of first and second normalized intervals  $(\tau_1, \tau_2)$  for 136 "3+1" codas: Data sets 1-8 (■), data set 9 (×), data set 10 (+) and data set 11 (△).

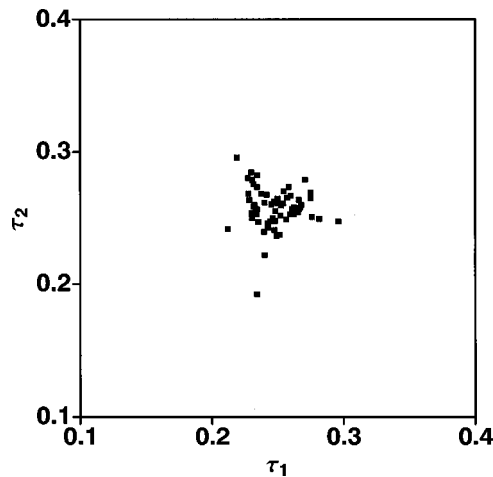


FIG. 12. Scatter plot of first and second normalized intervals  $(\tau_1, \tau_2)$  for the codas in data sets 1-8.

indicates a substantial deviation of the coda rhythms in data set 10 from those of data sets 1-8 (see also Fig. 11).

Most of the codas in data set 10 were produced by a relatively small whale whose length was estimated at 9.3 m, based on visual estimates and confirmed by a measured IPI of 3.07 ms (Pavan *et al.*, 1997c). According to Berzin (1971), this whale might be a 5-year-old male or a 7-year-old female. We also note that the codas produced by this whale were interspersed with single clicks and with pairs of clicks having the same interval as the first two clicks of the codas; the latter may have been incomplete codas or possibly 2-click codas.

The codas in data set 9 were unusually short and had highly variable durations. In addition, the signal-to-noise ratio for this data set was relatively low, and the click waveforms were noticeably distorted; this probably resulted from a long propagation path. These factors may have contributed, through measurement errors, to the untypical values of the normalized intervals for data set 9. The rhythmic patterns of the codas in data set 11 were intermediate between those of data sets 1-8 and those of data set 9 (see Fig. 11).

Table IV shows the correlation coefficients among the three coda parameters  $\tau_1, \tau_2$  and  $D$  for data sets 1-8 and for data set 10. (The sample sizes were too small to report correlation coefficients for data sets 9 and 11.) Within each of these two data sets, the normalized intervals  $\tau_1$  and  $\tau_2$  had low correlations with each other and with the coda duration  $D$ , consistent with random variations in the production of a stereotyped interval ratio pattern. In data sets 1-8, the varia-

TABLE III. Means ( $\mu$ ), standard deviations ( $\sigma$ ), and coefficients of variation ( $\sigma/\mu$ ) of the three coda parameters  $\tau_1, \tau_2$ , and  $D$  for data sets 1-8 ( $N=66$ ) and for data sets 9 ( $N=8$ ), 10 ( $N=54$ ), and 11 ( $N=8$ ).

Data set(s)	$\tau_1$			$\tau_2$			$D$		
	$\mu$	$\sigma$	$\sigma/\mu$	$\mu$	$\sigma$	$\sigma/\mu$	$\mu$	$\sigma$	$\sigma/\mu$
1-8	0.248	0.016	0.066	0.258	0.015	0.058	0.970	0.166	0.171
9	0.229	0.014	0.063	0.180	0.023	0.130	0.486	0.016	0.032
10	0.291	0.030	0.100	0.281	0.026	0.094	0.903	0.112	0.125
11	0.217	0.018	0.082	0.215	0.023	0.110	0.851	0.060	0.070



TABLE IV. Number of samples ( $N$ ) and correlation coefficients among the three coda parameters  $\tau_1$ ,  $\tau_2$ , and  $D$  for data sets 1–8, and data set 10. Correlation coefficients are not estimated for data sets 9 and 11 because of the small sample sizes.

Data set(s)	$N$	$\rho(\tau_1, \tau_2)$	$\rho(\tau_1, D)$	$\rho(\tau_2, D)$
1–8	66	-0.07	-0.02	0.25
9	8	...	...	...
10	54	0.21	-0.05	0.05
11	8	...	...	...

tions are those among 66 codas produced by nine or more individuals over a span of 11 years. In data set 10, the variations are those among 54 codas recorded on a single occasion; most of these were produced by one individual.

The coda parameters  $\tau_1$ ,  $\tau_2$ , and  $D$  were generally stable within each coda series. The averages (over series) of the coefficients of variation of  $\tau_1$ ,  $\tau_2$ , and  $D$  within each of the 20 series were 0.044, 0.047, and 0.039, respectively.

#### IV. SUMMARY AND DISCUSSION

A total of 138 sperm whale codas, recorded in the central Mediterranean Sea in the years 1985–1996, were analyzed to characterize statistically their series lengths, repetition patterns, and click interval patterns. Codas were heard and recorded only when two or more whales were acoustically detectable in the area, consistent with the hypothesis of a social function of sperm whale codas, as suggested by Watkins and Schevill (1977). Out of 121 codas found in continuous recordings, 117 belonged to coda series, each consisting of up to 16 codas. Each series consisted of one or more ‘bursts’ of between 1 and 13 codas spaced fairly regularly in time.

Nearly all of the recorded codas had the same ‘‘3+1’’ (/// /) time pattern. Although the coda duration varied, the ratios of the interclick intervals remained highly stable for 8 of the 11 data sets and corresponded closely to the 1:1:2 interval ratio pattern perceived aurally. This suggests that it is information in the pattern of clicks within the coda rather than the coda duration which is conserved in temporally and spatially distributed encounters. Normalized interclick intervals had low correlations with each other and with coda duration. Codas with interval ratios substantially different from the nominal 1:1:2 pattern, recorded in 1996, were associated with a young whale. A small number of codas with very short durations and variant interval ratios were recorded in 1995.

The small size of the coda repertoire reported here and the extent to which it is dominated by a single distinctive pattern (‘‘3+1’’) is notable. Even some of the few ‘‘aberrant’’ codas had a ‘‘3+1’’ pattern embedded within them and might be considered to be occasional variations of the standard pattern. This contrasts with the extensive repertoires reported from other areas. Although ‘‘3+1’’ codas were reported from the Pacific, they were not particularly common there, accounting for only 1.5% of the codas analyzed by Weilgart and Whitehead (1997); neither are they among the more frequently heard codas in the Azores (Gordon, unpublished data).

Moore *et al.* (1993) suggested that codas might allow the identification of an assemblage of whales sharing, at least temporarily, the same area, while Weilgart and Whitehead (1997) proposed that codas were learnt within matrilineal social groups, so that group dialects reflected familial relationships. They also found that repertoires varied regionally.

A number of factors might contribute to the predominance of a single coda in observations made over a long period of time within the Mediterranean:

- (a) The codas reported here were produced in the context of, and usually at the end of, sequences of ‘‘usual’’ clicks made during long feeding dives. Recordings analyzed by other teams have mostly been made from large socializing groups, often including apparent coda exchanges. Although similar in structure, codas produced in these two different situations may serve different functions, associated with different sized repertoires. Larger coda repertoires consisting of various 3-, 4-, 5- and 7-click codas have been observed recently in nursery groups in the Tyrrhenian and Ionian seas (Drouot and Gannier, 1999).
- (b) Most of the codas analyzed here were produced by mature males; the smaller animal represented in data set 10 may have been an immature male. Males may have smaller repertoires than females, possibly reflecting the fact that once they leave their maternal groups they are less social. It is perhaps informative that on one occasion when variant codas were recorded (data set 11) a large number of whales were reported on different bearings, indicating the possible presence of a mixed group of females and immature males.
- (c) The sperm whale population in the Mediterranean may be small and at least partially segregated from Atlantic populations.

Very little is known about the constituents of the sperm whale stock in the Mediterranean Sea (Notarbartolo *et al.*, 1993), the migration patterns within the basin and the extent of any movements across the Strait of Gibraltar, although data from several sources show that sperm whales regularly appear in the southern Tyrrhenian and Ligurian Seas in late spring, summer and early autumn. To understand more, acoustic techniques could be applied to study sperm whales more extensively in the Mediterranean Sea and to investigate their exchange with the Atlantic Ocean population as well as their movements in the eastern Mediterranean basin. Population-genetic analyses similar to those reported by Whitehead *et al.* (1998), along with analysis of coda repertoires throughout the Mediterranean and in adjacent Atlantic waters, might provide further evidence of a possible partial segregation of the Mediterranean population.

#### ACKNOWLEDGMENTS

This work was supported by the University of Pavia, the Italian Ministry of the Environment (Inspectorate for Sea Protection) and by the Italian Navy. The work of the second

author was supported by funds from the Office of Naval Research. We thank W. A. Watkins and the Woods Hole Oceanographic Institution for providing the recordings made in 1985; Tethys Research Institute for the cruises in 1991, 1992 and 1993; the International Fund for Animal Welfare for the recordings in 1994; the Italian Navy for the recordings in 1995; the NATO Saclant Center (La Spezia, Italy) for the recordings in 1996; the association Ambiente Mare for the cruises in 1996 and 1997; Giuseppe Notarbartolo di Sciara for his help in the earlier stages of the research; and Valeria Teloni for her work on the recordings made in 1996. Special thanks are due Alenia Eltag Sistemi Navali for the design and construction of the excellent array used in the cruises from 1994 to 1997.

- Berzin, A. A. (1971). "The sperm whale," Israel program for scientific translations, Jerusalem, p. 394, translated from Russian.
- Borsani, J. F., Pavan, G., Gordon, J. C. A., and Notarbartolo di Sciara, G. (1997). "Regional vocalizations of the Sperm Whale: Mediterranean codas," *Eur. Res. Cetaceans*, Cambridge (UK) **10**, 78–81.
- Clarke, M. R. (1978). "Structure and proportions of the spermaceti organ in the sperm whale," *J. Mar. Biol. Assoc. U.K.* **58**, 1–17.
- D'Amico, A. (ed.) (1998). "Summary Record and Report, SACLANTCEN Bioacoustics Panel, La Spezia, Italy, 15–17 June 1998," SACLANTCEN M133, Saclant Centre for Undersea Research, La Spezia, Italy.
- Drouot, V. and Gannier, A. (1999). "New sperm whale vocalizations recorded in the Mediterranean Sea," *Eur. Res. Cetaceans*, Cambridge (UK) **13**, 30–32.
- Goold, J. C. (1996). "Signal processing techniques for acoustic measurement of sperm whale body lengths," *J. Acoust. Soc. Am.* **100**, 3431–3441.
- Gordon, J. C. D. (1991). "Evaluation of a method for determining the length of sperm whales *Physeter catodon* from their vocalizations," *J. Zool., London* **224**, 310–314.
- Gordon, J., Leaper, R., Hartley, F.G. and Chappell, O. (1992). "Effects of whale-watching vessels on the surface and underwater acoustic behaviour of sperm whales off Kaikoura, New Zealand," Science and Research Series Publication, Department of Conservation, Wellington, New Zealand, Vol. 52, p. 64.
- Leaper, R., Chappell, O., and Gordon, J. C. D. (1992). "The development of practical techniques for surveying sperm whale populations acoustically," *Rep. Int. Whal. Comm.* **42**, 549–640.
- Moore, K. E., Watkins, W. A., and Tyack, P. L. (1993). "Pattern similarity in shared codas from sperm whales (*Physeter catodon*)," *Mar. Mammal Sci.* **9**, 1–9.
- Mullins, J., Whitehead, H., and Weilgart, L. (1988). "Behaviour and vocalizations of two single sperm whales *Physeter macrocephalus*, off Nova Scotia," *Can. J. Fish. Aquat. Sci.* **45**, 1736–1743.
- Norris, K. S., and Harvey, G. W. (1972). "A theory for the function of the spermaceti organ of the sperm whale (*Physeter catodon*)," in *Animal Orientation and Navigation*, edited by S. R. Galler, K. Schmidt-Koenig, G. J. Jacobs, and R. E. Belleville, NASA special publication SP-262, National Aeronautics and Space Administration, Washington, DC, pp. 397–417.
- Notarbartolo-di-Sciara, G., Venturino, M. C., Zanardelli, M., Bearzi, G., Borsani, J. F., and Cavalloni, B. (1993). "Cetaceans in the central Mediterranean Sea. Distribution and sighting frequencies," *Boll. Zool.* **60**, 131–138.
- Pavan, G. (1992). "A portable DSP workstation for real-time analysis of cetacean sounds in the field," *Eur. Res. Cetaceans*, Cambridge (UK) **6**, 165–169. The DSPW software is available on the web at ([www.unipv.it/cibra/instru.html](http://www.unipv.it/cibra/instru.html)).
- Pavan, G. (1994). "A digital signal processing workstation for bioacoustical research," *Atti 6 Conv. Ital. Ornit., Torino*, 1991, 227–234.
- Pavan, G. (1996). "Cetacei nel Mediterraneo. Nuove strategie per lo studio dei mammiferi marini con la collaborazione della Marina Militare," *Rivista Marittima*, Roma, Anno CXXIX n. 5, 117–122.
- Pavan, G., and Borsani, J. F. (1997). "Bioacoustic research on cetaceans in the Mediterranean Sea," *Mar. Fresh. Behav. Physiol.* **30**, 99–123.
- Pavan, G., Nascetti, D., Manghi, M., Priano, M., Fossati, C., and Borsani, J. F. (1997a). "Bioacoustic research on Sperm Whales in cooperation with the Italian Navy," *Eur. Res. Cetaceans* **10**, 82–86.
- Pavan, G., Priano, M., Manghi, M., and Fossati, C. (1997b). "Software tools for real-time IPI measurements on sperm whale sounds," *Proceedings Underwater Bio-Sonar and Bioacoustics Symposium. Proc. I.O.A.* **19** (Pt. 9), Loughborough, UK, 157–164.
- Pavan, G., Priano, M., Manghi, M., Fossati, C., and Bergamasco, C. (1997c). "Sperm whales *Physeter macrocephalus* L. off the north-west coast of Corsica, France, in summer 1996, acoustic and surface behaviours," *European Research on Cetaceans* **11**, 218–221.
- Pavan, G., Fossati, C., Manghi, M., and Priano, M. (1999). "Acoustic measure of body growth in a photo-identified sperm whale," *Eur. Res. Cetaceans*, Cambridge (UK) **12**, 254–258.
- Priano, M., Pavan, G., Manghi, M., and Fossati, C. (1997). "The cetacean sound library of the Interdisciplinary Centre for Bioacoustics and Environmental Research of the University of Pavia," *Proceedings Underwater Bio-Sonar and Bioacoustics Symposium, Proc. I.O.A.* **19** (Pt. 9), Loughborough, UK, 245–249. Samples of the recordings can be accessed on the web at ([www.unipv.it/cibra/sounds.html](http://www.unipv.it/cibra/sounds.html)).
- Watkins, W. A., and Schevill, W. E. (1977). "Sperm whale codas," *J. Acoust. Soc. Am.* **62**, 1485–1490 (with phonograph disk).
- Watkins, W. A., Moore, K. W., and Tyack, P. (1985). "Sperm whale acoustic behaviors in the southeast Caribbean," *Cetology* **49**, 15.
- Watkins, W. A., Fristrup, K. M., Daher, M. A., and Howald, T. (1992). "SOUND Database of Marine Animal Vocalizations; Structure and Operations," Technical Report WHOI-92-31, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts.
- Weilgart, L. S., and Whitehead, H. (1988). "Distinctive vocalizations from mature male sperm whales (*Physeter macrocephalus*)," *Can. J. Zool.* **66**, 1931–1937.
- Weilgart, L.S. (1990). "Vocalizations of the sperm whale, *Physeter macrocephalus*, off the Galapagos Islands as related to behavioural and circumstantial variables," Ph.D. thesis, Dalhousie University Halifax, Nova Scotia.
- Weilgart, L. S., and Whitehead, H. (1993). "Coda communication by sperm whales (*Physeter macrocephalus*) off the Galapagos Islands," *Can. J. Zool.* **71**, 744–752.
- Weilgart, L. S., and Whitehead, H. (1997). "Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whale," *Behav. Ecol. Sociobiol.* **40**, 277–285.
- Whitehead, H., Dillon, M., Dufault, S., and Weilgart, L. S. (1998). "Non-geographically based population structure of South Pacific sperm whales: Dialects, fluke-markings and genetics," *J. Anim. Ecol.* **67**, 253–262.

# Temporary shift in masked hearing thresholds of bottlenose dolphins, *Tursiops truncatus*, and white whales, *Delphinapterus leucas*, after exposure to intense tones

Carolyn E. Schlundt

Science Applications International Corporation, Maritime Services Division, 3990 Old Town Avenue, Suite 105A, San Diego, California 92110

James J. Finneran, Donald A. Carder, and Sam H. Ridgway

Marine Mammal Program, Space and Naval Warfare Systems Center, 49620 Beluga Road, Room 204, San Diego, California 92152-6266

(Received 21 July 1999; accepted for publication 2 February 2000)

A behavioral response paradigm was used to measure masked underwater hearing thresholds in five bottlenose dolphins and two white whales before and immediately after exposure to intense 1-s tones at 0.4, 3, 10, 20, and 75 kHz. The resulting levels of fatiguing stimuli necessary to induce 6 dB or larger masked temporary threshold shifts (MTTSs) were generally between 192 and 201 dB *re*: 1  $\mu$ Pa. The exceptions occurred at 75 kHz, where one dolphin exhibited an MTTS after exposure at 182 dB *re*: 1  $\mu$ Pa and the other dolphin did not show any shift after exposure to maximum levels of 193 dB *re*: 1  $\mu$ Pa, and at 0.4 kHz, where no subjects exhibited shifts at levels up to 193 dB *re*: 1  $\mu$ Pa. The shifts occurred most often at frequencies above the fatiguing stimulus. Dolphins began to exhibit altered behavior at levels of 178–193 dB *re*: 1  $\mu$ Pa and above; white whales displayed altered behavior at 180–196 dB *re*: 1  $\mu$ Pa and above. At the conclusion of the study all thresholds were at baseline values. These data confirm that cetaceans are susceptible to temporary threshold shifts (TTS) and that small levels of TTS may be fully recovered. [S0001-4966(00)00106-5]

PACS numbers: 43.80.Lb [WA]

## INTRODUCTION

The order Cetacea includes the Odontocetes (toothed whales) and the Mysticetes (baleen whales). These animals are marine mammals that are born and spend their entire lives in water. Many cetaceans have exceptional hearing and sound production capabilities (Johnson, 1967; Norris, 1968; Nachtigall, 1986; Watkins and Wartzok, 1985) and employ sound to forage, communicate, and navigate (Green *et al.*, 1994). There is growing concern that human-made sounds from sources such as maritime shipping, geophysical surveys, dredging, offshore drilling and exploration, oceanographic testing, and military testing are a potentially serious auditory problem for cetaceans (Richardson *et al.*, 1995); however, available data are inadequate to allow confident predictions regarding the parameters of sound that should be of concern (Green *et al.*, 1994).

Short exposure to intense sound may produce elevated hearing thresholds or threshold shifts (TSs). If the TS recovers after several minutes, hours, or days it is known as a temporary threshold shift (TTS). By contrast, a permanent threshold shift (PTS) is characterized by no recovery in hearing thresholds. TTS and PTS studies were instrumental in establishing noise exposure limits in humans. Currently, no published TTS or PTS data exist for cetaceans. Green *et al.* (1994) suggested that by measuring TTS in some cetaceans, it may be possible to safely determine maximum levels of sound to which these marine mammals can be exposed without risk of hearing damage. Richardson *et al.* (1995) identi-

fied studies of TTS in marine mammals as a “major data need.”

This report presents the results of a series of eight experiments designed to measure TTS in bottlenose dolphins (*Tursiops truncatus*) and white whales (*Delphinapterus leucas*) exposed to 1-s pure tones at frequencies between 0.4 and 75 kHz. The data presented here were collected between May 1996 and August 1998 and include and expand upon the original pure-tone TTS data for bottlenose dolphins provided by Ridgway *et al.* (1997) in a technical report. This technical report included TTS data for bottlenose dolphins exposed to 1-s tones at frequencies of 20, 75, and 3 kHz. These tests correspond to experiments I–III of the present report; however, the current study employed a different data analysis technique compared to Ridgway *et al.* (1997), thus some differences in results were obtained from the same raw data (see Sec. III).

The experimental methodology was briefly as follows: A behavioral response paradigm was used to measure masked underwater hearing thresholds in five bottlenose dolphins and two white whales before and immediately after exposure to intense 1-s tones. Thresholds measured before and after exposure were compared to determine if the subject experienced a TTS. Masking noise was used to provide a floor effect in the presence of variable ambient noise in San Diego Bay. Studies of TTS in humans (e.g., Humes, 1980; Parker *et al.*, 1976) have shown that the presence of masking noise results in elevated hearing thresholds and decreases the amount of TTS observed. The relationship between masking

TABLE I. Subject taxa, identification, and signalments.

Species	Animal identification	Gender	Weight (kg)	Length (cm)	Age (yrs)
<i>Tursiops truncatus</i>	APR	F	150	240	12–14
<i>Tursiops truncatus</i>	BEN	M	230	260	33–35
<i>Tursiops truncatus</i>	MUU	F	180	240	19–21
<i>Tursiops truncatus</i>	NEM	M	230	270	31–33
<i>Tursiops truncatus</i>	TOD	F	250	270	38–40
<i>Delphinapterus leucas</i>	MUK	F	550	350	29–31
<i>Delphinapterus leucas</i>	NOC	M	660	400	20–22

noise level and TTS in cetaceans is unknown. To indicate that the threshold shifts presented in this report were measured in the presence of masking noise, we use the term masked temporary threshold shift (MTTS) to identify these data.

## I. METHOD

### A. Subjects

The subjects for this study consisted of five *T. truncatus* and two *D. leucas* that ranged in age from 12 to 40 years (Table I). These species are members of the superfamily Delphinoidea, which includes more than half of all cetaceans (Ridgway, 1997). Figure 1 shows behavioral audiograms measured for *D. leucas* (Awbrey *et al.*, 1988; Johnson *et al.*, 1989; White *et al.*, 1978) and *T. truncatus* (Johnson, 1967). These species have hearing ranges and sensitivities equivalent to or better than many marine mammals (Fay, 1988; Richardson *et al.*, 1995) and may presumably be impacted by a wide range of anthropogenic acoustic devices.

The test animals were housed in netted enclosures (10 × 10 m to 13 × 25 m) in San Diego Bay, California. Individual diets consisted of a specific amount of herring, mackerel, capelin, smelt, and squid in order to maintain a healthy weight relative to the animal's age and gender. The amount of food delivered per day was the same regardless of perfor-

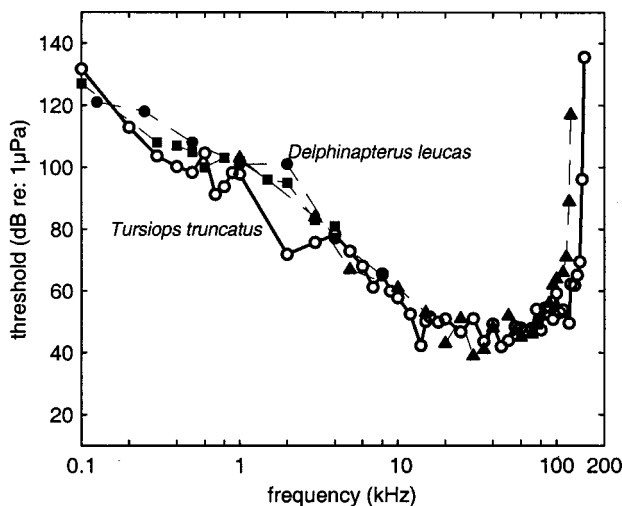


FIG. 1. Behavioral audiograms measured for the bottlenose dolphin, *Tursiops truncatus*. (open circles: Johnson, 1967) and white whale, *Delphinapterus leucas* (filled triangles: White *et al.*, 1978; filled squares: Johnson *et al.*, 1989; filled circles: Awbrey *et al.*, 1988).

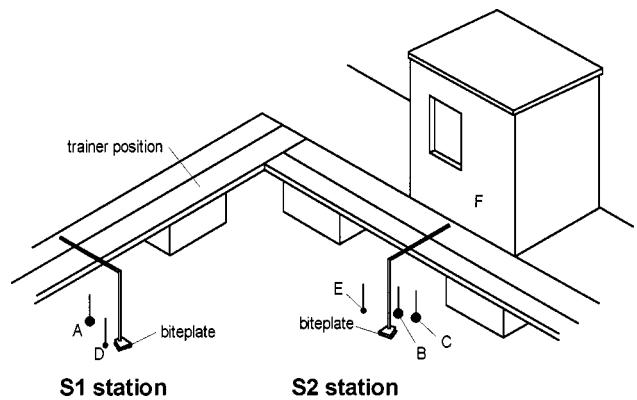


FIG. 2. The test enclosure showing the location of two underwater listening stations, S1 and S2. A—S1 projector, B—S2 projector, C—masking noise projector, D—S1 receiving hydrophone, E—S2 receiving hydrophone, F—equipment shelter.

mance or of the number of sessions in which an animal participated during the day. All animals were under constant veterinary supervision in accordance with applicable Federal regulations. The study followed a protocol approved by the Institutional Animal Care and Use Committee under guidelines of the Association for the Accreditation of Laboratory Animal Care. At the time of this study, all subjects were healthy and were not known to have any hearing loss at any of the fatiguing stimulus frequencies. One subject (MUK) had elevated baseline and preexposure thresholds at 40 kHz relative to those at 35 and 45 kHz, indicating a possible preexisting hearing loss at that frequency.

### B. Apparatus and sound presentation equipment

#### 1. Apparatus

Figure 2 illustrates the testing environment. The test enclosure, 10 × 10 m, similar in type to the “home” enclosures, was equipped with two underwater listening stations, designated “S1” and “S2.” Each station was constructed of polyvinyl chloride tubing and supported a biteplate on which the subjects stationed during testing. The stations were located at depths between 1 and 2.5 m, depending on the particular experiment. An underwater video camera was placed above the S1 station to allow the operator to observe the animal's position during low-visibility conditions. Each station was equipped with an underwater sound projector (Fig. 2; projectors A and B) and a receiving hydrophone (Fig. 2; hydrophones D and E). An additional sound source at the S2 station projected masking noise (Fig. 2; projector C). The S1 station was the site for the presentation of a “start” signal, or “S1 signal,” which was at times elevated for use as the fatiguing stimulus; the S2 station was the site for the hearing tests. The individual hearing test tones are referred to as “S2 tones.” The frequencies of the S1 signal and the S2 tones are referred to as the “S1 frequency” and “S2 frequency,” respectively.

#### 2. Test conditions

Table II lists the S1 and S2 frequencies, the masking noise level, and the sound projectors employed during each of the eight experiments. The parameters of the fatiguing

TABLE II. Combinations of S1 frequency and level, S2 frequency, masking noise level, and sound projectors used in the study. The last column indicates the number of bottlenose dolphins (*Ti*) and white whales (*Di*) tested at each condition.

Experiment	S1 frequency (kHz)	S1 projector	S1 levels (dB <i>re</i> : 1 $\mu$ Pa-rms)	S2 frequency (kHz)	S2 projector	Masking noise (dB <i>re</i> : 1 $\mu$ Pa <sup>2</sup> /Hz)	Masking noise projector	<i>n</i> ( <i>Ti</i> / <i>Di</i> )
I	20	ITC 1001	160–197	20, 30, 40	ITC 1032	75	EDC 6166	2/0
II	75	ITC 1042	160–194	75, 85, 100	LC-10	70	EDO 6166	2/0
III	3	ANSSQ62B	160–202	3, 4.5, 6	J9	67	ITC 1001	2/2
IV	10	ITC 1001	180–197	10, 15, 20	ITC 1032	63	J9	2/2
V	20	ITC 1001	180–201	20, 30, 40	ITC 1032	75	ITC 1042	2/2
VI	20	ITC 1001	178–202	30	ITC 1032	63	ITC 1042	2/2
VII	3	ITC 2015	180–201	4.5	ITC 1032	90	ITC 1032	2/2
VIII	0.4	XF4	179–193	0.6	ITC 1001	95	Custom	2/2

stimuli were chosen to be representative of certain sonar signals. Experiments I–IV were designed to measure MTTs at S1 frequencies of 20, 75, 3, and 10 kHz, respectively. Experiment V was designed to replicate experiment I with different subjects. Experiments VI and VII were designed to repeat the 20- and 3-kHz tests, respectively, using different masking noise levels. Experiment VIII featured a 0.4-kHz fatiguing stimulus.

This study was conducted over a 2.3-year period during which time equipment was maintained in an air-conditioned hut adjacent to the test enclosure (Fig. 2, F). During this time, some equipment was repaired and/or replaced, thus multiple pieces of equipment may be listed in Table II for the same task. The S1 signals, S2 tones, and the masking noise are described in more detail next.

*a. S1 signals.* All S1 signals were 1 s in duration with approximately 0.5-ms rise and fall times. The S1 signal amplitudes were 141 dB *re*: 1  $\mu$ Pa for baseline, preexposure, and recovery sessions (see Sec. IC). The amplitude of the first S1 signal presented during post-exposure sessions was elevated (i.e., was greater than 141 dB *re*: 1  $\mu$ Pa) and served as the fatiguing stimulus. The remaining post-exposure S1 signals were at 141 dB *re*: 1  $\mu$ Pa.

The S1 signals were generated using a programmable function generator (Wavetek model 164 or 178), filtered (Ithaco 4302), amplified (Crown Macro-Tech 2400, Hafler Pro 5000 or P7000, BGW Performance Series 4, or Instruments, Inc. LDV2-5 or S11-24), and used to drive a piezoelectric underwater sound projector (ITC 1001, ITC 1042, ITC 2015, ANSSQ62B, or XF4). The S1 sound projector was positioned directly in front of the subject at a distance of 1 to 2 m (depending on the experiment), at the same depth as the S1 biteplate. The amplitude of each S1 signal presented to the subjects was measured using a receiving hydrophone (B&K 8103) and charge amplifier (B&K 2635), digitized at 320 or 500 kSamples/s (Tucker-Davis AP2, Keithley-Metrabyte DAS-1800, or National Instruments PCI-MIO-16E-1), and saved for later analysis. The sensitivity of the B&K 8103 hydrophone was regularly checked with a B&K 4223 pistonphone. The sound pressure level (SPL) distance correction (from the hydrophone position to the approximate location of the animals' ears) was obtained from experimental measurements conducted prior to testing, without animals present.

At the higher S1 levels the amplifiers and sound projec-

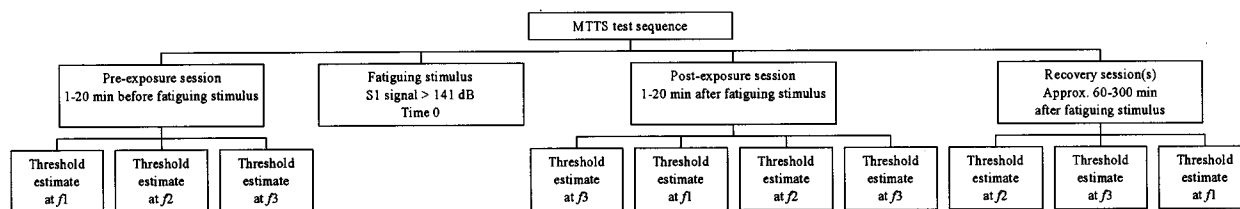
tors were often operating at near-capacity and the ability to specify the S1 level *a priori* decreased; however, recorded S1 amplitudes were generally within 3 dB of the intended level. It is important to note that each individual S1 was recorded so the exact S1 signal level was measured during each exposure. S1 levels cited in this report have units of dB *re*: 1  $\mu$ Pa [root mean square (rms)]; these are received levels at the estimated position of the animals' ears.

*b. S2 tones.* All S2 tones were 250 ms in duration, including 5-ms rise and fall times. Hearing thresholds were measured at three S2 test frequencies during experiments I–V and at a single S2 frequency for experiments VI–VIII (see Table II). When three S2 frequencies were used, the first was the same as the S1 frequency; the second and third were higher frequencies within one octave of the S1 frequency. If only a single S2 frequency was tested, it was at 1/2-octave above the S1 frequency. The use of multiple frequencies for hearing tests was based on human research showing that TTS may occur at frequencies above that of the fatiguing stimulus, often at a frequency 1/2-octave above (e.g., Dancer *et al.*, 1992; Green, 1976). The order in which multiple S2 frequencies were presented to the subject was counterbalanced between sessions and days.

The S2 tones were digitally generated and output through the D/A converter on a Tucker-Davis QDA2 or National Instruments PCI-MIO-16E-1 multifunction board. The D/A output was filtered (Ithaco 4302 and 4212), amplified (BGW Performance Series 2 or HP 467A), and used to drive an underwater sound projector (Fig. 2, Projector B). A number of different sound projectors were required to cover the range of S2 frequencies used; S2 projectors included piezoelectric spheres (ITC 1001, ITC 1032), piezoelectric cylinders (Celesco LC-10), and a moving-coil type (Chesapeake J9).

The S2 projector was positioned approximately 1.5 m in front of the animal for experiments I–VI and 1.5 m below the animal's ears, projecting upward, for experiments VII and VIII. A receiving hydrophone (B&K 8103, located at the estimated position of the animal's ears), charge amplifier (B&K 2635), and digital signal analyzer (HP 3561A) were used to measure the S2 tone amplitudes and calibrate the sound system prior to testing each day. During the actual hearing tests, this S2 receiving hydrophone was moved to a position above the S2 biteplate and was used to measure the

(a)



(b)

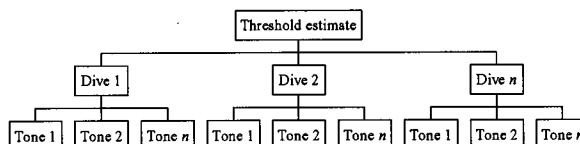


FIG. 3. A block diagram showing (a) the sequence of preexposure, post-exposure, and recovery sessions relative to the presentation of the fatiguing stimulus during MTTs testing when three S2 frequencies (designated  $f_1$ ,  $f_2$ ,  $f_3$ ) were tested, (b) the relationship between multiple dives and multiple S2 tone presentations that make up a threshold estimate (baseline and MTTs testing).

background noise (ambient as well as masking) and any sounds produced by the test subjects.

*c. Masking noise.* A noise-free test environment was not possible in San Diego Bay, therefore band-limited white noise (masking noise) was introduced to help create a floor-effect and to keep thresholds consistent despite variations in ambient noise. Masking noise was presented continuously during a session (i.e., was not turned “on” and “off” to delineate trials). If ambient noise levels temporarily exceeded the masking noise level, testing was suspended until ambient noise levels again fell below the masking noise.

Masking noise for experiments I and II was generated using a custom-built noise generator. Masking noise for the other experiments was generated using a Wavetek 132 function generator. Masking noise was amplified (HP 467A or BGW Performance Series 4) and output through an underwater sound projector (EdoWestern 6166, ITC 1001, ITC 1032, ITC 1042, Chesapeake J9, or a custom-built piezoelectric sphere). The masking noise projector was positioned approximately 1.5 m directly in front of the subject, at the same depth as the S2 biteplate. The noise spectral density was flat within  $\pm 5$  dB (experiments I–IV) or  $\pm 3$  dB (experiments V–VIII) over the range of S2 frequencies tested.

### C. Procedure

During each day of testing, subjects participated in one or more sessions. Each session belonged to one of four categories: (1) baseline, (2) preexposure, (3) post-exposure, and (4) recovery. Baseline sessions were conducted as part of “baseline testing” and took place at the beginning of each experiment over several days prior to presentation of the fatiguing stimulus. These sessions provided baseline information regarding hearing thresholds at each S2 frequency with the appropriate level of masking noise to which comparisons of future thresholds could be made. The latter three session

types were conducted as part of “MTTs testing” and took place within one test day either before or after presentation of a fatiguing stimulus. Preexposure sessions were conducted immediately before exposure to the fatiguing stimulus; post-exposure sessions began immediately after exposure; recovery sessions were conducted beyond one hour post exposure. Figure 3(a) illustrates the relationship between preexposure, post-exposure, and recovery sessions relative to the presentation of the fatiguing stimulus. Within each session, hearing thresholds were estimated at one or more S2 frequencies. Regardless of the session type, the hearing test procedure was essentially the same. The details of this procedure are described next, followed by more detailed descriptions of the four session types.

#### 1. Hearing test procedure

The hearing test procedure was based on the Method of Free Response, or MFR (Egan *et al.*, 1961). In this situation, the listener is presented with a number of brief tones during a relatively long observation period. The time interval between tones is randomized and the listener does not know when the next tone will occur. The listener is provided with a single “response key,” which he is instructed to press each time he hears a tone, and to do nothing otherwise.

The MFR was modified appropriately to allow its use with marine mammals. Subjects were trained to produce an audible “whistle” if they heard a tone and to remain quiet otherwise (Ridgway and Carder, 1997). The time interval between tones (i.e., measured from onset to onset) was randomized between 5 and 8 s and the animal did not know when the next tone would occur. Each threshold estimate was divided into several observation periods, called “dives”; each dive consisted of two parts: presentation of the S1 signal at the S1 station, followed by presentation of multiple S2 hearing test tones at the S2 station, as illustrated

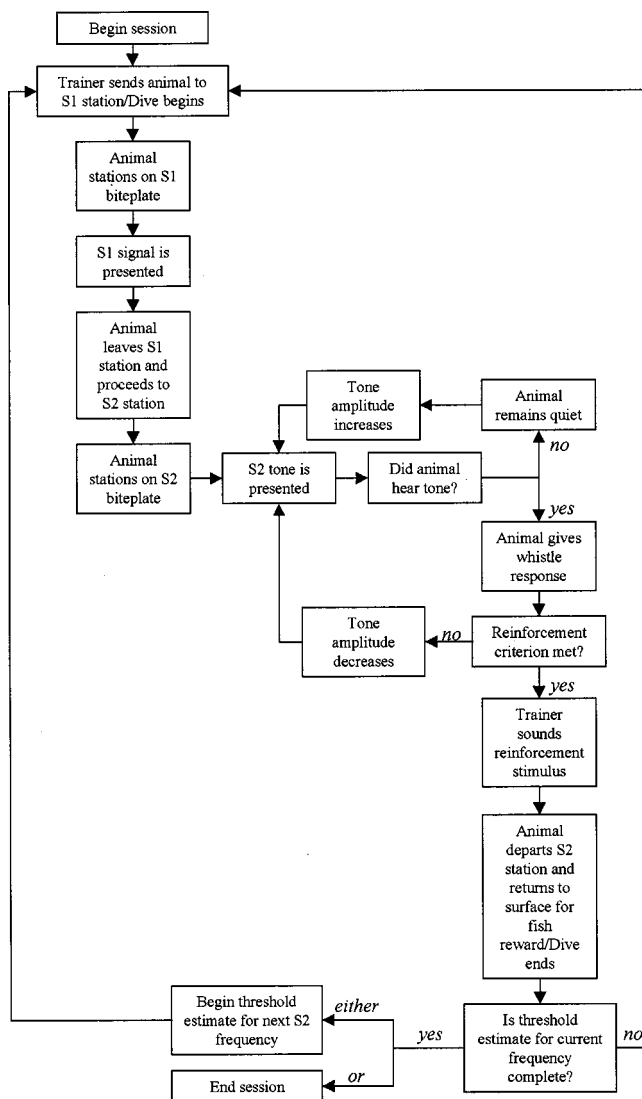


FIG. 4. A flowchart of the hearing test procedure.

in Fig. 3(b). The use of multiple dives was required to allow the animals to periodically surface for air and to receive fish reward.

Figure 4 shows the basic flow of the hearing test procedure. Each dive began with the trainer cueing the animal to the S1 station. The subject was trained to remain on the S1 station until presented with the S1 signal. The underwater camera was used to verify that the animal remained on the station until the signal was presented. Upon hearing the S1 signal, the subject proceeded directly to the S2 station for hearing tests. If the subject left the S1 station during the S1 signal, the dive continued without interruption (the distance moved during the 1-s S1 signal was considered to have a negligible effect on the received level); if the animal left the S1 station before the signal was projected the dive was restarted.

Once the animal was positioned at the S2 station, S2 tones were presented. The amplitudes of the S2 tones were adjusted using a modified staircase procedure (e.g., Cornsweet, 1962; Rosenberger, 1970). If the subject heard a tone and whistled (a hit), the amplitude of the next tone was decreased by 3 dB (experiment I) or 4 dB (experiments II–

VIII); if the subject did not hear a tone that was projected and remained quiet (a miss), the amplitude of the next tone was increased by 2 dB. A dive was terminated when the trainer sounded an underwater buzzer, a reinforcement stimulus, which signaled the animal to leave the S2 station and return to the surface for fish reward. Typically, the dolphins remained on the S2 station for an average of 12 S2 tone presentations. The whales remained on the S2 station for an average of 15 S2 tone presentations. Subjects were only reinforced after hits. Following reinforcement, the next dive was then begun, if necessary to complete a threshold estimate.

The time period between 0.05 and 2.05 s immediately following the onset of a tone was designated the “hit interval” and was included in the 5–8-s interstimulus interval measured from tone onset to onset. Only whistles occurring within the hit interval were recorded as hits. Any whistle response by a subject not occurring within a hit interval was recorded as a false alarm. Time catch trials (Ljungblad *et al.*, 1982), or no-tone periods, were randomly introduced while the animal was on the S2 station. No-tone periods ranged between 8 and 90 s.

## 2. Session types

At the start of each experiment listed in Table II, baseline hearing thresholds were measured to establish each subject’s masked hearing threshold at the relevant S2 frequencies and masking noise level over several days prior to any exposure to the fatiguing stimulus at the associated S1 frequency. Baseline sessions followed the hearing test procedure outlined above. All S1 signals were presented at 141 dB *re*: 1  $\mu$ Pa during baseline sessions.

After baseline thresholds had been established for each subject, MTTs testing began [Fig. 3(a)]. During MTTs testing, each subject was tested in the following sequence: pre-exposure session, post-exposure session, and recovery session (experiments V–VIII only). Preexposure sessions were conducted immediately (normally 1–20 min) before the subject’s exposure to the fatiguing stimulus. The procedure was identical to the baseline session procedure. All S1 signals were presented at 141 dB *re*: 1  $\mu$ Pa during preexposure sessions.

Post-exposure sessions were identical to baseline and preexposure sessions with the following exception: the amplitude of the first S1 signal presented during post-exposure sessions was elevated (i.e., was greater than 141 dB *re*: 1  $\mu$ Pa). This single, “loud” S1 signal served as the fatiguing stimulus to induce MTTs. The remaining S1 signals during post-exposure sessions were delivered at 141 dB *re*: 1  $\mu$ Pa. The post-exposure threshold at the first S2 frequency was generally obtained within 1–3 min following exposure to the fatiguing stimulus. When three S2 frequencies were tested, thresholds at the remaining frequencies were normally obtained within 20 min after exposure. After all three S2 frequencies had been tested, testing was repeated at the first S2 frequency [see Fig. 3(a)]. The order in which the frequencies were tested was counterbalanced between days, e.g., during experiment I the post-exposure frequencies may have been tested in the order 20, 30, 40, 20 kHz, 30, 40, 20, 30 kHz,

TABLE III. Information about MTTs testing including start dates, animals tested, the number of MTTs test sequences (S) run (i.e., fatiguing stimulus presentations), and the number of thresholds estimated (TE) in the corresponding preexposure, post-exposure, and recovery sessions over the 2.3-year duration of the study.

Animal	Date								Total S TE
	8/96 EXP I S TE	10/96 EXP II S TE	12/96 EXP III S TE	6/97 EXP IV S TE	9/97 EXP V S TE	1/98 EXP VI S TE	5/98 EXP VII S TE	6/98 EXP VIII S TE	
MUU	12 57								12 57
APR			8 44						8 44
TOD	15 101	11 82							26 183
BEN				5 29	8 82	5 38	5 40	6 48	29 237
NEM		12 93	9 53	6 35	5 44	7 50	5 39	3 25	47 339
MUK			8 46	6 33	9 91	6 48	5 39	3 24	37 281
NOC			7 40	5 30	8 81	7 56	5 39	4 32	36 278
Total	27 158	23 175	32 183	22 127	30 298	25 192	20 157	16 129	195 1419

and 40, 20, 30, 40 kHz on three successive days. If no MTTs was observed (see Sec. I 3), the S1 level was increased and testing continued the next day. S1 levels were increased by 10-dB steps through 180 dB *re*: 1  $\mu$ Pa, 6-dB steps through 192 dB *re*: 1  $\mu$ Pa, and 2- to 4-dB steps for S1 levels above 196 dB *re*: 1  $\mu$ Pa. If an MTTs was observed, testing was repeated at the same S1 level on subsequent days so that each S2 frequency could be tested immediately after exposure.

Auditory testing of animals using a behavioral response paradigm may be very time consuming, thus the number of S2 hearing test frequencies was limited. Recovery of MTTs over time (within a few minutes for small levels of MTTs) and declining motivation of the subjects after repeated testing practically limited the number of sessions that could be run in a single day. Therefore, during experiments I–IV, baseline thresholds were measured at all three S2 frequencies; however, preexposure thresholds were only measured for the first S2 frequency to be tested immediately after exposure. For those frequencies with no preexposure thresholds, post-exposure thresholds were compared to baseline values to determine if an MTTs had occurred. As proficiency in the test methodology increased, more sessions were conducted each day, e.g., during experiment V preexposure and post-exposure sessions were conducted at all three S2 frequencies.

Recovery sessions were conducted 60–300 min after exposure to the fatiguing stimulus to confirm that thresholds returned to (or stayed within) baseline levels within the same test day. The procedure was identical to that of baseline and preexposure sessions. Recovery sessions were only conducted during experiments V–VIII. Usually two recovery sessions were conducted at least one hour apart from each other.

### 3. Threshold estimation and MTTs criterion

Hearing thresholds were defined as the average SPL of the first ten hit-miss or miss-hit reversal points within the staircase data. A subject's threshold at a particular frequency could usually be estimated within 1 to 4 dives, or 20 to 30 total tones [Fig. 3(b)]. Note that this methodology differs from that previously used by Ridgway *et al.* (1997), who estimated thresholds as the average SPL of the five lowest

hits. Furthermore, Ridgway *et al.* quantized thresholds into 3-dB bins. Because of these methodological differences, some disagreement may be seen between the results presented here and the earlier data presented by Ridgway *et al.* (1997).

In some cases, it was not possible to obtain ten reversals from previously acquired post-exposure or recovery threshold data [e.g., results previously presented by Ridgway *et al.* (1997) were based on the mean SPL of the five lowest hits and so ten reversal points were not necessarily present]. Post-exposure or recovery threshold data with fewer than five reversal points were excluded from the current MTTs determination. Any MTTs arising from post-exposure or recovery thresholds based on 5–9 reversals are identified as such Sec. II.

Post-exposure and recovery thresholds were compared to preexposure (or baseline, when preexposure thresholds were not available) hearing thresholds to determine whether an animal experienced an MTTs, defined as a 6-dB (or larger) statistically significant increase. The 6-dB difference was considered the smallest shift that was clearly distinguishable from 3–4 dB variability in baseline thresholds measured day-to-day and session-to-session.

## 4. Behavioral alterations

In an attempt to document any behavioral reactions to the fatiguing stimuli, detailed notes were recorded regarding each subject's movements, swimming directions, and postures during each session. Each session was videotaped underwater and in-air. The time taken to swim from the S1 station to the S2 station was recorded for each dive. Any acoustic responses by the test subjects were noted as well.

## II. RESULTS

### A. Threshold shifts

During the 2.3-year test period over which the eight experiments were conducted, there were 195 MTTs test sequences (i.e., presentations of a fatiguing stimulus) and 1419 preexposure, post-exposure, and recovery threshold estimates from the seven test subjects. Table III provides a breakdown of the number of test sequences and the corre-



TABLE IV. S1 exposure levels at which MTTs was observed in dolphins and white whales as well as the magnitude of the shift, the S2 shift frequency, and the number of false alarms committed during threshold estimation at the shift frequency both pre- and post-exposure. Thresholds estimated before and immediately after exposure from which the magnitude of the MTTs was calculated were based on the ten reversals unless noted.

Exp.	S1 frequency (kHz)	Subject	S1 level (dB <i>re</i> : 1 $\mu$ Pa)	MTTS (dB)	S2 shift frequency (kHz)	No. of false alarms preexposure	No. of false alarms post-exposure		
I	20	TOD	193	8	40	0	1		
II	75	TOD	182	8	100	0	0		
III	3	NEM	194	7 <sup>a</sup>	3	0	0		
				16 <sup>b</sup>	4.5	0	0		
				17 <sup>c</sup>	6	0	0		
IV	10	BEN	192	7	15	0	0		
				MUK	192	7	20	0	0
				MUK	197	8	40	0	1
V	20	MUK	200	6	40	1	0		
				201	10	30	0	0	
				200	12	20	0	0	
				NOC	200	12	20	0	0
VI	20	BEN	196	6	30	1	1		

<sup>a</sup>7 reversals.

<sup>b</sup>9 reversals.

<sup>c</sup>8 reversals.

sponding threshold estimates for each animal by experiment, as well as the dates on which MTTs testing began for each experiment. Eleven MTTs were observed in six of the experiments. Table IV summarizes the MTTs observed during each of these experiments. At the conclusion of each experiment, all thresholds were within 3 dB of baseline values; there were no PTSs. Details regarding the levels of the S1 tones that produced the MTTs, the amount of MTTs, and the S2 frequencies at which the MTTs occurred follow.

During experiment I, an 8-dB MTTs at 40 kHz was measured in the oldest dolphin TOD after exposure to a 20-kHz, 1-s tone at 193 dB *re*: 1  $\mu$ Pa. This shift occurred 4 min post-exposure and had recovered (i.e., within 3 dB of baseline/preexposure threshold at that frequency) by the next time she was tested, the following day. During experiment II, TOD also showed an 8-dB MTTs at 100 kHz after exposure to a 75-kHz tone at 182 dB *re*: 1  $\mu$ Pa. This shift occurred 2 min post-exposure and recovered within 11 min. The shifts observed during experiments I and II occurred at frequencies above those of the fatiguing stimuli. No white whales were tested during experiments I and II.

During experiment III, the dolphin NEM had MTTs of 7, 16, and 17 dB at 3, 4.5, and 6 kHz, respectively, after exposure to a 3-kHz tone at 194 dB *re*: 1  $\mu$ Pa. These shifts occurred at 18, 10, and 5 min post-exposure, respectively. The whale NOC had a 12-dB shift at 4.5 kHz after exposure to a 3-kHz tone at 195 dB *re*: 1  $\mu$ Pa. This shift occurred 15 min post-exposure. Both the animals' thresholds had recovered by the next test day. Note that the post-exposure thresholds for NEM were based on fewer than ten reversals and must be interpreted with some caution. APR, the youngest dolphin tested, showed no MTTs at levels up to 201 dB.

Experiment IV produced MTTs of 7 dB after exposure to 10-kHz tones at 192 dB *re*: 1  $\mu$ Pa in both the dolphin BEN (15-kHz shift frequency, 4 min post-exposure) and the whale MUK (20-kHz shift frequency, 13 min post-exposure).

BEN's threshold recovered within 16 min, MUK's threshold had recovered by the next time she was tested, the following day.

During experiment V, the whale MUK had an MTTs of 8 dB at 40 kHz after exposure to a 20-kHz tone at 197 dB *re*: 1  $\mu$ Pa. This shift occurred 15 min post-exposure and had fully recovered by the next time she was tested, five days later. MUK also had MTTs of 6 dB at 40 kHz (2 min post-exposure, recovered within 6 min), and 10 dB at 30 kHz (2 min post-exposure, recovered within 17 min) after exposure to 20-kHz tones at 200 and 201 dB *re*: 1  $\mu$ Pa, respectively. The whale NOC also had an MTTs of 7 dB at 20 kHz 2 min after exposure to a 20-kHz tone at 200 dB *re*: 1  $\mu$ Pa. This shift increased to 9 dB 5 min after exposure, and to 12 dB 16 min post-exposure. NOC's threshold recovered to within 5 dB by the last recovery session of the day and was fully recovered by the next opportunity for testing three days later.

The dolphin BEN had a 6-dB MTTs at 30 kHz during experiment VI after exposure to a 20-kHz tone at 196 dB *re*: 1  $\mu$ Pa. This shift occurred 2 min post-exposure and recovered within 5 min. No threshold shifts were observed during experiments VII (3 kHz S1, 90 dB *re*: 1  $\mu$ Pa<sup>2</sup>/Hz masking noise, maximum S1 of 201 dB *re*: 1  $\mu$ Pa) or experiment VIII (0.4 kHz S1, 95 dB *re*: 1  $\mu$ Pa<sup>2</sup>/Hz masking noise, maximum S1 of 193 dB *re*: 1  $\mu$ Pa).

Table IV also provides the number of false alarms that occurred during the pre- and post-exposure thresholds estimated at the S2 frequency where MTTs was observed. The numbers are representative of the very low false alarm occurrences seen throughout baseline and MTTs test sessions. Generally, each threshold estimate (10 reversals) for a given S2 frequency was complete within an average of 30 S2 tone presentations. Usually two or fewer false alarms were committed during any threshold estimate. A total of 195 MTTs test sessions were conducted over the eight experiments.

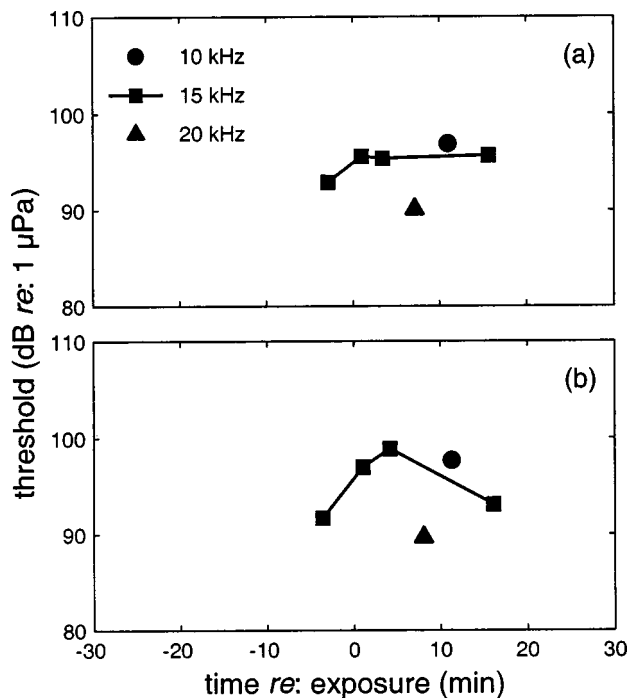


FIG. 5. Hearing thresholds measured from the dolphin BEN during experiment IV before and after exposure to 10-kHz tones at (a) 180 dB *re*: 1  $\mu$ Pa (no MTTs) and (b) 192 dB *re*: 1  $\mu$ Pa (7-dB MTTs at 15 kHz). Baseline thresholds were 97, 93, and 87 dB *re*: 1  $\mu$ Pa at 10, 15, and 20 kHz, respectively.

Eleven of those sessions produced MTTs (Table IV). In those 11 sessions, a total of 99 thresholds were estimated in preexposure, post-exposure, and recovery sessions combined. Three or more false alarms were committed in only 11 of 99 threshold estimates. Eight thresholds during which three or more false alarms were committed occurred during experiment V (BEN—three false alarms at 20- and 40-kHz preexposure to 181 dB; MUK—three false alarms at 30 kHz pre- and post-exposure to 200 dB, and at 30 kHz recovery to 201 dB; NOC—four false alarms at 40 kHz preexposure and three false alarms at 30 and 40 kHz post-exposure to 197 dB). The remaining three thresholds with three or greater false alarms occurred during experiment VIII (BEN—seven false alarms during preexposure at 600 Hz and three false alarms in two recovery sessions to 191 dB).

Figures 5 and 6 are some representative plots of masked hearing thresholds estimated during MTTs testing presented as a function of time relative to the fatiguing stimulus (time 0). Figures 5(a) and (b) show masked hearing thresholds measured from the dolphin BEN during experiment IV, before and after exposure to 10-kHz S1s at 180 and 192 dB *re*: 1  $\mu$ Pa, respectively. The abscissa is time relative to the exposure and the ordinate is the measured threshold. Baseline masked thresholds were 97, 93, and 87 dB *re*: 1  $\mu$ Pa at 10, 15, and 20 kHz, respectively. Figure 5(a) shows essentially flat thresholds at 15 kHz despite the fatiguing stimulus exposure at time zero. Figure 5(b) shows a 7-dB MTTs approximately 4 min after exposure and the eventual recovery, within approximately 16 min. For both Fig. 5(a) and (b), preexposure thresholds at 15 kHz and post-exposure thresholds at 10 and 20 kHz were within 3 dB of the baseline

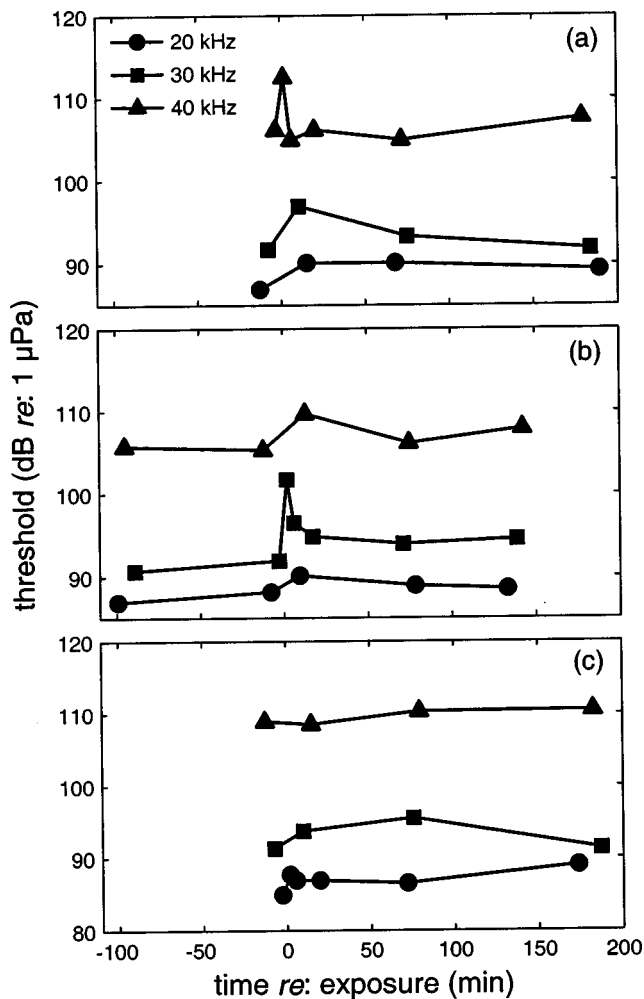


FIG. 6. Preexposure, post-exposure, and recovery thresholds measured for the whale MUK during experiment V. Exposures were to 20 kHz S1s at (a) 200 dB *re*: 1  $\mu$ Pa (6-dB MTTs at 40 kHz), (b) 201 dB *re*: 1  $\mu$ Pa (1-dB shift at 30 kHz), and (c) 200 dB *re*: 1  $\mu$ Pa (no MTTs). Baseline thresholds were 87, 93, and 109 dB *re*: 1  $\mu$ Pa at 20, 30, and 40 kHz, respectively.

values. Note that the data for the S1 at 180 dB *re*: 1  $\mu$ Pa were chosen for comparison because this was the other occasion when 15-kHz thresholds were measured immediately before and after exposure with this subject during experiment IV (see Sec. IC 2).

Figure 6 shows preexposure, post-exposure, and recovery thresholds measured for the whale MUK during experiment V. As in Fig. 5, the abscissa is time relative to the exposure and the ordinate is the measure threshold. Figures 6(a), (b), and (c) show thresholds measured before and after exposure to 20-kHz S1s at 200, 201, and 200 dB *re*: 1  $\mu$ Pa, on three test days (the intended S1 level was 200 dB *re*: 1  $\mu$ Pa). Baseline masked thresholds were 87, 93, and 109 dB *re*: 1  $\mu$ Pa at 20, 30, and 40 kHz, respectively. Figure 6(a) shows a 6-dB MTTs at 40 kHz and Fig. 6(b) shows a 10-dB shift at 30 kHz. Both shifts were measured within 2 min after exposure and recovered within approximately 6–17 min. Recovery thresholds were within 3 dB of preexposure thresholds at all S2 frequencies. No MTTs is seen in Fig. 6(c)—all thresholds were within 4 dB of preexposure thresholds.

Figure 6 illustrates the sequence in which the S2 frequencies were tested: in Fig. 6(a), preexposure thresholds

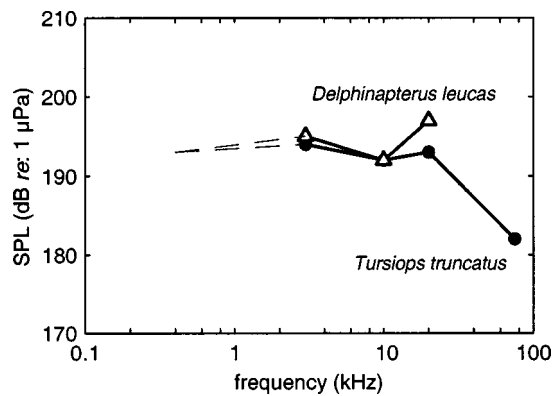


FIG. 7. Minimum fatiguing stimulus levels required to induce MTTs at each frequency, regardless of shift frequency or masking noise level. No MTTs was observed at 0.4 kHz at the highest level achieved, 193 dB *re*: 1  $\mu$ Pa.

were measured in the order 20, 30, 40 kHz, then post-exposure thresholds were measured in the order 40, 40, 30, 20, 40 kHz, so that 40 kHz was tested immediately before and after exposure. Similarly, Figs. 6(b) and (c) show thresholds at 30 and 20 kHz, respectively, measured immediately before and after exposure. It is interesting to note that MTTs (according to our criteria) were only observed in the thresholds measured immediately after exposure, while smaller (<6 dB) TSs appear to occur at some of the frequencies tested later [e.g., 30 kHz in Fig. 6(a) and 40 kHz in Fig. 6(b)]. Higher TSs may have been seen if these frequencies were tested sooner; however, there appear to be no TSs at 20 kHz even when it was tested immediately after exposure. Also of note is the 6-dB threshold shift at 40 kHz, compared to 10 dB at 30 kHz, and no shift at 20 kHz, even though MUK appeared to have elevated baseline and preexposure hearing thresholds at 40 kHz.

Figure 7 shows the minimum fatiguing stimulus levels that produced MTTs at each of the S1 frequencies for both *Tursiops* (filled circles) and *Delphinapterus* (open triangles). These data are pooled at each S1 frequency regardless of the shift frequency or masking noise level. The dashed line indicates that no MTTs were observed after exposure to the 0.4-kHz fatiguing stimulus at the highest level achieved, 193 dB *re*: 1  $\mu$ Pa.

## B. Behavioral reactions

As the S1 level gradually increased each day during MTTs testing, animals began to exhibit alterations in their trained behaviors. In this report we define a behavioral alteration as a deviation from an animal's trained behaviors as a result of exposure to intense sound. This is in contrast to field observations, where the reaction of naïve animals to novel stimuli is often difficult to interpret; a behavioral reaction in these circumstances may occur at levels corresponding to the animal's detection of the sound, rather than a level which may produce TTS (Green *et al.*, 1994). In the present study, changes began at levels below those that induced MTTs and continued as S1 signals increased to MTTs levels. Of the 195 MTTs test sequences, behavioral alterations were recorded for 129 of them.

Some of the most dramatic behavioral alterations were observed in the seconds immediately following exposure to the louder S1 signal and can be described as a general disorientation. The calm, smooth departure from the S1 station in the direction of the S2 station was replaced with an abrupt, quick departure from S1 usually in the direction opposite S2. Dolphins sometimes swam erratically around the test enclosure and performed tail slaps and jaw pops (Connor and Smolker, 1996; Finneran *et al.*, 2000; Smolker and Richards, 1988). Whales quickly backed off the S1 station and also circled the enclosure, often vocalizing as they swam. The male subjects often replied to the S1 with series of pulse bursts. In a more aggressive display, one dolphin (NEM at 3 kHz, 194 dB *re*: 1  $\mu$ Pa) attacked the S1 station after hearing the louder S1 signal. Several animals seemed not to recognize the loudest S1 signals as the tonal cue to go to S2 and, after having swum around the enclosure, voluntarily returned to the S1 station and required a second S1 signal (at 141 dB *re*: 1  $\mu$ Pa) before they would proceed to the S2 station.

The most apparent consequence of these behavioral alterations was an increase in the travel time between the S1 and S2 stations. We looked at a random sample of 15 S1–S2 time intervals over a minimum of three test days for each animal during each experiment ( $n=420$ ) to determine the average time it took to swim between the S1 and S2 stations during baseline testing. The average S1–S2 time was 14 s (*s.d.* 2.4). We looked at the S1–S2 travel time in the dive during which the fatiguing stimulus was presented for 125 of the MTTs sequences in which altered behavior was observed (four of the times were not available). At the higher S1 levels as animals began to deviate from their trained behaviors the average S1–S2 travel time increased to 28 s (*s.d.* 42.4). A linear regression was performed on the S1–S2 times sampled during baseline and immediately following the fatiguing stimulus in the altered behavior sequences during MTTs testing ( $n=545$ ). S1–S2 time intervals were significantly higher immediately after exposure to the louder S1 levels than during baseline testing [ $F(1,544)=45.89$ ,  $p<0.001$ ]. Figure 8 shows the mean S1–S2 time interval for each animal over the eight experiments both during baseline and during the dive immediately following the fatiguing stimulus for the MTTs test sequences in which altered behavior was observed.

A fairly common example of a behavioral alteration was the reluctance to return to the S1 station for dives following the fatiguing stimulus. This apparent reluctance was exhibited in several ways: (1) an animal might ignore the trainer's cue to go to the S1 station and instead proceed directly to the S2 station for hearing tests; (2) an animal might not station properly on the S1 biteplate (e.g., biting only the outermost edge or at an angle); or (3) an animal might leave the S1 station early before actually hearing the S1 signal.

Another common behavioral alteration was that of "breaking station" or leaving the S2 station prematurely (not under the direction of a trainer). Breaking station after a louder S1 signal usually occurred after only one or two trials and was suspected to be caused by a hypersensitivity to tones following the louder signal. After several instances in which animals broke station at S2 following louder S1 signals, we

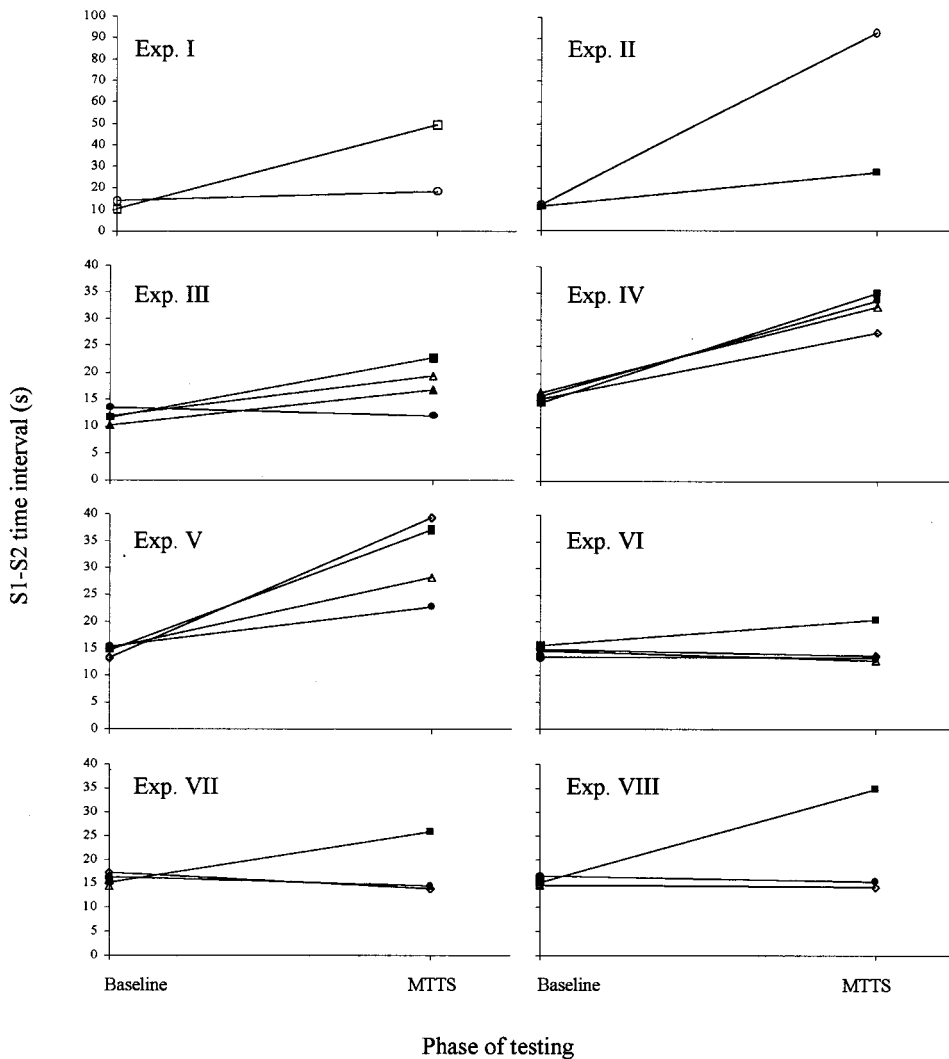


FIG. 8. The mean S1–S2 time interval for each animal over the eight experiments both during baseline and during the dive in which the fatiguing stimulus was presented for MTTs test sequences in which altered behavior was observed (open square: MUU; open circle: TOD; closed square: NEM; open triangle: MUK; closed triangle: APR; closed circle: NOC; open diamond: BEN).

began initiating S2 tone presentation at levels 4–6 dB below threshold for trials immediately following the louder S1. Instances of breaking station at S2 were greatly diminished. Further support for a hypersensitivity effect as a result of exposure to louder signals was evidenced by several cases of negative TTS, i.e., lower post-exposure thresholds (Hodge and McCommons, 1966; Marshall and Brandt, 1974; Silbiger, 1965). It has been suggested that negative TTS may be a good predictor for levels that are approaching those that elicit TTS. In this study, both whales and dolphins often experienced slight increases in hearing sensitivity, decreases in threshold, following louder S1s. On most occasions, the improvement in sensitivity was still within the “normal” threshold variation of 3–4 dB. On one occasion the whale NOC had a 7-dB increase in hearing sensitivity at 6 kHz after exposure to a 3-kHz tone at 193 dB *re*: 1  $\mu$ Pa. This negative shift occurred 2 min post-exposure and recovered within 14 min.

Figure 9 shows the minimum fatiguing stimulus levels at which the dolphins and white whales began to display altered behavior at each S1 frequency. These levels ranged from 178–193 dB *re*: 1  $\mu$ Pa. In general, the white whales tolerated louder levels better than the dolphins and were more subdued in their behavioral reactions.

### III. DISCUSSION

This report presents the first direct evidence of TTS in odontocetes. Using 1-s signals between 3 and 75 kHz, MTTs between 6 and 17 dB were induced in dolphins and white whales. Because of the limited amount of data and intersubject variability, pooling the data (regardless of shift

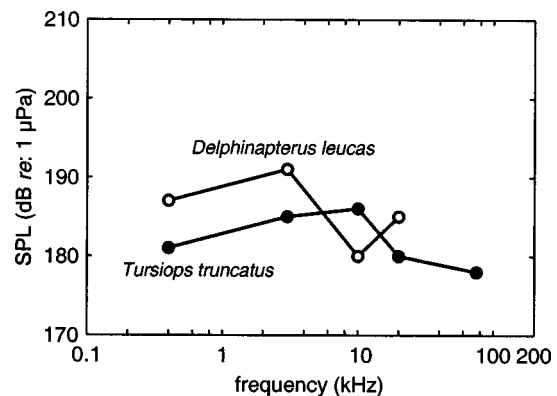


FIG. 9. Minimum fatiguing stimulus levels required to produce behavioral alterations at each frequency.

frequency or masking noise level) and reporting the minimum S1 levels that induced MTTTS seems to be the most attractive way of presenting the data. These data are shown in Fig. 7. The levels of fatiguing stimuli necessary to induce MTTTS in the test subjects were consistently within the range of 192 to 201 dB *re*: 1  $\mu$ Pa. The exceptions occurred at the 75-kHz fatiguing stimulus frequency, where one dolphin exhibited an MTTTS after exposure at 182 dB *re*: 1  $\mu$ Pa and the other dolphin did not show any shift after exposure to a maximum level of 194 dB *re*: 1  $\mu$ Pa, and at 0.4 kHz, where no subjects exhibited shifts after exposure to maximum levels of 193 dB *re*: 1  $\mu$ Pa. The shifts that were observed most often occurred at frequencies above that of the fatiguing stimulus. The single occasion on which a negative MTTTS was observed occurred at an S1 level just below that which caused a positive MTTTS in the same animal (NOC, 3-kHz fatiguing stimuli of 193 dB *re*: 1  $\mu$ Pa and 195 dB *re*: 1  $\mu$ Pa, respectively). At the conclusion of the study all thresholds were within 3 dB of baseline/preexposure values.

Hearing thresholds were measured using a unique method (based on the MFR) that featured a vocal response emitted by the test subjects. The modified MFR employed in this study allowed thresholds to be very quickly measured from the test subjects following exposure to the fatiguing stimulus. The first threshold estimate following exposure was routinely obtained within 1–3 min of the exposure, e.g., Figs. 5 and 6. This is particularly important in studies of small TSSs, since recovery from small TSSs may be rapid (e.g., Kryter, 1994).

As seen in Table III, four of the subjects (BEN, NEM, MUK, and NOC) were used in the majority of the experiments. Ideally, naïve subjects of varying age and gender would be used in every experiment; however, availability of trained marine mammals for studies such as this precluded this as an option, a problem inherent to most marine mammal studies. Furthermore, the difficulty of maintaining trained behaviors in the face of aversive stimulus presentation further restricted our subject pool. Ultimately, we were limited not only by which animals were available to us, but also by which animals would continue to do the task. A related difficulty is that individual differences between subjects is especially pronounced when working with a small sample size. With the exception of testing at 3 and 20 kHz, these results are based on an *n* of 2 for each species. This fact makes it difficult to establish absolute levels that will cause a reliable TS in these marine mammals; these data are most useful as a basis to estimate maximum exposure levels below which no large TSSs are likely to occur.

Studies of TTS in humans (e.g., Humes, 1980; Parker *et al.*, 1976) have shown that the presence of masking noise results in elevated hearing thresholds and decreases the amount of TTS observed. We employed masking noise out of necessity: the test location in San Diego Bay has rather high, and variable, ambient noise. There are currently no data concerning the relationship between masking noise and TTS in cetaceans. Preexposure hearing thresholds for the test subjects ranged from approximately 20–40 dB above those measured in *Tursiops* and *Delphinapterus* in quiet conditions, thus we must acknowledge that larger shifts may have

been seen without the masking noise. No MTTTSs were seen during repeated testing at 3 kHz with a higher masking noise level (experiment VII), which also seems to support this statement. The limited data from this study do not support additional conclusions. A more systematic effort with a larger sample size is likely needed to fully explore the effects of masking noise on TTS in cetaceans.

The number of false alarms committed throughout these experiments was low, indicating that the subjects were very conservative. The intent of this study was not to establish absolute thresholds or to compare these thresholds to those obtained elsewhere. More important to this study was that the thresholds be stable and repeatable and that the number of false alarms committed be consistent throughout the study. The fact that the number of false alarms committed post-exposure showed little variability to those committed during preexposure eliminates a change in response bias as a confound for a change in hearing thresholds. Of course it is possible that the occasions on which there were no false alarms committed in either pre- or post-exposure session (see Table IV) could mask a change in response bias, both being arbitrarily high, although this seems unlikely.

Deviations from the animals' trained behaviors were observed at S1 levels approaching those that produced MTTTS. These behaviors continued at S1 levels which produced MTTTS; however, this may have been due to order effects. The lowest SPLs at which these altered behaviors were observed ranged from 178–193 dB *re*: 1  $\mu$ Pa for the dolphins and 180–196 dB *re*: 1  $\mu$ Pa for the white whales. The abundance of baseline data and the animal's previous test history were paramount in establishing a standard by which "altered" behaviors were registered. The trainers' familiarity with the test subjects and the degree of close observation and recording made it possible to not only document behavioral changes, but to relate them directly to the presence of the acoustic signal. Each session was videotaped both underwater and in air. Movements, swimming directions, and postures were recorded during each baseline and MTTTS session. The interval it took the animal to swim between the S1 and S2 underwater stations was timed. Animal acoustic responses just after the S1 were noted as well.

One of the fundamental goals in studying TTS in marine mammals is the identification of acoustic stimulus parameters that affect hearing loss. The peak pressure, total energy, frequency content, rise time, and duration may all be significant with regard to auditory fatigue and TTS, yet the relative importance of and interaction between these properties is unknown. TTS research may often be forced to strike a balance between employing acoustic signals representative of those used operationally by military or industrial sources (which have immediate application to the establishment of safe noise exposure guidelines), or idealized sources that allow one to better explore the effects of the various parameters noted above. For this study, we used 1-s tonal signals at frequencies of 0.4, 3, 10, 20, and 75 kHz. These frequencies were selected in part for their prevalence in military sonars and other anthropogenic acoustic sources. The fatiguing stimulus duration was fixed at 1 s. This duration had the advantage of being short enough that the animal's position

was essentially constant throughout the exposure, thus making estimates of the received level particularly accurate. The disadvantage of the S1 signal duration was that relatively high SPLs were required to produce TTS. The high SPL and very short rise time may have contributed to the behavioral alterations (producing a startle or annoyance response) at the higher S1 levels. Future studies employing longer tone durations should allow TTS to be observed at lower stimulus levels, where behavioral reactions may not be as pronounced; however, longer tone durations produce additional complexities in estimating the received level for a moving subject, especially if the animal surfaces for air. Even brief interruptions in the noise exposure may allow the ears to recover and decrease the amount of TTS observed. Future studies of TTS in marine mammals should also employ impulsive sources such as those commonly encountered in marine seismic surveys and military testing.

#### IV. CONCLUSIONS

A behavioral response paradigm was used to measure masked underwater hearing thresholds in five bottlenose dolphins and two white whales before and immediately after exposure to intense 1-s tones at frequencies of 0.4, 3, 10, 20, and 75 kHz. The levels of fatiguing stimuli necessary to induce MTTs in the test subjects were generally within the range of 192 to 201 dB *re*: 1  $\mu$ Pa. The exceptions occurred at the 75-kHz fatiguing stimulus frequency, where one dolphin exhibited an MTTs after exposure at 182 dB *re*: 1  $\mu$ Pa and the other dolphin did not show any shift after exposure to maximum levels of 193 dB *re*: 1  $\mu$ Pa, and at 0.4 kHz, where no subjects exhibited shifts after exposure to a maximum level of 193 dB *re*: 1  $\mu$ Pa. The shifts that were observed most often occurred at frequencies above that of the fatiguing stimulus. Dolphins began to exhibit altered behavior at levels of 178–193 dB *re*: 1  $\mu$ Pa and white whales displayed altered behavior at 180–196 dB *re*: 1  $\mu$ Pa and above. At the conclusion of the study all thresholds were at baseline values. These data confirm that cetaceans are susceptible to TTS and that small levels of TTS may be fully recovered.

#### ACKNOWLEDGMENTS

We thank Tricia Kamolnick, Jennifer Carr, Monica Chaplin, Lauryn Crosthwaite, Jack Gonzales, and Mark Todd for their assistance with animal training and other technical aspects of the project. Software for computer control of stimuli and recording of animal responses was developed by William Root and was modified by Wesley Elsberry for this project. Tim McBride and Cy Frazier of Program Executive Officer for Undersea Warfare [PEO(USW)], Dr. Frank Stone of N45, and Dr. Robert Gisiner of Office of Naval Research (ONR) gave us frequent encouragement during the work and provided helpful technical review at various stages. Stacie Andersen, Kim De Paul, Tom Fong, and Michelle Reddy were also helpful in facilitating this work. A custom-built noise generator was provided by Dr. Whitlow Au. Funding

was provided by PEO(USW), Naval Air Systems Command (NAVAIR), Major Range Test Facility Base (MRTFB), and ONR. James Finneran was supported by a National Research Council Research Associateship.

- Awbrey, F. T., Thomas, J. A., and Kastelein, R. A. (1988). "Low-frequency underwater hearing sensitivity in belugas, *Delphinapterus leucas*," J. Acoust. Soc. Am. **84**(6), 2273–2275.
- Connor, R. C., and Smolker, R. A. (1996). "'Pop' goes the dolphin: A vocalization male bottlenose dolphins produce during consortships," Behaviour **133**, 643–662.
- Cornsweet, T. N. (1962). "The staircase method in psychophysics," Am. J. Psychol. **75**, 485–491.
- Dancer, A. L., Henderson, D., Salvi, R. J., and Hamernik, R. P., eds. (1992). *Noise-Induced Hearing Loss* (Mosby Year Book, St. Louis, MO).
- Egan, J. P., Greenberg, G. Z., and Schulman, A. I. (1961). "Operating characteristics, signal detectability, and the method of free response," J. Acoust. Soc. Am. **33**, 993–1007.
- Fay, R. R. (1988). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay, Winnetka, IL).
- Finneran, J. J., Oliver, C. W., Schaefer, K. M., and Ridgway, S. H. (2000). "Source levels and estimated yellowfin tuna (*Thunnus albacares*) detection ranges for dolphin jaw pops, breaches, and tail slaps," J. Acoust. Soc. Am. **107**, 649–656.
- Green, D. M. (1976). *An Introduction to Hearing* (Lawrence Erlbaum, New Jersey).
- Green, D. M., DeFerrari, H. A., McFadden, D., Pearse, J. S., Popper, A. N., Richardson, W. J., Ridgway, S. H., and Tyack, P. L. (1994). *Low-Frequency Sound and Marine Mammals: Current Knowledge and Research Needs* (National Academy, Washington, D.C.).
- Hodge, D. C., and McCommons, R. B. (1966). "Reliability of TTS from impulse-noise exposure," J. Acoust. Soc. Am. **40**(4), 839–846.
- Humes, L. E. (1980). "Temporary threshold shift for masked pure tones," Audiology **19**, 335–345.
- Johnson, C. S. (1967). "Sound detection thresholds in marine mammals," in *Marine Bio-Acoustics*, edited by W. N. Tavolga (Pergamon, New York), Vol. 2, pp. 247–260.
- Johnson, C. S., McManus, M. W., and Skaar, D. (1989). "Masked tonal hearing thresholds in the beluga whale," J. Acoust. Soc. Am. **85**(6), 2651–2654.
- Kryter, K. D. (1994). *The Handbook of Hearing and the Effects of Noise* (Academic, San Diego).
- Ljungblad, D. K., Scoggins, P. D., and Gilmartin, W. G. (1982). "Auditory thresholds of a captive Eastern Pacific bottle-nosed dolphin, *Tursiops spp.*," J. Acoust. Soc. Am. **72**(6), 1726–1729.
- Marshall, L., and Brandt, J. F. (1974). "Temporary threshold shift from a toy cap gun," J. Speech Hear Disord. **39**(2), 163–168.
- Nachtigall, P. E. (1986). "Vision, audition, and chemoreception in dolphins and other marine mammals," in *Dolphin Cognition and Behavior: A Comparative Approach*, edited by R. J. Schusterman, J. A. Thomas, and F. G. Wood (Lawrence Erlbaum, Hillsdale, NJ), pp. 79–113.
- Norris, K. S. (1968). "The evolution of acoustic mechanisms in odontocete cetaceans," in *Evolution and Environment*, edited by E. T. Drake (Yale University, Connecticut), pp. 298–323.
- Parker, D. E., Tubbs, R. L., Johnston, P. A., and Johnston, L. S. (1976). "Influence of auditory fatigue on masked pure-tone thresholds," J. Acoust. Soc. Am. **60**, 881–885.
- Richardson, W. J., Greene, Jr., C. R., Malme, C. I., and Thomson, D. H. (1995). *Marine Mammals and Noise* (Academic, San Diego).
- Ridgway, S. H. (1997). "Who are the whales?," Bioacoustics **8**, 3–20.
- Ridgway, S. H., and Carder, D. A. (1997). "Hearing deficits measured in some *Tursiops truncatus*, and discovery of a deaf/mute dolphin," J. Acoust. Soc. Am. **101**(1), 590–594.
- Ridgway, S. H., Carder, D. A., Smith, R. R., Kamolnick, T., Schlundt, C. E., and Elsberry, W. R. (1997). "Behavioral responses and temporary shift in masked hearing threshold of bottlenose dolphins, *Tursiops truncatus*, to 1-second tones of 141 to 201 dB *re* 1  $\mu$ Pa," Tech. Rep. No. 1751, Naval Command, Control and Ocean Surveillance Center, RDT&E Division, San Diego, CA.
- Rosenberger, P. B. (1970). "Response-adjusting stimulus intensity," in *Animal Psychophysics: The Design and Conduct of Sensory Experiments*,

- edited by W. C. Stebbins (Appleton-Century-Crofts, New York), pp. 161–184.
- Silbiger, H. R. (1965). “Temporary threshold shifts due to single clicks,” *J. Acoust. Soc. Am.* **38**, 937–938.
- Smolker, R., and Richards, A. (1988). “Loud sounds during feeding in Indian Ocean bottlenose dolphins,” in *Animal Sonar: Processes and Performance*, edited by P. E. Nachtigall and P. W. B. Moore (Plenum, New York), pp. 703–706.
- Watkins, W. A., and Wartzok, D. (1985). “Sensory biophysics of marine mammals,” *Mar. Mamm. Sci.* **1**(3), 219–260.
- White, M. J., Jr., Norris, J., Ljungblad, D., Baron, K., and di Sciara, G. (1978). “Auditory thresholds of two beluga whales (*Delphinapterus leucas*),” Hubbs/Sea World Research Institute Tech. Rep. 78-108.

# Calibration and comparison of the acoustic location methods used during the spring migration of the bowhead whale, *Balaena mysticetus*, off Pt. Barrow, Alaska, 1984–1993

Christopher W. Clark<sup>a)</sup>

Cornell Lab of Ornithology, Bioacoustics Research Program, 159 Sapsucker Woods Road, Ithaca, New York 14850

William T. Ellison<sup>b)</sup>

Marine Acoustics, Inc., P.O. Box 340, Litchfield, Connecticut 06759

(Received 3 September 1999; accepted for publication 23 February 2000)

Between 1984 and 1993, visual and acoustic methods were combined to census the Bering-Chukchi-Beaufort bowhead whale, *Balaena mysticetus*, population. Passive acoustic location was based on arrival-time differences of transient bowhead sounds detected on sparse arrays of three to five hydrophones distributed over distances of 1.5–4.5 km along the ice edge. Arrival-time differences were calculated from either digital cross correlation of spectrograms (old method), or digital cross correlation of time waveforms (new method). Acoustic calibration was conducted *in situ* in 1985 at five sites with visual site position determined by triangulation using two theodolites. The discrepancy between visual and acoustic locations was <1%–5% of visual range and less than 0.7° of visual bearing for either method. Comparison of calibration results indicates that the new method yielded slightly more precise and accurate positions than the old method. Comparison of 217 bowhead whale call locations from both acoustic methods showed that the new method was more precise, with location errors 3–4 times smaller than the old method. Overall, low-frequency bowhead transients were reliably located out to ranges of 3–4 times array size. At these ranges in shallow water, signal propagation appears to be dominated by the fundamental mode and is not corrupted by multipath. © 2000 Acoustical Society of America.

[S0001-4966(00)01006-7]

PACS numbers: 43.80.Lb, 43.80.Ka, 43.30.Sf [WA]

## INTRODUCTION

The use of passive acoustic methods to assist in population assessment and behavioral studies of large whales has been of interest to researchers for many years (Clark, 1982; Payne and Guinee, 1983; Schevill *et al.*, 1964; Thomas *et al.*, 1986; Thompson and Friedl, 1982; Watkins, 1981; Winn *et al.*, 1975). Between 1984 and 1993, six surveys of the Bering-Chukchi-Beaufort population of the bowhead whale, *Balaena mysticetus*, were conducted off Point Barrow, Alaska (71° 23' N and 156° 26' W). These springtime surveys combined passive acoustic methods with traditional visual census methods to describe acoustic behavior and to estimate population size and trends (Clark *et al.*, 1996; Clark and Ellison, 1989; Clark *et al.*, 1986a; Raftery and Zeh, 1998; Raftery *et al.*, 1990; Würsig and Clark, 1993; Zeh *et al.*, 1993). For all surveys, the acoustic method relied on a sparse array of hydrophones, with one hydrophone located near the visual observation perch. Other hydrophones were distributed along the ice edge to either side of the perch out to distances of 0.5 to 3 km from the perch. Ice conditions and safety considerations primarily dictated the total size and configuration of the array. The perch was typically located

on a grounded 5–10-m-high ice ridge that was 6–15 km from the shoreline, and the ice edge extended in both directions approximately parallel to the shoreline.

The method for locating a vocalizing animal relied on measurement of the difference in time-of-arrival of the same sound on each of three to five hydrophones in the array (Clark *et al.*, 1986b; Cummings and Holliday, 1985; Watkins, 1976; Watkins and Schevill, 1972). Verification of the method was necessary to empirically quantify the errors in range and bearing estimates to the transient source. In most marine mammal research relying on arrays, calibration of the arrival-time difference method has relied on the use of broadband impulsive sounds, or high-frequency bursts produced by pingers within close proximity of the array (Janik *et al.*, 2000; Watkins and Schevill, 1972), or empirical use of visually positioned singer locations (Frankel *et al.*, 1995). There have been no calibration tests using low-frequency, frequency-modulated (FM) transients in the shallow water, arctic environment. This paper describes the technical aspects of two different arrival-time-difference methods used to locate vocalizing bowhead whales during the spring migration. It provides the results of a field calibration test conducted *in situ* in 1985 during the bowhead whale migration. The 1985 calibration data and a sample of 217 bowhead calls from 1988 were reanalyzed and used as the basis for comparing and evaluating the reliability of the two methods.

<sup>a)</sup>Electronic mail: cwc2@cornell.edu

<sup>b)</sup>Electronic mail: wtemai@aol.com



## I. METHODS

In the arctic environment through which bowhead whales migrate during the spring, the acoustic propagation path is relatively simple. The water is shallow ( $<100$  m) and isothermal ( $0 \pm 1$  °C), and the sound speed profile is uniform with depth ( $1437 \pm 2$  m/sec). Occasionally, the propagation conditions become complex and unpredictable; for example, when deep-keeled, multiyear ice floes create a highly reflective environment and different types of ice create horizontal refraction (Ellison *et al.*, 1987; George *et al.*, 1989). However, the dominant springtime acoustic propagation condition over ranges of a few tens of kilometers off Alaska's North Slope is relatively stable and homogeneous.

Bowhead whales produce sounds that are transient and highly variable. Signal energy is primarily in the 80–500-Hz band with single utterances having bandwidths of 50–150 Hz, durations on the order of 0.5–4 s, and spectrum source levels of 155–189 dB *re* 1  $\mu$ Pa at 1 m (Clark and Johnson, 1984; Cummings and Holliday, 1985; Ljungblad *et al.*, 1982). Almost all calls have a structured spectrum consisting of discrete frequency bands of acoustic energy that vary in amplitude and frequency over the duration of the sound. Sounds can be simple monotonies with very little frequency modulation (FM), swept FM calls where a fundamental frequency rises or falls with time throughout the call's duration, amplitude modulated (AM) purrs, harmonically rich screams, and complex mixtures of FM, AM, and broadband pulses (Clark and Johnson, 1984; Ljungblad *et al.*, 1982; Würsig and Clark, 1993).

The irregular ice edge in the spring restricts array geometry to a roughly linear orientation aligned approximately parallel with the direction of whale migration. One drawback to a linear array is that sources to either side of and near the axis of the array (i.e., endfire) are difficult to locate. As a result, for a sound source outside of the array but near the array axis, a small uncertainty in time-of-arrival measurement results in a large uncertainty in range to the source, while bearing to the source remains relatively stable. This restricts the area within which locations from a linear array are reliably accurate. For the purposes of the arctic arrays used in this research, the area of reliable locations was restricted to a sector of  $120^\circ$  centered on the line normal to the array axis at the array center (Clark *et al.*, 1986b).

A vocalizing whale's location was estimated using measurements of the arrival-time difference between the occurrence of the same sound on different pairs of hydrophones (Fig. 1). The acoustic location task was simplified to locating a transient sound in two dimensions. This was because water depths were relatively shallow (15–100 m) compared to the ranges to vocalizing animals (100–20000 m). It was also assumed that the whale was an omnidirectional source, and sound propagated at a constant speed in water that was homogeneous, isotropic, and nondispersive.

Over the years of the acoustic survey effort, two different methods for measuring arrival-time difference and for computing an estimate for the location of a vocalizing whale were developed and implemented. The first method was used to analyze the array recordings collected during the acoustic

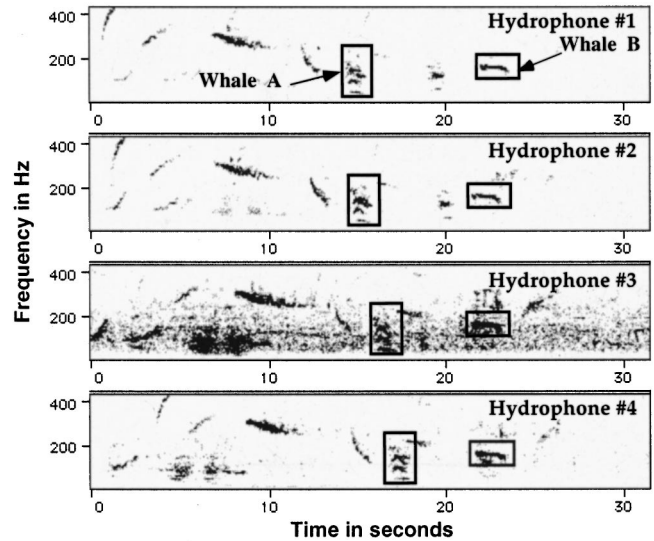


FIG. 1. Multichannel spectrogram example from 1988 array deployed off Pt. Barrow, Alaska illustrating arrival-time differences for bowhead calls at four different hydrophones. Calls from two different bowhead whales are highlighted.

surveys in 1984, 1985, 1986, and 1988, while the second method was used for the 1992 and 1993 acoustic studies.

### A. Old location method

In the first location method (old method), arrival-time difference was estimated by cross correlation of filtered spectrograms (Altes, 1980; Clark *et al.*, 1986a; Clark *et al.*, 1987; Janik *et al.*, 2000; McGregor *et al.*, 1997). A multichannel digital spectrogram representing the time-varying frequency distribution (i.e., short-time Fourier transform, STFT) was computed and displayed for the set of received pressure waveforms containing the sound of interest [1000-Hz sampling rate, 256-point fast Fourier transform (FFT) rectangular window function, with sampling overlap (99%) equal to  $\pm 2$ -ms time resolution, and 3.9-Hz frequency resolution]. An analyst identified the lower and upper frequencies, and start and end times containing the sound, and these were used to restrict the frequency band and signal duration in subsequent analysis. Pairs of band-limited and duration-limited spectrograms were cross correlated.

The following notation describes this spectrogram cross-correlation process:

$$C(\Delta T) = \frac{1}{n \cdot k} \sum_n \sum_k A[f(k), n\Delta t] \times B[f(k), n\Delta t + \Delta T], \quad (1)$$

where  $A[f(k), n\Delta t]$  and  $B[f(k), n\Delta t]$  are digital spectrograms of signals from hydrophone A and B, respectively,  $f(k)$  is the normalized amplitude spectrum at time  $n\Delta t$ ,  $\Delta t$  is the time interval between successive spectra,  $n$  is the index for successive spectra, and  $\Delta T$  is the time offset between the two signals.

The cross correlation of two digital spectrograms results in a time-dependent function,  $C(\Delta T)$ , for which the time of the maximum correlation value ( $\Delta T_{\max}$ ) was taken as the best estimate of the arrival-time difference,  $\tau$ , between the occurrence of the same sound at the two hydrophones. In

practice, the time interval between successive spectra,  $\Delta t$ , was set to 2 ms by having a 254-sample overlap between successive 256-point spectra.

In a cylindrical coordinate system centered at the midpoint between a pair of hydrophones, a given arrival-time difference,  $\tau$ , corresponds to the loci of points  $(R, \phi)$  referred to as a hyperbolic bearing line, such that,

$$\tau = \tau_{\max} \{ [R^2 + R/d(\cos(\phi)) + 1/4]^{1/2} - [R^2 - R/d(\cos(\phi)) + 1/4]^{1/2} \}, \quad (2)$$

where  $\phi$  is the bearing from the array's origin to the sound,  $R$  is the range from the array's origin to the sound,  $\tau_{\max}$  is  $d/c$  (the maximum possible time-difference between the hydrophone pair),  $c$  is sound velocity, and  $d$  is the distance between the two hydrophones.

For large values of  $R$ , the asymptotic result is

$$\cos(\phi) = \tau / \tau_{\max}, \quad R/d \gg 1. \quad (3)$$

The symmetry of the cosine in (2) creates an ambiguity in determining the proper half-space in which to place the hyperbolic bearing line. This ambiguity was resolved by knowing at which hydrophone the sound first arrived. A second ambiguity existed as to whether the sound was arriving from the frontside or backside of the array. In practice, this front-back ambiguity was resolved by deploying the array along the edge of the shore-fast ice, thereby restricting whales to the offshore side of the array. The addition of one or more hydrophones at other locations resulted in additional arrival-time differences, where the number of hydrophones,  $N_{hy}$ , yielded  $N_{hy} * (N_{hy} - 1) / 2$  arrival-time differences. Ideally, arrival-time differences are perfectly accurate and the bearing lines all cross at a single point. However, noise in the transmission paths and the recording system, variability in sound speed, and discretization in the analysis processes can introduce errors in arrival-time difference estimates which result in uncertainty in the source location. In practice, the acoustic analyst subjectively evaluated spectrograms and the aural quality of all available channels, and only the three best channels were used for arrival-time difference analysis. A graphical solution was used to estimate the location as illustrated in Fig. 2. The location was estimated as the centroid of the triangle defined by the intersections of the three bearing lines calculated from arrival-time differences. The acoustic location range error was the distance between the center of the triangle and the farthest vertex. The acoustic location bearing error was half the maximum angular width of the triangle.

## B. New location method

In the second location method (new method), arrival-time difference was estimated by cross correlation of filtered waveforms from pairs of hydrophones (Clark *et al.*, 1986a; Marple, 1987; Mitchell and Bower, 1995; Spiesberger and Fristrup, 1990). The time offset associated with the maximum value of the complex demodulated, cross-correlation function was taken as the best estimate of the arrival-time difference. A user-friendly, array location analysis program specifically written for the task of locating and tracking vo-

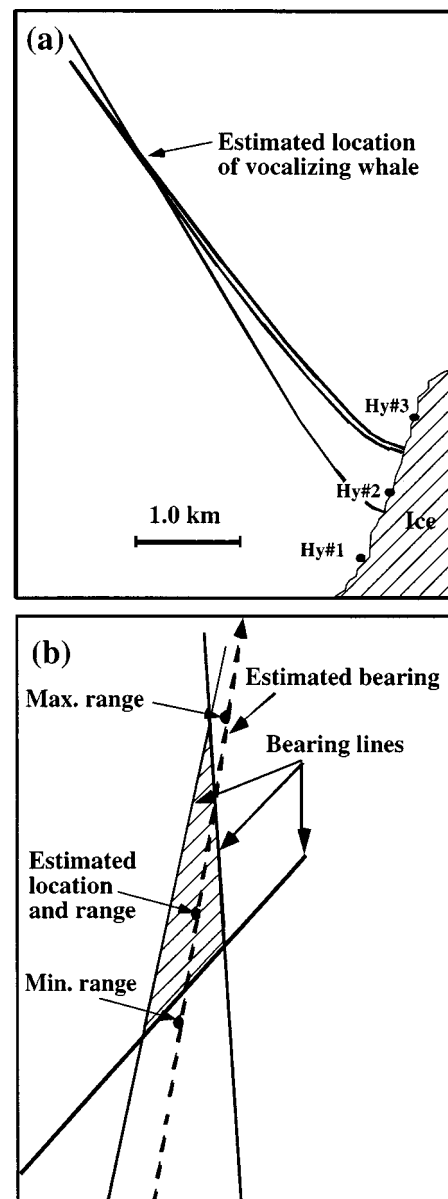


FIG. 2. Example of old method for determining an acoustic location. (a) The three hyperbolic bearing lines representing the spectrographic cross-correlation time delays from a three-hydrophone array. (b) An expanded portion of the bearing line intersections showing the estimated location of the vocalizing whale, and the errors associated with this location.

calizing bowhead whales facilitated this process. With this analysis system, the analyst viewed scrolling, multichannel spectrograms while listening to any of the various channels. The analyst was responsible for recognizing which signals were bowhead calls, selecting a time-frequency box around the call in one channel, and time-shifting the selection in the remaining channels so as to identify the same sound in each of the channels. The lowest and highest frequency values then set bandpass filter parameters for filtering out unwanted noise, while the start and end times delimit the portions of the waveforms containing the sound of interest.

The acoustic location was determined from the consistency of redundant information provided by the arrival-time-difference estimates for all pairs of hydrophones (see Clark *et al.*, 1996, Appendix 1). Source location was computed us-

ing a minimum sum-of-squares error technique based on the set of measured arrival-time-difference values and sets of theoretical arrival-time-difference values. Theoretical arrival-time-difference values were those that would occur for each point in a gridded  $x,y$  space for the given array geometry and speed of sound conditions. In practice, the gridded space had a resolution of 30 m, equivalent to a bow-head call bandwidth of 30 Hz.

The minimization was done with the Levenberg-Marquardt algorithm. The starting point of the minimization algorithm was obtained by finding the intersections of the asymptotes of the hyperbolic bearing lines for each arrival-time difference, and choosing as the starting point the intersection with the smallest error.

Since errors were present in the acoustic location process, each acoustic location had a confidence region associated with it, where the confidence region represented some area of geographic positions (around the estimated location) whose arrival-time-difference errors were less than a minimum factor, or confidence region, of 80%. The boundary extremes of this region were used to compute the minimum range, maximum range, and bearing error statistics. In Fig. 3(a) the 80% confidence region associated with an acoustic location is shown as an almost elliptical area surrounding the acoustic location point.

To resolve the problem of including poor or obviously aberrant arrival-time differences in the location process, a screening algorithm was developed, tested, and programmed into the new location analysis process. This algorithm, referred to as the “sigma minimizer” program, minimized the sum-of-squares error (sigma) in the location process. The sigma minimizer analyzed the set of all measured arrival-time differences and discarded either zero, one, or two differences to obtain the best fit (i.e., minimum error) between the set of measured arrival-time differences and the set of theoretical arrival-time differences for the estimated location. This eliminated outlying arrival-time differences resulting from cross correlations of acoustic data that were corrupted by multipath sound propagation, excessive ambient noise, or surface reverberation. An example of the improvement in acoustic location error as a result of this sigma minimizer method is illustrated by comparing Fig. 3(a) and (b).

The old location method was completed on a Digital Equipment Corporation (DEC) 11/23+ minicomputer with a Sky array processor. The customized software was developed in collaboration with Kim Beeman (Clark *et al.*, 1986a). The new location method was completed on a Macintosh-based workstation using customized DSP hardware boards. The location software was written by Tom Calupca, Sean Cunningham, and Steve Mitchell and merged with the Canary software package (Charif, 1995).

### C. 1985 Calibration experiment

One of the foremost concerns regarding the use of acoustic techniques for locating bowhead whales was the lack of a full scale, *in situ* validation of the arrival-time-difference method for locating sounds. In order to address this deficiency, a calibration experiment was conducted dur-

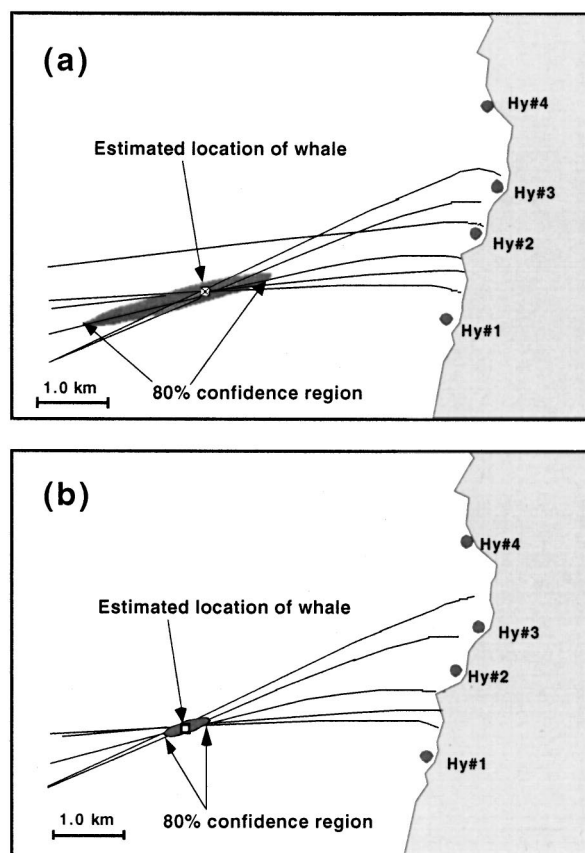


FIG. 3. (a) Estimation of location and 80% confidence region for a calling whale using all six time-delay bearing lines from a four-channel hydrophone array. The large confidence region around the whale's location is due principally to the bearing line from the time delay between hydrophones 1 and 4, which intersects the other bearing lines far from the estimated position. The 80% confidence region is considered the error region and defines the range and bearing errors associated with the location of the vocalizing whale. (b) Effect of eliminating a corrupted bearing line on a whale's acoustic location and associated error region for the same whale call as shown in (a). This illustrates the reduction in the range and bearing errors when an erroneous bearing line is excluded.

ing the 1985 field season to determine the accuracy of the sound location method relative to a reliable visual method.

The calibration test plan assumed an array of at least three hydrophones and a total of ten sites, three each at ranges of 500, 1000, and 2000 m from the array's center on bearings of 30, 90, and 135° from the array axis, with an additional site at a minimum distance of 4000 m, if feasible.

The array was established within the visual census reference frame based on bearings relative to Magnetic North, and a fixed position relative to the RACON Beacon at Point Barrow, Alaska (NOAA, 1976). Only three hydrophones were installed due to difficult ice conditions. The array geometry and the “true” position of each playback site were determined using two theodolites and a Cubitape DM-60 modulated infrared light source theodolite sighting system with an accuracy of  $\pm 5$  mm. The “true” position of each playback site was determined using triangulation from the two theodolites.

Three approaches were used to calculate the local sound velocity. Historical measurements in this area documented temperature stability ( $-1.7 \pm 0.2$  °C), salinity stability

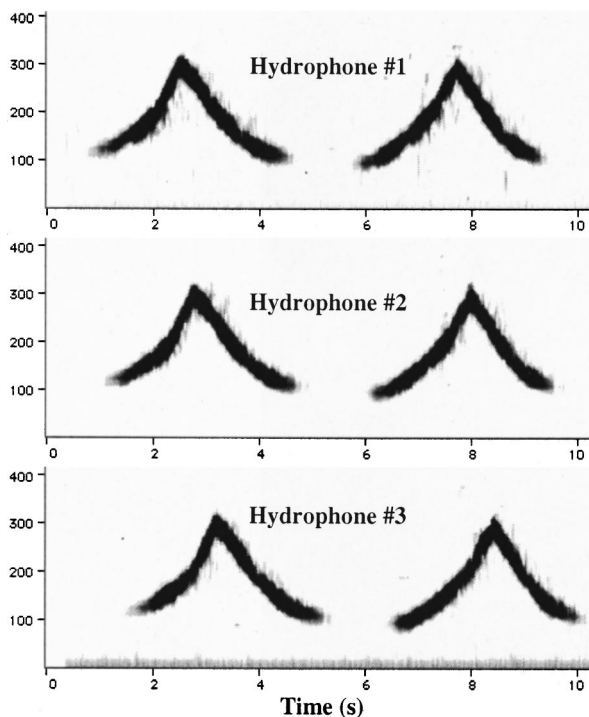


FIG. 4. Example of the two calibration sweeps as recorded on the array during the calibration test experiment on 28 April 1985. The J-11 source was at site E, 4420 m from the visual observation perch.

( $32 \pm 2.0\%$ ), and a speed of sound between 1440 and 1438 m/s (Garrison, 1976), resulting in a sound velocity of  $1437 \pm 14.4$  m/sec. In 1982, Cummings *et al.* (Cummings *et al.*, 1983) used a sound velocimeter and measured a sound velocity of 1438.2 m/sec at a depth of 1 m with a maximum excursion to 1437.2 m/sec at a depth of 9 m. In 1984, at a location close to the 1985 calibration site, we calculated a sound velocity of  $1441 \pm 4$  m/s using measurements of maximum arrival-time differences between two widely spaced hydrophones for an impulsive sound source. Based on all these data, the value of 1437 m/sec was used in all location analyses.

The sound playback transducer was a J-11 Audio Frequency Generator (Naval Research Lab, 1975). This is a laboratory standard, calibrated transducer specifically designed for experimental use. It operates over a frequency band of 20–12000 Hz and is capable of generating intensities of up to 162 dB *re* 1  $\mu$ Pa at 1 m. A variable-frequency function generator was used to synthesize two bowhead like, frequency-modulated sounds lasting a total of approximately 8 sec (Fig. 4). This alternating pair of sounds was repeated multiple times onto a continuous 15-min tape. The tape was played back on a Sony cassette recorder through a Crown preamplifier and a 300-watt Crown amplifier into the J-11. Most sounds were played back at approximately mid water depth at maximum source level. However, during the lowering and raising of the source, careful attention was paid to the depth of the source at the time of a transmission, resulting in a small set of transmissions at various depths. Electricity was provided by a portable gas powered generator. The output power of the Crown 300-watt amplifier was

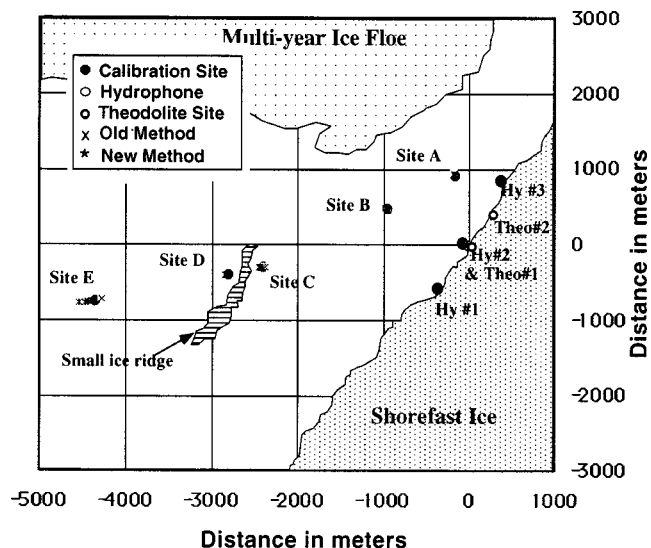


FIG. 5. Schematic of 1985 calibration geometry and acoustic locations for five test sites (A–E) based on old and new acoustic location methods. “Theo#1” and “Theo#2” refer to theodolite sites.

monitored by an alternating current (AC) voltmeter across the output leads to the J-11.

Underwater sounds were collected using modified AN/SSQ-57A sonobuoy units deployed along the ice edge (Cummings and Holliday, 1985). Each unit operated on a different radio frequency. A multichannel receiver and recording system was maintained in a small portable hut a safe distance away from the ice edge (Clark *et al.*, 1986b). This equipment included radio receivers, signal amplifiers, and a TEAC R61-D four-channel cassette tape recorder using normal-bias tape. The recording system had a flat frequency response  $\pm 1$  dB between 100 and 5000 Hz with a low-frequency roll-off of 20 dB for the 10–100 Hz band.

In 1985, ten calibration sounds from each of the five calibration sites were randomly selected and analyzed using the old method. More recently, the exact same 50 transmission samples as analyzed in 1985, as well as all other transmissions from each calibration site, were reanalyzed using the new method.

#### D. Reanalysis of 1988 whale data using new methods

To compare the two acoustic location methods using bowhead sounds, 38 hours of four-channel array data from the 1988 field study were reanalyzed using the new method.

## II. RESULTS

### A. 1985 calibration experiment

The calibration test area was located north northeast of Point Barrow immediately offshore of a grounded pressure ridge in an area very typical of bowhead census sites. The water depth in the calibration test area varied between 15 and 100 m. Two theodolites were established on the pressure ridge on an axis of  $19.7^\circ$  magnetic, separated by a distance of 482.3 m. A sparse three-element array was installed on 27 April, with hydrophones deployed along the edge of the grounded pressure ridge (Fig. 5). The middle hydrophone

(Hy#2) of the array was installed close to the southern theodolite site. Hydrophones were lowered to a depth of 8–10 m, which was approximately mid water depth. The measured array had a 1616-m aperture with spacings between adjacent pairs of hydrophones of 670 and 946 m. The local ice conditions on the day of the experiment, 28 April, were very good. There was a thin sheet of 10–30-cm pan ice in front of the grounded ridge system. This pan reduced local ambient noise from wind, waves, and hydrophone motion. Interspersed were small pockets of open water and some working fracture zones. A large floe of grounded multiyear ice was directly off shore, approximately 1–2 km northwest of the visual observation perch (George *et al.*, 1989), and a small (3–5-m high) pressure ridge ran parallel to the array approximately 2 km to the west of the perch. Due to time and safety constraints, calibration tests were completed at only five sites.

Calibration sounds were played into the water at each of the five sites, with the time spent at each site varying between 9 and 27 min, depending on safety considerations. The J-11 was suspended to a maximum of 12-m depth at sites A–D, and 24-m depth at site E. In a few cases, source level at the start of a transmission session was 150 dB and increased by approximately 3 dB transmission until maximum power (162 dB) was achieved. Sites A–C had direct paths from the transducer to all three hydrophones, while sites D and E were on the far side of the 3–5-m ice ridge. Site E, at a range of 4420 m, was just barely visible to the observers on the perch. This determined the range limit of the calibration test.

A research assistant completed location analysis using the old method in early May 1985. Ten randomly selected calibration sounds were analyzed for each of the five sites. Later in the month one of us (C.W.C) reanalyzed the same 50 randomly selected calibration sounds and obtained identical results. In 1998, all calibration sounds from each of the five sites were analyzed using the new method. This allowed comparison of the analyses of the ten random sounds from each site using both location methods.

Table I lists the results from the analyses of the calibration data using both location methods. All the calibration results for the five sites using both methods are plotted in Fig. 5. Bearing results are very consistent for both acoustic methods, varying by only a fraction of a degree for all five sites. There is a slight reduction in the standard deviations for bearing using the new method [mean old standard deviation (s.d.) 0.29°, mean new s.d. 0.03°,  $n=50$ , five sites, based on ten samples per site; mean new standard error (s.e.) 0.16°,  $n=5$  sites, based on all samples per site].

Range results are remarkably consistent even at the farthest distance of 4420 m (site E). The old method estimates a slightly greater average range than the new method for site A, but lower average ranges for sites B–E.

The percent differences between the visual and acoustic ranges relative to the visual range using the old method were between 1% and 5%, with coefficients of variation (CV) between 1 and 4%. The percent differences and the CVs for the new method were between 1% and 2% and less than 1%, respectively, for the samples of ten. For the complete set of

TABLE I. Ranges in meters ( $\pm$ s.d.) and bearings in degrees ( $\pm$ s.d.) for five 1985 calibration sites relative to the visual observation perch as determined by visual and acoustic methods (see Fig. 5). Acoustic ranges and bearings were computed using the spectrogram cross-correlation (old acoustic) and waveform cross-correlation (new acoustic) methods. “Range” is the average distance from the visual observation perch, origin (0,0), to the location of the acoustic source, while “bearing” is the average bearing to the source from the perch relative to magnetic North. Time in parentheses is the duration spent at the site broadcasting the calibration sounds.

	<i>N</i>	Range $\pm$ s.d. (m)	Bearing $\pm$ s.d. ( $^{\circ}$ mag.)
Site: A (11 min)			
Visual		924	349.8
Old acoustic	10	959 $\pm$ 15.8	349.3 $\pm$ 0.58
New acoustic	10	941 $\pm$ 1.4	349.8 $\pm$ 0.06
New acoustic all	129	942 $\pm$ 4.9	349.8 $\pm$ 0.20
Site: B (26 min)			
Visual		1063	296.3
Old acoustic	10	1050 $\pm$ 12.9	295.6 $\pm$ 0.38
New acoustic	10	1082 $\pm$ 8.9	296.5 $\pm$ 0.02
New acoustic all	231	1074 $\pm$ 17.6	296.5 $\pm$ 0.30
Site: C (17 min)			
Visual		2417	262.8
Old acoustic	10	2404 $\pm$ 54.6	262.9 $\pm$ 0.15
New acoustic	10	2455 $\pm$ 8.0	263.0 $\pm$ 0.02
New acoustic all	159	2450 $\pm$ 14.7	263.1 $\pm$ 0.10
Site: D (9 min)			
Visual		2829	261.8
Old acoustic	10	2679 $\pm$ 94.2	261.7 $\pm$ 0.22
New acoustic	10	2863 $\pm$ 12.2	261.9 $\pm$ 0.04
New acoustic all	106	2849 $\pm$ 22.6	262.0 $\pm$ 0.10
Site E (4 min)			
Visual		4420	260.3
Old acoustic	10	4341 $\pm$ 168.0	260.6 $\pm$ 0.11
New acoustic	10	4425 $\pm$ 36.7	260.4 $\pm$ 0.03
New acoustic	47	4421 $\pm$ 21.1	260.5 $\pm$ 0.10

samples using the new method, the percent differences and CVs were between 0% and 2% and less than 1%, respectively. Overall the new method resulted in slightly greater precision (lower standard deviations) and greater consistency (less variability in percent differences) than the old method. The greatest improvement between the two methods occurred at the two most distant positions, sites D and E. At site B, where the calibrations occurred over a 26-min period and theodolite fixes were taken only at the start of the test, there was a small but steady increase in range, suggesting that perhaps the ice pan from which the calibration equipment was deployed was moving northward. Times spent at the other sites were too short to determine if this occurred elsewhere. Overall, there was some indication that location and associated errors decreased slightly with transmitter depth, but not with source level.

## B. Reanalysis of 1988 bowhead calls using new methods

A second comparison of the old and new methods was based on the analysis of 38 hours of bowhead whale recordings from a four-channel array in 1988. The 1988 array had a 3031-m aperture with spacings between hydrophones of

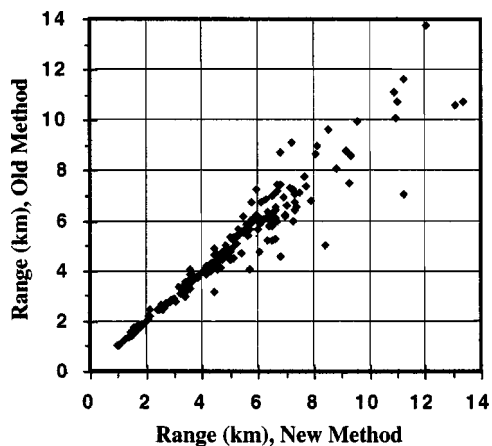


FIG. 6. Comparison of range estimates for the same 217 bowhead calls recorded during the 1988 acoustic census, as computed using the old and new methods.

900, 1150, and 990 m. Reanalysis of the 1988 data using the new method located 236 of the same calls that had been located by the old method. Nineteen of the calls located by the old method failed to yield acceptable locations by the new method. Ten of these 19 unacceptable locations were within a few degrees of the edges of the 120° sector in which locations are considered acceptable. The location data for the acceptable 217 calls were used to compare the old and new methods.

Figure 6 shows a comparison of the range estimates for the old and new method for each of the 217 calls located in the 1988 data. Range estimates were similar for the two methods ( $r^2=0.96$ ), and as expected range differences increased as a function of range. Bearings estimated by the two methods were essentially identical ( $r^2=0.999$ ).

Range error was expected to increase with increasing range. Therefore, range errors for the two methods were compared as a function of range. Since ranges calculated with the new method were considered better estimates of actual range, the new range estimate was used as the independent variable in this comparison. Figure 7 shows differ-

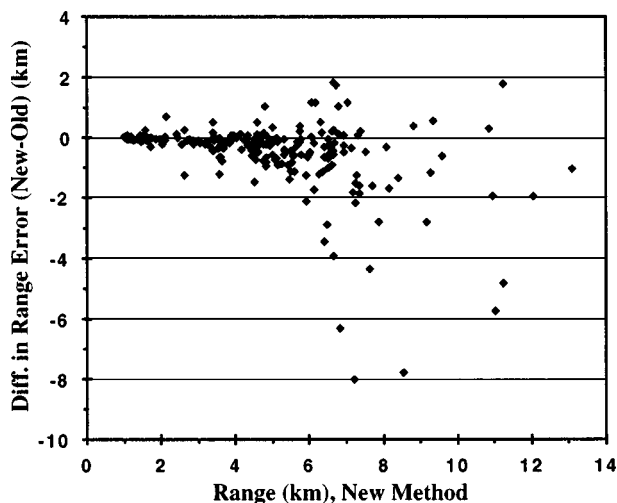


FIG. 7. Difference in range errors (new-old) for the 217 bowhead calls from 1988 relative to the range estimated by the new method. A negative difference indicates that the new method yielded a lower error value.

ences between range errors for the new and old methods as a function of the new range for each pair of 217 locations. The predominantly negative range error differences indicate that the new range errors were consistently less than the corresponding old range errors. Average range error for the old method was  $1138 \pm 2094$  m compared to the average range error of  $275 \pm 420$  m for the new method. When considering range error as a percentage of the range, the averages were  $17.9\% \pm 21.9\%$  for the old method and  $4.7\% \pm 5.8\%$  for the new method. Bearing error was  $>1^\circ$  for both methods, but greater for the old method ( $0.6 \pm 0.56^\circ$ ) than the new method ( $0.2 \pm 0.22^\circ$ ). These results indicate that range and bearing errors for the new method were approximately 3–4 times smaller and less variable than the errors using the old method.

### III. DISCUSSION

The primary motivation for the calibration experiment and comparative analysis was to empirically evaluate the reliability of methods developed to locate transient vocalizations from bowhead whales. Each spring these animals migrate through the shallow water, marginal ice zone along Alaska's north slope, an environment that is very dynamic and for which there are few data on acoustic transmission. The signal processing methods implemented for this work were relatively straightforward. However, their bioacoustic application under the demanding arctic conditions was novel and has led to significant progress in a quantitative assessment of the bowhead population (Raftery and Zeh, 1998; Raftery *et al.*, 1990; Zeh *et al.*, 1993). It is noteworthy that this calibration was conducted under typical field conditions using standard acoustic data collection procedures. No special precautions were in effect for this experiment. Ice conditions on the day of the calibration were similar to the conditions that existed throughout much of the 1984–1993 bowhead census studies. Whale sounds were recorded simultaneously with the calibration sounds, and there was no immediately obvious change in vocal activity associated with the calibration transmissions.

The reanalysis of 38 hours of 1988 data ( $n=217$  locations) leads to two conclusions. First, the two methods yield similar range distributions, with absolute differences in range estimates increasing as the distance to the vocalizing whale increased (see Fig. 6). Overall, both methods had approximately the same bias in location estimates of the vocalizing whale. Second, the range errors associated with a location were four times smaller with the new method than with the old method, while the bearing errors were three times smaller (old method,  $1138 \pm 2094$  m,  $0.6 \pm 0.56^\circ$ ; new method,  $275 \pm 420$  m,  $0.2 \pm 0.22^\circ$ ).

The calibration results comparing the acoustic location methods with locations obtained by theodolite cross bearings showed that the new method yielded locations with less variability but only slightly less bias than locations using the old method. The conclusion that both methods have approximately the same bias is supported by the results showing that the percent difference between the visual and acoustic ranges to the five acoustic source locations using the old method was 1%–5%, while it was  $<2\%$  for the new method.

The conclusion that the new method was less variable was supported by the results showing that the coefficients of variation for the new method were 0.5–2% compared to 1–4% for the old method. On average, the new method was about 2–3 times more precise than the old method.

For either location method there was a high degree of accuracy out to the maximum calibration range of 4.4 km. The range differences between the visual and the two acoustic methods were less than 5% of the range, and the bearing errors were less than a degree. There was no indication that range errors increased with range, although the maximum calibration range was less than three times the array aperture (1616 m). At first glance, the high degree of accuracy and similarity of calibration analyses results for the two methods seems surprising. However, this is most probably attributable to the FM structure of the test signals. These were ascending and descending frequency swept signals (200-Hz bandwidth, 3–4-sec duration, slew rate  $\sim 100$  Hz/sec) with relatively high levels of time-bandwidth structure in their STFT representations, which, when cross correlated, yielded reliable estimates of time delay. These test signals are similar to natural bowhead whale calls, supporting the conclusion that the calibration results are directly transferable to those based on bowhead calls (Clark and Johnson, 1984).

These improvements were not altogether unexpected. The method of determining time differences between the occurrence of the same sound at different hydrophones was improved in the new method by using the entire waveform rather than its STFT. In the old method, information in the phase spectra was ignored. Furthermore, and of greater importance here, the STFT had the effect of smoothing the two signal representations, thereby reducing time resolution. The use of a high level of overlap (99%) in the STFT process increased time resolution computationally (i.e., one could compute  $\Delta T$  functions with 1–2 ms resolution), but did not necessarily improve measurement accuracy. Location estimation with the new method was improved by including both amplitude and phase information. This, coupled with the 1000-Hz sampling rate, increased time resolution and reduced the uncertainty in the arrival-time differences, resulting in lower range and bearing errors (see Figs. 6 and 7). Part of this improvement was also due to the new method including all possible pairwise arrival-time differences, whereas the old method only included three, and to using an optimization routine rather than a graphical solution to estimate location. As a result, on average, the precision of a location was improved by a factor of 4 with the new acoustic location method.

The small range errors for the two acoustic methods as applied to the calibration data were within the error expected based on the time resolutions ( $\pm 1$ –2 ms) of the computation systems alone. This result supports the conclusion that noise errors introduced by the transmission paths were relatively small and that the medium provided a very coherent and stable path. In a related arctic acoustic experiment, DiNapoli *et al.* (1984) found a very high level of coherence (0.77–0.96 cross-correlation coefficient) for a 97-Hz signal received on two different hydrophones separated by as much as 1300 m. In their experiment, however, the source was located at a

distance of over 250 km from the hydrophones. The very high level of coherence observed with a signal transmitted over such distances indicates that the arctic water column is stable.

It should be noted that at the lower-frequency range of concern here (<500 Hz), and with the water depth being only several wavelengths in extent, the acoustic path of interest is best described as a plane wave ducted mode rather than a ray path. The bulk of the acoustic energy is being transmitted in the fundamental mode. However, the actual energy partition between the fundamental mode and any higher modes in the shallow water will require further experimentation.

These methods were specifically designed to take advantage of the broadband nature of the whale signals. The old method of using spectrogram cross correlation was originally selected because spectrograms readily displayed a sound's time-varying FM structure while waveform did not. In the spectrographic rendering, signal structure was conserved even with pronounced interference from propagation effects in the shallow water, arctic environment. For this reason we expected that spectrogram cross correlation would be more robust than waveform cross correlation and allow range and bearing estimation out to greater distances. A related presentation on the benefits from ranging with broadband signal was presented by Harrison and Harrison (1995), showing that a key benefit of using broadband signals is to reduce the effect of nulls on reception, thereby effectively increasing communication range. It is assumed that bowhead signals are well adapted to their arctic habitat, and that some of these signals serve a long-range communicative function (Clark and Johnson, 1984; Würsig and Clark 1993). It would therefore not be surprising to find that the simple FM call structures of bowheads are optimized to increase detection range and recognition by minimizing distortions from propagation effects (e.g., interference nulls), while the bowhead's auditory sensory system is adapted to extract information from such propagation effects.

## ACKNOWLEDGMENTS

Funding for this study was provided by the State of Alaska through the Department of Natural Resources, U.S. Bureau of Indian Affairs, North Slope Borough, and National Marine Fisheries Service through the Alaska Eskimo Whaling Commission (AEWC). Dr. Tom Albert of the Department of Wildlife Management of the North Slope Borough provided significant guidance and inspiration for the project. We are very grateful to all the people of Barrow for their support and hospitality over the many years of research in their magnificent arctic homeland. Advice and assistance received from the AEWC and from members of the Barrow Whaling Captains Association were of enormous help in designing and carrying out these studies. Diana Dronenburg, Perry Anashugak, Phillip Feebeck, and Joe Burgener ably provided base camp communications and logistic support. Special thanks go to all the people who assisted in the field work: Dave Blomstrom, John Bower, Maria Brower, Thomas Calupca, Geoff Carroll, Ray Dronenburg, Adam Frankel, Christine Gabriele, John Craighead George, Mia

Grifalconi, Ruth Johnson, Brendan Kelly, Donna McDonald, Mary Moore, Lori Quakenbush, and Dave Smith. Kim Beeman, Tom Calupca, Sean Cunningham, and Steve Mitchell did a remarkable job of hardware and software engineering. Katherine Dunsmore, Carol Carson, Janet Doherty, and Russ Charif were extraordinarily diligent in the laboratory analysis and organization of the 1988 data set. Thomas Parks and Talal Shamoon from Cornell College of Engineering gave considerable guidance and advice leading to improvements in signal processing. Special appreciation is owed to the production and engineering staff at Sippican Ocean Systems, Inc. for their effort in producing the hydrophones and circuitry for the sonobuoys. Peter Tyack at WHOI graciously loaned the Crown preamplifier and 300-watt amplifier. We thank Charles Greene, Jr., and one anonymous reviewer for helpful and constructive comments.

- Altes, R. (1980). "Detection, estimation, and classification with spectrograms," *J. Acoust. Soc. Am.* **67**, 1232–1246.
- Charif, R. A., Mitchell, S., and Clark, C. W. (1995). *Canary 1.12 User's Manual* (Cornell Laboratory of Ornithology, Ithaca, NY), p. 229.
- Clark, C. W. (1982). "The acoustic repertoire of the Southern right whale, a quantitative analysis," *Anim. Behav.* **30**, 1060–1071.
- Clark, C. W., Charif, R., Mitchell, S., and Colby, J. (1996). "Distribution and behavior of the Bowhead whale, *Balaena mysticetus*, based on the analysis of acoustic data collected during the 1993 spring migration off Point Barrow, Alaska," *Rep. int. Whal. Commn.* **46**, 541–552.
- Clark, C. W., and Ellison, W. T. (1989). "Numbers and distributions of bowhead whales, *Balaena mysticetus*, based on the 1986 acoustic study off Pt. Barrow, Alaska," *Rep. Int. Whal. Commn.* **39**, 297–303.
- Clark, C. W., Ellison, W. T., and Beeman, K. (1986a). "Acoustic tracking of migrating bowhead whales," *IEEE Oceans '86 Conference Proceedings*, 341–346.
- Clark, C. W., Ellison, W. T., and Beeman, K. (1986b). "A preliminary account of the acoustic study conducted during the 1985 spring bowhead whale, *Balaena mysticetus*, migration off Point Barrow, Alaska," *Rep. int. Whal. Commn.* **36**, 311–316.
- Clark, C. W., and Johnson, J. H. (1984). "The sounds of the bowhead whale, *Balaena mysticetus*, during the spring migrations of 1979 and 1980," *Can. J. Zool.* **62**, 1436–1441.
- Clark, C. W., Marler, P., and Beeman, K. (1987). "Quantitative analysis of animal vocal phonology: An application to swamp sparrow song," *Ethology* **76**, 101–115.
- Cummings, W. C., and Holliday, D. V. (1985). "Passive acoustic location of bowhead whales in a population census off Point Barrow, Alaska," *J. Acoust. Soc. Am.* **78**, 1163–1169.
- Cummings, W. C., Holliday, D. V., Ellison, W. T., and Graham, B. J. (1983). "Feasibility of passive acoustic location of bowhead whales in population studies off Point Barrow, Alaska," unpublished report to the North Slope Borough, p. 169.
- DiNapoli, F. R., Nielson, R., Potter, D., and Stocklin, P. L. (1984). "TRISTEN/FRAM IV CW spatial coherence and temporal stability," Naval Underwater Systems Center Technical Document 7095, Newport, Rhode Island, p. 37.
- Ellison, W. T., Clark, C. W., and Bishop, G. C. (1987). "Potential use of surface reverberation by bowhead whales, *Balaena mysticetus*, in under-ice navigation: Preliminary considerations," *Rep. int. Whal. Commn.* **37**, 329–332.
- Frankel, A. S., Clark, C. W., Herman, L. M., and Gabriele, C. M. (1995). "Spatial distribution, habitat utilization, and social interactions of humpback whales, *Megaptera novaeangliae*, off Hawaii, determined using acoustic and visual techniques," *Can. J. Zool.* **73**, 1134–1146.
- Garrison, G. R. (1976). "Chukchi sea oceanography, 1975 measurements and a review of coastal current properties," Report No. 7614, Applied Physics Laboratory, University of Washington, Seattle, Washington, p. 78.
- George, J. C., Clark, C., Carroll, G. M., and Ellison, W. T. (1989). "Observations on the ice-breaking and ice navigation behavior of migrating bowhead whales (*Balaena mysticetus*) near Point Barrow, Alaska, Spring 1985," *Arctic* **42**, 24–30.
- Harrison, C., and Harrison, J. (1995). "A simple relationship between frequency and range averages for broadband sonar," *J. Acoust. Soc. Am.* **97**(2), 1314.
- Janik, V. M., Van Parijs, S. M., and Thompson, P. M. (2000). "A two-dimensional acoustic localization system for marine mammals," *Mar. Mamm. Sci.* **16**, 437–447.
- Ljungblad, D. K., Thompson, P. O., and Moore, S. E. (1982). "Underwater sounds recorded from migrating bowhead whales, *Balaena mysticetus*, in 1979," *J. Acoust. Soc. Am.* **71**, 477–482.
- Marple, S., Jr. (1987). *Digital Spectral Analysis with Applications* (Prentice-Hall, Englewood Cliffs, NJ), p. 492.
- McGregor, P. K., Dabelsteen, T., Clark, C. W., Bower, J. L., Tavares, J. P., and Holland, J. (1997). "Accuracy of a passive acoustic location system: Empirical studies in terrestrial habitats," *Ethol. Ecol. Evol.* **9**, 269–286.
- Mitchell, S., and Bower, J. (1995). "Localization of animal calls via hyperbolic methods," *J. Acoust. Soc. Am.* **97**, 3352–3353.
- National Oceanographic and Atmospheric Administration. (1976). "National ocean survey chart, Point Barrow and vicinity. No. 16082," 5th ed., U.S. Department of Commerce, Washington, DC.
- Naval Research Laboratory Instruction Book Number 151. (1975). *USRD Type J11 Audio Frequency Transducer*, Naval Research Laboratory, Underwater Sound Reference Detachment, Orlando, Florida, 1 May 1975.
- Payne, R., and Guinee, L. (1983). *Humpback whale (Megaptera novaeangliae) songs as an indicator of "stocks"* (Westview, Boulder, CO), pp. 333–358.
- Raftery, A. E., and Zeh, J. E. (1998). "Estimating bowhead whale population size and rate of increase from the 1993 census," *J. Am. Stat. Assoc.* **93**, 451–463.
- Raftery, A. E., Zeh, J. E., Yang, Q., and Styer, P. E. (1990). "Bayes empirical Bayes interval estimation of bowhead whale, *Balaena mysticetus*, population size based upon the 1986 combined visual and acoustic census off Point Barrow, Alaska," *Rep. int. Whal. Commn.* **40**, 393–409.
- Schevill, W. E., Watkins, W. A., and Backus, R. H. (1964). *Marine Bio-Acoustics*, edited by W. N. Tavolga (Pergamon, Oxford), pp. 147–152.
- 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**, 107–153.
- Thomas, J. A., Fisher, S. R., and Ferm, L. M. (1986). "Acoustic detection of cetaceans using a towed array of hydrophones," *Rep. int. Whal. Commn.* **8**, 139–148.
- Thompson, P. O., and Friedl, W. A. (1982). "A long term study of low frequency sounds from several species of whales off Oahu, Hawaii," *Cetology* **45**, 1–19.
- Watkins, W. A. (1976). "Biological sound-source locations by computer analysis of underwater array data," *Deep-Sea Res.* **23**, 175–180.
- Watkins, W. A. (1981). "Activities and underwater sounds of fin whales," *Sci. Rep. Whales Res. Inst.* **33**, 83–117.
- Watkins, W. A., and Schevill, W. E. (1972). "Sound location by arrival times on a non-rigid three dimensional hydrophone," *Deep-Sea Res.* **19**, 691–706.
- Winn, H. E., Edell, R. K., and Taruski, A. G. (1975). "Population estimate of the humpback whale (*Megaptera novaeangliae*) in the West Indies by visual and acoustic techniques," *J. Fish. Res. Board Can.* **32**, 499–506.
- Würsig, B., and Clark, C. (1993). "Behavior," *The Bowhead Whale* (The Society for Marine Mammalogy, Lawrence, KS), pp. 157–199.
- Zeh, J. E., Clark, C. W., George, J. C., Withrow, D., Carroll, G. M., and Koski, W. R. (1993). "Current Population Size and Dynamics," *The Bowhead Whale* (The Society for Marine Mammalogy, Lawrence, KS), pp. 409–489.



# Recognizing transient low-frequency whale sounds by spectrogram correlation

David K. Mellinger<sup>a)</sup>

Cooperative Institute for Marine Resources Studies, Oregon State University, 2030 South Marine Science Drive, Newport, Oregon 97365, Monterey Bay Aquarium Research Institute, P.O. Box 628, Moss Landing, California 95039, and Bioacoustics Research Program, Cornell University, 159 Sapsucker Woods Road, Ithaca, New York 14850-1999

Christopher W. Clark

Bioacoustics Research Program, Cornell University, 159 Sapsucker Woods Road, Ithaca, New York 14850-1999

(Received 25 January 1999; accepted for publication 26 February 2000)

A method is described for the automatic recognition of transient animal sounds. Automatic recognition can be used in wild animal research, including studies of behavior, population, and impact of anthropogenic noise. The method described here, spectrogram correlation, is well-suited to recognition of animal sounds consisting of tones and frequency sweeps. For a sound type of interest, a two-dimensional synthetic kernel is constructed and cross-correlated with a spectrogram of a recording, producing a recognition function—the likelihood at each point in time that the sound type was present. A threshold is applied to this function to obtain discrete detection events, instants at which the sound type of interest was likely to be present. An extension of this method handles the temporal variation commonly present in animal sounds. Spectrogram correlation was compared to three other methods that have been used for automatic call recognition: matched filters, neural networks, and hidden Markov models. The test data set consisted of bowhead whale (*Balaena mysticetus*) end notes from songs recorded in Alaska in 1986 and 1988. The method had a success rate of about 97.5% on this problem, and the comparison indicated that it could be especially useful for detecting a call type when relatively few (5–200) instances of the call type are known. © 2000 Acoustical Society of America. [S0001-4966(00)01706-9]

PACS numbers: 43.80.Lb, 43.80.Ka [WA]

## INTRODUCTION

Automatic recognition of animal vocalizations is promising as a tool for investigating free-ranging animals in their natural habitats. It is also challenging as a problem in signal processing, since difficulty arises from the nonstationary nature of the signals involved; from the highly variable nature of animal sounds at the individual, intraspecific, and interspecific levels of analysis; and from the characteristics of the noise environments encountered in field recordings.

Primary applications of acoustical methods for studying marine animals include species range and distribution surveys (Clark and Mellinger, 1994; Clark and Fristrup, 1996; Moore *et al.*, 1998; Stafford *et al.*, 1998), behavior studies (Buck and Tyack, 1993; Frankel *et al.*, 1995), population measurement (Leaper *et al.*, 1992; Zeh *et al.*, 1993; Raftery and Zeh, 1998), and studies of the impact of noise on animals (Clark *et al.*, 1998). The more traditional methods of approaching these problems involve visual observation, but acoustic recognition methods have certain advantages over visual ones, especially when augmented with techniques for tracking animals in their native habitat (Watkins and Schevill, 1972; Clark *et al.*, 1986; Clark, 1989; Clark *et al.*, 1991, 1996; McDonald *et al.*, 1995). Some environments, including dense forests and oceans, are inaccessible or inhospitable

to visual observation but more accommodating to acoustic methods. Visual observation is possible only during daylight hours, while acoustic methods can be used for species that are active anytime during the 24-hour day. Such species include ones active nocturnally (Griffin, 1964; Graber, 1968; Larkin, 1978; Terres, 1980) or throughout the 24-hour day, such as right whales (*Eubalaena* sp.), bowheads, and fin whales (*Balaenoptera physalus*) (Clark, 1983; Watkins *et al.*, 1987; Würsig and Clark, 1993). With automated recognition methods, extensive sound recordings can be made and processed relatively quickly, while visual observation is usually very labor-intensive and requires trained observers. Conditions may make visual observation difficult, as in the case of fog, or ice cover in the Arctic. For many species, especially ones in the marine environment that rely heavily on the acoustic modality for communication and information gathering, acoustic methods may offer more appropriate insights into the biology of the animals than visual methods.

Acoustic recognition can be done by persons with trained ears, but automated methods have certain advantages. They are unbiased, or rather their bias is constant rather than possibly changing from time to time and place to place. They can be used to process large amounts of data; this is quite important in field work, where thousands of hours of sound may need to be analyzed. In real-time processing situations, automated methods can operate on many sound channels at once, allowing simultaneous monitoring of several widely

<sup>a)</sup>Electronic mail: mellinger@pml.noaa.gov

distributed different locations, or of sounds from the same source received at several different sensors. Sounds above or below the frequency range of human hearing can be processed, as can sounds that change too quickly or too slowly for humans to hear clearly. Automatic methods may also be relatively inexpensive for long-term monitoring.

A variety of techniques has been used for automatic recognition of animal calls. Spectrogram matched filtering, or cross-correlation of the spectrograms of a known sound and an unknown recording, has been used for classifying and comparing animal sounds (Clark *et al.*, 1987; Chabot, 1988). A variant of matched filtering uses synthetic waveforms (Stafford *et al.*, 1998) or synthetic spectrograms (Mellinger and Clark, 1993) instead of sounds edited out of recordings. Another approach is to measure a number of characteristics from the sound and use these in a statistical classifier (Fristrup, 1992; Pinkowski, 1994); this method has been successful at recognizing calls in a large database of marine mammal sounds (Fristrup and Watkins, 1994). The frequency contour of a vocalization can be tracked (Newman *et al.*, 1978; Clark, 1982; Goedecking, 1983; Buck and Tyack, 1993; Nyamsi *et al.*, 1994) for comparison to other contours. Neural networks, typically using spectrogram values as input, work well in many cases for detecting the sounds of a species of interest (Ramani and Patrick, 1992; Gaetz *et al.*, 1993; Moore *et al.*, 1991; Patrick *et al.*, 1994). Speech methods (Rabiner and Juang, 1993), in particular hidden Markov models, have been applied as well (Weisburn *et al.*, 1993; Sturtivant and Datta, 1997). Methods employing machine learning in combination with statistical feature extraction have been used successfully for monitoring frog sounds in Australia (Taylor, 1995; Taylor *et al.*, 1996), and a variety of *ad hoc* signal-processing systems has been used as well (Whitehead and Weilgart, 1990; Leaper *et al.*, 1992).

What characteristics of animal vocalizations make the automatic recognition problem tractable? Many vocalizations have evolved to be heard against background noise, and thus contain changing parameters that make them audible to the perceptual systems receiving them. One type of change is frequency modulation (FM); that is, a change in a signal's narrow-band, instantaneous frequency over time. Narrow-band FM signals, also known as frequency sweeps, are found in the vocalizations of many animals. Many acoustic signals of primates incorporate FM (Moody and Stebbins, 1989), as do those of birds, whose long-range advertisement songs "almost invariably consist of frequency-modulated tones" (Wiley and Richards, 1982). Most cetacean species, comprising the toothed or odontocete whales and the baleen or mysticete whales, produce a great variety of sounds, covering an extensive frequency band and including both FM signals and pulse sequences (Schevill, 1964; Thompson *et al.*, 1979; Herman and Tavolga, 1980; Clark, 1990).

Whales of the mysticete group, which includes only 11 species, show a wide range of inter- and intraspecific variability in their vocalizations. Within the more accessible coastal species, the sounds from bowhead, humpback (*Megaptera novaeangliae*), and right whales are the best documented, and all produce a wide variety of FM and pulsed sound structures (Schevill and Watkins, 1962; Payne

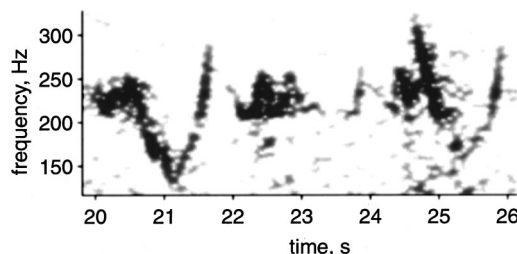


FIG. 1. Three successive bowhead song end notes from a song of one whale, recorded in Alaska in 1986. The first shows the frequency contour typical of end notes; the other two show typical variation. Spectrogram parameters for all figures: Sampling rate 2023 Hz, frame size 512 samples (zero-padded to an FFT size of 1024 samples), filter bandwidth 16 Hz, frame increment 32 ms, Hamming window.

and McVay, 1971; Payne and Payne, 1971; Clark, 1981, 1982; Payne *et al.*, 1983; Clark and Johnson, 1984; Chabot, 1984; Silber, 1986; Clark, 1991). In contrast, within the *Balaenoptera* genus that includes the blue (*Balaenoptera musculus*), fin, and minke (*B. acutorostrata*) whale species, signal documentation has been more limited since access has been difficult (Schevill *et al.*, 1964; Cummings and Thompson, 1971; Winn and Perkins, 1976; Edds, 1982; Watkins *et al.*, 1987; Thompson *et al.*, 1996; Rivers, 1997; Stafford *et al.*, 1999).

The method described here for recognizing animal vocalizations focuses on finding frequency sweeps. The method was motivated by psychoacoustic work showing the salience of frequency sweeps in human hearing (Kay and Matthews, 1972), and by neurophysiological studies showing that there are neurons in the mammalian auditory system selectively responsive to frequency sweeps (Whitfield and Evans, 1965; Møller, 1977; Mendelson and Cynader, 1985; deCharms *et al.*, 1998). The method's image-processing technique is inspired by vision research that has identified neurons responsive to certain orientations of lines in the visual field (Hubel and Wiesel, 1962, 1977), analogous to certain orientations of frequency sweeps in the auditory field.

In this paper, we describe a specific call-recognition problem, present a solution in the form of a method for recognizing frequency sweeps, describe ways to tune parameters to make the method most effective, test the method on sets of sample data, compare results to those from other methods for animal sound recognition, and offer some improvements that raise performance of the method.

## I. PROBLEM STATEMENT

The specific problem addressed here is the recognition of the end notes of bowhead whale songs. These notes are portions of a bowhead's song that occur one or more times in succession at the end of each song repetition, and are distinctly different from the preceding portions of the song. Songs by one individual typically last 60–70 s, with pauses of 5–15 s between songs (Würsig and Clark, 1993). Bowhead song end notes were chosen because they are relatively loud, and they typically occur several times per song (Clark, 1991). Figure 1 shows, in the leftmost of the three calls, a typical end note from the 1988 song for the Bering-Chukchi-Beaufort Sea stock of the bowhead whale, as well as two

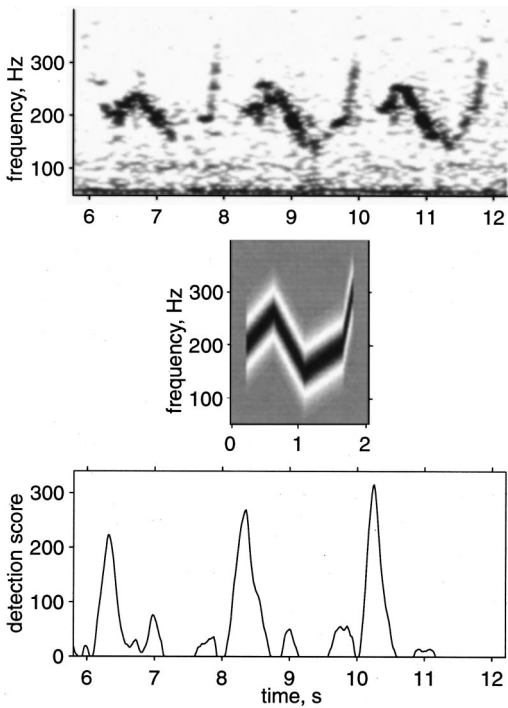


FIG. 2. Spectrogram correlation example. (a) Spectrogram of a novel recording containing three bowhead song end notes. (b) Spectrogram correlation kernel, with black representing positive areas, white representing negative areas, and gray representing the zero background. (c) The resulting recognition score function. The units of the recognition score function are arbitrary.

other end notes illustrating the amount of internote variation that can be present in sounds recorded from one whale. Recognizing sounds from bowhead whales provides a signal-processing challenge in that the sonic environment of the recordings contains many types of interfering noise, both stationary and transient.

## II. METHOD

The method for recognizing animal vocalizations is described here in general terms, with specific application to the recognition of bowhead song end notes to make implementation of the fundamental technique concrete. The method operates on spectrograms computed from the time-series waveform of a sound (Oppenheim and Schaffer, 1975; Altes, 1980). Examples of the sound type (the bowhead song end note, in this case) are used to construct a correlation kernel for the vocalization. To recognize the vocalization of interest in a novel recording, a spectrogram of the recording is made and cross-correlated with the kernel representing the signal of interest. The result is a recognition function—a time series of recognition values—that represents the closeness of the match between the kernel and the novel recording at each time increment of the spectrogram. Larger values in the time-varying recognition function represent higher likelihood that a bowhead song end note was present in the novel recording. An example of the process is shown in Fig. 2.

To use this method, a correlation kernel is made for the vocalization type of interest, as follows. A set of training examples of the vocalization of interest is obtained. Spectrogram time/frequency parameters are chosen that make the

TABLE I. Measured mean ( $\pm$ s.d.) characteristics of the bowhead song end notes from 1986; values were measured in a spectrogram computed with the same parameters as Fig. 1. The bandwidth value is an approximate estimation (see the text). The mean values were used for constructing the recognition kernel  $k$  containing four contiguous segments: an initial frequency upswing, a downswing, and two successive upsweeps. Since the frequency sweeps in this case were contiguous throughout the vocalization, the end frequency of each section is the same as the start frequency of the next section.

Section no.	Duration, s	Start freq., Hz	End freq., Hz	Bandwidth, Hz
1	$0.395 \pm 0.079$	$197.6 \pm 9.3$	$257.2 \pm 15.1$	50
2	$0.458 \pm 0.110$	$257.2 \pm 15.1$	$146.4 \pm 12.2$	50
3	$0.557 \pm 0.139$	$146.4 \pm 12.2$	$196.7 \pm 15.5$	50
4	$0.143 \pm 0.105$	$196.7 \pm 15.5$	$295.7 \pm 19.4$	50

vocalization's frequency contours clearly visible in the spectrogram. Bradbury and Vehrencamp (1998) further discuss choice of spectrogram parameters for animal vocalizations. Our experience has been that parameters that reveal frequency contours for visual inspection also work well for spectrogram correlation. Next, the vocalization is divided into *sections* in which the FM rate is relatively constant. The kernel is made up of a series of *segments*, where each segment corresponds to one section of the vocalization. For instance, the first bowhead song end note of Fig. 1 has four sections: a short initial FM upswing, a downswing, an upswing, and another, steeper, upswing. The corresponding kernel made for this song end note therefore would have four segments.

To construct a kernel for the vocalizations in the training set, time and frequency characteristics of the sections that make up each sample vocalization are measured. The start time of the vocalization is set to zero. For each section in the vocalization, the times and frequencies of the section's endpoints are measured. This measurement may be done using widely available computerized spectrogram-display tools, such as CANARY (Charif *et al.*, 1995), OSPREY (Mellinger, 1995), or SIGNAL (Engineering Design, 1997). Also, the bandwidth of a each section is measured—the difference between the upper and lower edges of the section visible in the spectrogram at one instant. The measurements for each endpoint are averaged for all sample vocalizations to obtain characteristics of an "average call" that is used to make up the kernel.

For recognizing bowhead song end notes, songs from two whales recorded 5 days apart were used, and five end notes from each of four total songs were selected, for a total of 20 training notes. The endpoints and bandwidths of the four sections in each note were measured, then averaged. These averaged values are shown in Table I.

The kernel is made up of several segments, one per FM section in the target vocalization type. Each kernel segment is a two-dimensional array of values in the time–frequency space of spectrograms; it contains positive- and negative-valued regions. The kernel for the bowhead song end notes is shown in the center of Fig. 2. Its positive region extends between the averaged start and end frequencies of the measured FM sections, has a duration equal to the average duration of the sections, and has a frequency spread approxi-

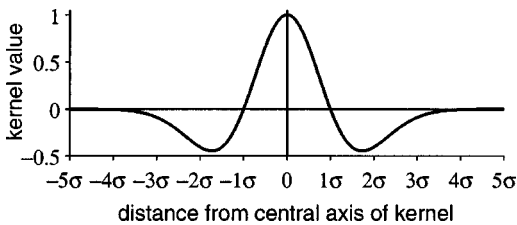


FIG. 3. Values in a vertical cross section of the kernel segment of Fig. 2. The horizontal axis is the distance from the axis of the FM sweep in the kernel, in units of the spread factor  $\sigma$ . The vertical axis is the kernel value, displayed here as a curve instead of gray-scale values as in Fig. 2.

mately equal to the bandwidth of the measured sections. Each segment also has two negative regions, at higher and lower frequency bands above and below the central positive region, respectively. These negative regions aid in the rejection of noise sounds. These are shown in the kernel of Fig. 2 as lighter areas. The shape of the positive and negative areas in the kernel is defined as a hat function, the second derivative of a Gaussian distribution as shown in Fig. 3. The kernel value  $k$  at a given time and frequency point  $(t, f)$  is specified by

$$x = f - \left( f_0 + \frac{t}{d}(f_1 - f_0) \right), \quad (1)$$

$$k(t, f) = \left( 1 - \frac{x^2}{\sigma^2} \right) \exp\left( -\frac{x^2}{2\sigma^2} \right), \quad (2)$$

where  $x$  = distance of this  $(t, f)$  point from the central axis of the segment at time  $t$ ,  $f_0$  = start frequency of the segment,  $f_1$  = end frequency of the segment,  $d$  = duration of the segment, and  $\sigma$  = instantaneous bandwidth of the segment at time  $t$ .

The positive and negative values of a segment sum to zero. This is important for rejection of noise and of interfering sounds. When a spectrogram of uniform background noise is made and cross-correlated with such a zero-sum kernel, the result is zero, since the noise aligns with equal regions of positive and negative values of the kernel. Similarly, if a frequency sweep of a different FM sweep rate ( $df/dt$ ) is cross-correlated with this kernel, the sweep will intersect both positive and negative regions of the kernel, producing a summed output response of zero. Other types of noise such as wideband clicks and pops, commonly produced by both cracking ice and recording glitches, also intersect positive and negative regions and produces a summed output response of zero. As the next section shows, this characteristic has proven effective at rejecting many types of interfering noise—a useful feature in the acoustically cluttered arctic environment of the bowhead call-recognition example. A further feature is that many noisy sounds can co-occur with bowhead song end notes and not affect the correlation output, because the correlation operation is linear and the interfering sounds simply add zero to its output. Of course, if an interfering noise has frequency sweeps that closely match the sound of interest (the bowhead song end note, in our example), then the correlation value will be undesirably large.

The value of the bandwidth parameter  $\sigma$  may be varied to accommodate a greater or lesser amount of variation in the

instantaneous bandwidth of frequency sweeps. Greater kernel bandwidth results in a less discriminatory recognition function that includes more types of frequency sweeps. That is, a kernel segment constructed with a large  $\sigma$  value, when cross-correlated with a wide variety of input frequency sweeps, will produce relatively high recognition values, while small bandwidths in the kernel segment will produce equivalent recognition values only for frequency sweeps that are more exactly aligned with the kernel axis. As a rule of thumb,  $\sigma$  values of  $0.7 \pm 0.2$  times the instantaneous bandwidth of the actual frequency sweep have been found to work well. Since small variations ( $\pm 10\%$ ) in  $\sigma$  were found to have little effect on recognizer performance, exact measurement and averaging of the instantaneous bandwidth of many frequency sweeps was not necessary. However, different choices of spectrogram parameters, especially fast Fourier transform (FFT) size and window function, will lead to different instantaneous bandwidths of frequency sweeps, and will require different values of  $\sigma$ .

To produce the correlation kernel, several segments are concatenated. Given the kernel  $k(t, f)$  and the spectrogram, the output of the cross-correlation process for each frame of the spectrogram is given by

$$\alpha(t) = \sum_{t_0} \sum_f k(t_0, f) S(t - t_0, f), \quad (3)$$

where  $S(t, f)$  is the spectrogram, and the limits of summation are specified by the size of  $k$  in time and frequency. Note that the correlation is in the time dimension only, not in frequency as would be used for two-dimensional cross-correlation. This one-dimensional correlation is significantly faster than a two-dimensional one, and works well provided that the vocalizations to be recognized are not frequency shifted (see Clark *et al.*, 1987). This is true of bowhead whale vocalizations, and of many other species' sounds as well.

The result of the cross-correlation operations is a time series with a "recognition score"  $\alpha(t)$  for each point in time, i.e., a value representing the closeness of match between the bowhead song end note and the kernel. The bottom of Fig. 2 shows an example of the recognition score function. This function,  $\alpha(t)$ , is not strictly a probability, as it does not vary between only 0 and 1, but its value is near 0 when a bowhead vocalization is absent and increases in value when a vocalization is present. A zero minimum is applied to the recognition score, that is, negative values are changed to zero. Because spectrogram levels are not normalized before the cross-correlation is computed, the maximum value of  $\alpha(t)$  is arbitrarily large. Also, because of the definition given by Eq. (3), the time in the recognition function when the peak occurs is at the very beginning of each song end note.

The recognition score function  $\alpha(t)$  may be turned into a sequence of discrete *detection events*—occasions when the sound of interest is determined to be present—by setting a threshold and registering an event each time the score goes above the threshold.

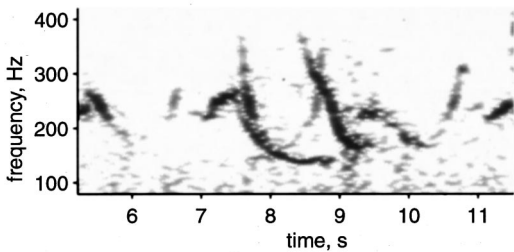


FIG. 4. Bearded seal sounds with some similarities to bowhead song end notes. The two prominent curving frequency downsweeps from 7.5–9 s and from 8.5–9.5 s are bearded seal sounds. The first of these has a frequency contour similar to sections of the end notes of Fig. 1.

### III. RESULTS

#### A. Qualitative analysis

As part of an acoustic census program, bowhead whales were recorded in Spring 1986 and 1988 off Pt. Barrow, Alaska during their seasonal migration (Clark and Ellison, 1989; Clark *et al.*, 1996; Clark and Ellison, in press). Sounds were received by hydrophones placed along the edge of the shorefast ice in water 15–20 m deep, transmitted by FM radio to a field recording hut, and recorded with a Teac R-61D cassette data recorder. The recordings were later brought into the laboratory and digitized by a Teac RD-135T digital tape recorder for computer analysis. The recording systems' frequency responses were flat in the 14–5000 Hz frequency band.

The 1986 recordings contain bowhead songs that include the end notes of the type shown at the left in Fig. 1. The recordings also contain a variety of types of noise in the same frequency range as the bowhead song end notes: hydrophone cable flutter, wind and wave noise, ice noise (including squeals, creaks, and pops), and vocalizations of beluga whales (*Delphinapterus leucas*), bearded seals (*Erignathus barbatus*), and other bowhead whales.

A kernel was designed for the bowhead song end notes using the methods described above, and the spectrogram-correlation operation of Eq. (3) was computed. In preliminary tests, the spectrogram correlation method worked well at recognizing bowhead song end notes, producing high recognition scores when a note was present and low values at other times. Sometimes notes were missing one or more FM sections, in which case the recognition score was lower but still high relative to background noise.

The rating of a recognition score as either "low" or "high" depends on the size of the kernel, the gain of the sound acquisition system, and other factors. They are essentially arbitrary values that can be easily determined by inspection of the detector output. In this case a low score is one below 50, and a high score one above 100. Recognition scores, as expected, were low for white and colored background noise. The scores were also low for most types of interfering sounds, including hydrophone cable flutter, ice sounds, and most other marine mammal sounds. A few bearded seal sounds did result in high values ( $\approx 150$ ) of the spectrogram correlation in the correlation output. These sounds, two of which are shown in Fig. 4, had frequency contours similar to those of bowhead song end notes. This

result is not altogether surprising, as human listeners often confused this type of seal sound with bowhead sounds.

The handling of variation in frequency contours is discussed more extensively and quantitatively in Sec. IV, below. Qualitatively, the method gave reliable results for bowhead song end notes that contained some variation in the frequency contour, but not a large amount. The amount of frequency variation that could be accommodated depended primarily on the bandwidth  $\sigma$  of the kernel segments as discussed above. As expected, variation in the duration of an end-note section, or in the timing of different sections, caused some problems for the method. A small amount of timing variation was accommodated by the nature of the up- and downsweeps of the kernel. For example, even when one of the FM sections in a bowhead song end note was slightly delayed, it still intersected part of the corresponding positive segment of the kernel.

#### B. Quantitative comparison

To test the spectrogram-correlation method, its performance was compared against that of a matched filter and a hidden Markov model (HMM) using a set of sample data. The sample data set for the comparison was selected by examining spectrograms of the Pt. Barrow recordings to determine when bowhead song end notes were present. Recordings were divided into short samples of a few seconds duration, with each sample containing a bowhead song end note. A total of 114 samples was selected this way.

Matched filtering is the optimum method for detecting a known signal in white Gaussian noise (van Trees, 1968). When detecting animal sounds, however, the signal is not known precisely since it varies from one occurrence to the next, and the background noise rarely if ever has a flat spectrum, so a matched filter is not necessarily the optimal solution. To compute the matched filter kernel, the time-series waveforms of several high-quality (good signal-to-noise ratio) notes were selected. Each time series was normalized to have zero mean and unit energy, and all time series were placed in a matrix. A sample covariance matrix was calculated, and the eigenvector corresponding to the maximum eigenvalue was computed. This calculation corresponds to a principal components projection of the notes (Jain and Dubes, 1988). The resulting time-series kernel was then cross-correlated with each of the test sound samples to produce output functions. As with the spectrogram kernel-correlation method, the maximum correlation value for a given sound sample was used as the recognition score.

HMM techniques have been used widely in speech recognition (Rabiner and Juang, 1986; Lee *et al.*, 1990; Rabiner and Juang, 1993) and it was thought that an HMM would perform well for the task of recognizing the complexity of bowhead song notes. The HMM used has been described elsewhere (Weisburn *et al.*, 1993); briefly, it was trained by picking peaks in several example calls and using their frequencies to trigger state transitions in the model.

The spectrogram correlation, matched filter, and HMM methods each produced a recognition score for each of the 114 sounds in the sample set. In order to determine if the score would be considered a detection event, it was neces-

TABLE II. Comparison of three recognition methods for a set of 114 bowhead song end notes recorded in 1986. The detection thresholds were set as described in the text.

Method	Correct detections	Error rate
Matched filter	96	15.8%
Hidden Markov model	111	2.6%
Spectrogram correlation	113	0.9%

sary to choose a detection threshold for each of the three methods, such that scores above the threshold would be considered a detection event.

The choice of a detection threshold depends on the application. Choosing a low detection threshold causes few localizations to be missed, but can also cause more background noise sounds to be falsely “detected.” Choosing a high threshold eliminates some or all of the false detections, but can also cause some actual vocalizations to be missed. If the aim is to scan a large body of recordings looking for a vocalization that may possibly be present, then choosing a low detection threshold makes sense, so that the rare sound is not missed. Any sounds detected can be re-examined later to see if they are the vocalizations of interest. Alternatively, automatic recognition can be used in an acoustic census to count the number of calls present in a set of recordings—typically, a large number of calls. In this case, missing some fraction of the total calls is not a severe drawback, as long as the fraction missed is known. Counting background noise sounds as calls, however, can affect the census result, so a high detection threshold is more appropriate.

For the problem of recognizing bowhead song end notes, it was decided to use a relatively low threshold, so that 10% of noise samples would be detected as calls. Thresholds for each of the methods were chosen by synthesizing 114 white-noise sound samples (generated using a normally distributed pseudorandom number generator), using them as input to the algorithms, and setting the threshold for each method so that 10% of the white-noise sounds were detected as calls and the remaining 90% rejected.

Using these detection thresholds, the 114 bowhead calls were processed by each of the recognition methods. The numbers of detection events for each method are shown in Table II. These results show that the matched filter worked poorly at recognizing bowhead song end notes, missing 18 of them. In contrast, the HMM worked fairly well, missing only 3 of the 114 notes, and the spectrogram correlator worked best of all, missing only 1.

### C. Quantitative comparison to a neural network

Performance of the spectrogram-correlation method was compared to that of a neural network. For this comparison, a more extensive set of bowhead song end note recordings, made at Pt. Barrow in 1988, was used. Recording methods and equipment were similar to those in 1986, and the origins of noise were the same. This data set contains end notes from seven whales. Songs were identified as coming from an individual by acoustic tracking and by the timing of repeated songs. Spectrograms of the recordings were again examined to determine when bowhead song end notes were present, and recordings were edited to produce a set of 588 samples, each one 3.5 s in duration. Test samples that did not contain calls were also obtained; this was accomplished by editing out 3.5-s samples from Pt. Barrow recordings which did not contain bowhead sounds. A total of 888 noise samples containing ambient background noise, bearded seal sounds, ice noise, hydrophone flutter, and other sounds was obtained by this procedure.

Neural networks (Hebb, 1949; Rumelhart *et al.*, 1987) are often used in pattern-recognition systems where flexibility in the recognition of highly variable signals is required (Lippman, 1989; Ghosh *et al.*, 1992). The neural network architecture used for this comparison has been described elsewhere (Potter *et al.*, 1994). Briefly, it was a nonlinear network trained by backpropagation with momentum; the learning rate and tolerance values were decreased throughout the training period. The input layer was a spectrogram computed with a frequency range of 63–700 Hz, a filter bandwidth of 63.5 Hz, a duration of 2.7 s, and a time resolution of 0.128 s. A few selected pixels were deleted from the 11 × 21 spectrogram grid for a total of 192 input elements. The network’s hidden layer had four units, and the output layer had a single unit. This output unit’s value was compared to a threshold to determine whether or not a detection event had occurred. For training the network, each bowhead sound sample was time-aligned; noise samples were assigned a random time alignment. Approximately half of the data set was used for training, with the other half used as test data.

For an additional comparison, a matched filter was used to detect the bowhead song end notes as well. The matched-filter kernel for the 1988 end notes was constructed by the same method described above for the 1986 data.

The spectrogram correlation kernel was designed using the techniques described previously. Bowhead song end notes from 1988 had only three FM sections that were consistently present: a downsweep, followed by an upsweep, followed by a downsweep. Time and frequency values were measured for 48 calls from the training set. Time and fre-

TABLE III. Characteristics of the kernel used for recognizing bowhead song end notes from 1988. The kernel contains three segments, representing an initial FM downsweep, an upsweep, and a final downsweep. The FM sections in this case were not contiguous.

Section no.	Start time, s	End time, s	Start freq., Hz	End freq., Hz	Bandwidth, Hz
1	0	0.580	439.5	156.7	50
2	0.862	1.297	178.1	616.3	50
3	1.693	2.106	637.7	260.0	50

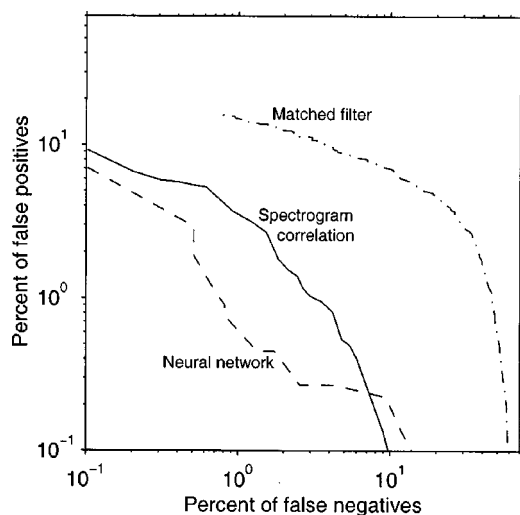


FIG. 5. Recognition error rates for the methods of spectrogram correlation, matched filter, and neural networks. Since error rates are plotted, curves lower and farther left represent better recognition performance.

quency averages are shown in Table III, along with the bandwidth used.

For this comparison many threshold values were used, instead of a single value as was done with the 1986 data. With a given setting of the threshold, two types of errors are possible. If a bowhead song end note is present in a sound and a recognition algorithm does not produce a recognition score above the threshold (and hence does not register a detection event), a “false negative” error occurs. If an end note is not present but the maximum recognition score is above threshold, a “false positive” error occurs. The false-positive and false-negative error rates are inversely related: if the threshold is raised, the number of false positive errors declines, but the number of false negative errors rises, and conversely for lowering the threshold. By setting the threshold to a series of increasing values, a succession of false-positive and false-negative error rates are obtained. The false-negative error rate may be plotted against the false-positive rate. Results of running the three recognizers—the spectrogram correlator, the matched filter, and the neural network—are shown in Fig. 5. In this figure, smaller error rates are represented by points toward the left and bottom of the plot.

As the figure shows, the matched-filter method worked only marginally well. The combined error rate, defined as the minimum of the sum of the false-positive and false-negative error rates, was 13.4% (9.0% false positive+4.5% false negative), a value too high for the matched filter to be of practical value.

The spectrogram-correlation method worked significantly better than the matched filter, perhaps well enough to make the method practical for some applications. The combined error rate was 3.6% (1.8% false positive+1.8% false negative). At this rate, the method could be used to screen long recordings for calls of interest.

The neural network had better performance, achieving a combined error rate of 1.6% (0.8% false positive+0.8% false negative). This error rate makes it the most accurate of the three methods discussed here for recognizing bowhead song end notes. As discussed below, however, there are oc-

casions when spectrogram correlation is the method of choice.

#### D. Computational issues

The most computationally intensive automatic call-recognition task is detecting calls of interest in a large body of sound. For instance, the recognition system discussed above could be employed to detect bowhead whale song end notes in several months of continuous recordings spanning one field season. High speed is important for such intensive applications.

How does spectrogram correlation compare in speed to other methods? As an example, in operation on the set of bowhead end notes from 1988, spectrogram correlation ran approximately 2.6 times as fast as the matched filter, and about the same speed as the neural network. (These times are for operation only, not including training.) The higher speed of spectrogram correlation and neural network probably occurred because their slowest step is calculating FFTs in making a spectrogram. Further steps involving dot products of two-dimensional images (the spectrogram correlation kernel or the input layer of the neural network) actually involved relatively few spectrogram cells in this case and were faster than the calculation of the spectrogram. Matched filtering also requires an FFT (and inverse FFT), but on a much longer sequence of sound samples, which have a much higher sampling rate than spectrogram frames. Spectrogram correlation and neural networks will slow down as the size of the kernel increases, but in use on many different projects to date, its speed has always been comparable to or faster than matched filtering.

### IV. TIME VARIATION

#### A. Problem description

One way in which the spectrogram-correlation method fails to work effectively is in handling timing variations within animal calls. Figure 1 shows bowhead song end notes that illustrate the large range of variation present in the data set. Some better means of accommodating the variation between calls might improve the performance of the automatic call recognizer. A similar problem exists in speech-recognition research, and several methods have been used to address it, including HMMs and dynamic time warping (Rabiner and Juang, 1993). The method used here (Mellinger, 1993) is a variant of dynamic time warping; it is described for a three-part bowhead song end note, but may be extended in an obvious way to sounds with any number of parts.

A call type is modeled, as above, as a sequence of separate sections, but now each section is separated from the next by an amount of time described by a probability distribution. For instance, the bowhead song end notes from 1988 can be modeled as three sections, as described previously. We now relax the assumption that the three sections occur in immediate succession, and instead describe them as separate FM sections occurring at different points in time. The amount of time between section  $i$  of the call and section  $i + 1$  is speci-

fied by  $\lambda_i(t)$ , a continuous distribution representing the probability that the two sections occur separated by time  $t$ .

To make a recognizer for a call type modeled this way, first recognizers for each separate section are constructed. Any type of recognizer works for processing these separate sections, provided that its output is a recognition score as a function of time. Here, a spectrogram correlator is assumed; constructing a spectrogram correlator amounts to making its kernel by the method outlined previously. Processing the training samples with these separate recognizers produces several recognition functions for each sample, one per section. Next, the time-delay probability distributions  $\lambda_i$  are estimated by measuring, for each training sample, times between the peaks of the recognition functions for the segments in the sound sample. The result is a set of time delays, one per training sample, that is used to estimate the time-delay probability functions. Typically, these time-delay probability estimates have a shape similar to a bell curve, and the distributions are therefore modeled as Gaussian.

The single-section recognizers and the time-delay probability functions make up the specification of the recognizer for the whole call type. After the separate section recognizers are run, the output functions, one per section, and the time-delay probability distributions are combined in a way described below to produce a single output recognition function.

To make a recognizer for the 1988 bowhead song end notes, three recognition kernels were made, one for each of the three sections. The kernels were constructed in the usual way, using measured time-frequency points from the training data of 48 end notes and zero-sum excitatory/inhibitory functions. Each input sound was correlated with the kernels as in Eq. (3), to produce three recognition functions  $\alpha_1(t)$  through  $\alpha_3(t)$ . Figure 6 shows an example of the three single-segment kernels and associated recognition scores.

Note that in  $\alpha_1(t)$ , the function computed by spectrogram correlation with the first kernel, the peak occurs early in the end note, at about 0.85 s, where the first FM section of the bowhead sound begins. The peak in  $\alpha_2(t)$  occurs later, at about 1.75 s, and the peak in  $\alpha_3(t)$  occurs still later, at about 2.5 s. The differences between the times of these successive peaks are examples of the measurements used to estimate the distributions  $\lambda_1(t)$  and  $\lambda_2(t)$ . These distributions were estimated from measurements of the trial set of 48 end notes. A histogram of the measured times between the initial upswing and the downswing—a sampling of the probability distribution  $\lambda_1(t)$ —is shown in Fig. 7. The assumption was made that the distribution was Gaussian and could be described by its mean and variance. This assumption was not required—the method will work with any distribution—but if the type of distribution is known, far fewer samples are needed to estimate it.

The recognition functions  $\alpha_i(t)$  are combined to make a single master recognition score  $\alpha$

$$\alpha = \max_{t_1, t_2, t_3} (\alpha_1(t_1) + \lambda_1(t_2 - t_1)\alpha_2(t_2) + \lambda_2(t_3 - t_2)\alpha_3(t_3)). \quad (4)$$

Every possible time of occurrence  $t_i$  of each section of

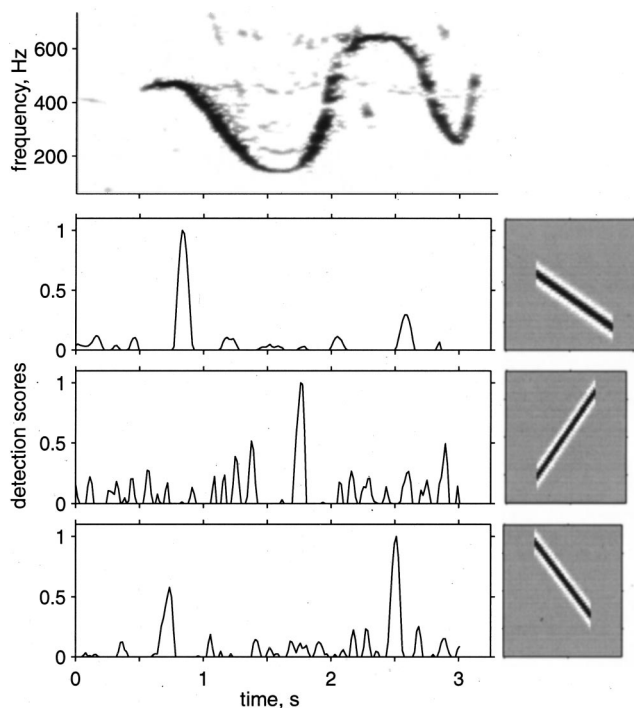


FIG. 6. Bowhead song end note from 1988, with (a) spectrogram and spectrogram correlator recognition functions and kernels for (b) the initial down-sweep, (c) the middle up-sweep, and (d) the final down-sweep. Note the progressive delays in timing between the peaks in (b) and (c), and the peaks in (c) and (d). Recognition functions are scaled to have a maximum value of 1. In the images, the axes of the four kernels are not precisely parallel to the axes of the frequency sweeps in the bowhead song end note because the parameters used to construct this kernel were averaged over many bowhead song end notes.

the bowhead song end note is used in this maximization. In it, the products of the recognition score for that section at that time [ $\alpha_i(t)$ ] and the probability of the associated time delay [ $\lambda_i(t)$ ] are computed. These products are summed, and the maximum of all such sums of products is taken as the overall recognition score.

The combination function  $\alpha$  has the characteristic that whenever the separate recognition functions  $\alpha_i(t)$  have peaks separated by the correct time delays, then  $\alpha$  has a large value. If the recognition functions  $\alpha_i(t)$  do not have large values, or if the delays between their large values differ from the high values in the distributions  $\lambda_i(t)$ , then  $\alpha$ 's value is small. Equation (4) defines  $\alpha$  for a three-section call type, but it can easily be extended to sounds with any number of sections. Essentially,  $\alpha$  is large when sections of the right

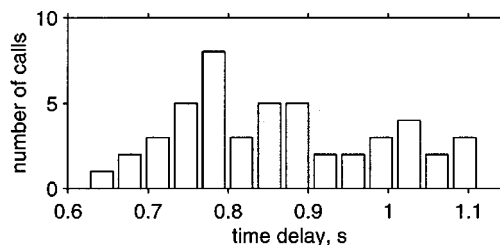


FIG. 7. Histogram of time delays between the first two sections of 1988 bowhead song end notes, as measured from the 48 end notes used for training.



type occur [the  $\alpha_i(t)$  are large], and they occur separated by the right time delays [the  $\lambda_i(t)$  are large].

As before, a threshold may be applied to this recognition score to determine detection events and then the associated false-positive and false-negative error rates.

## B. Computational issues

Equation (4) represents a three-dimensional optimization problem, one that can be quite computationally expensive. Furthermore, the cost increases exponentially with the number of FM sections in the sound type, so that sound types only a bit more complex are very computationally expensive. This cost could be a problem, since one of the uses of automatic recognition methods is real-time monitoring of animal sounds. Fortunately, the three-dimensional calculation of Eq. (4) can be simplified to two two-dimensional calculations. In general, an  $n$ -dimensional calculation similar to Eq. (4) is reduced to  $n-1$  two-dimensional calculations, greatly easing the computational burden.

To do this, note that the variable  $t_1$  in Eq. (4) is used in only the first two terms. Since these terms are summed in computing the maximum, they could be maximized separately from the rest of the equation *if* they were independent of it. They are not independent, but the maximization may still be done separately using the Viterbi algorithm (Rabiner and Juang, 1993), a form of dynamic programming. Define

$$f_2(t_2) = \max_{t_1} (\alpha_1(t_1) + \lambda_1(t_2 - t_1) \alpha_2(t_2)). \quad (5)$$

Then, from Eq. (4)

$$\alpha = \max_{t_2, t_3} (f_2(t_2) + \lambda_2(t_3 - t_2) \alpha_3(t_3)). \quad (6)$$

Similarly,

$$f_3(t_3) = \max_{t_2} (f_2(t_2) + \lambda_2(t_3 - t_2) \alpha_3(t_3)), \quad (7)$$

and

$$\alpha = \max_{t_3} (f_3(t_3)). \quad (8)$$

Computationally, the function  $f_2(t_2)$  is simple to calculate: for each  $t_2$ , the value of  $t_1$  that gives the maximum  $f_2$  is calculated. Similarly, once  $f_2(t_2)$  is known,  $f_3(t_3)$  may be calculated separately for each value of  $t_3$ . Last,  $\alpha$  can be calculated as a one-dimensional optimization problem.

Two steps of this procedure are two-dimensional computations: calculating  $f_2(t_2)$  requires a one-dimensional maximization (over  $t_1$ ) for each value of  $t_2$ , so for all values of  $t_2$  this is a two-dimensional optimization. Calculating  $f_3(t_3)$  is similarly two-dimensional, and the final maximization of  $\alpha$  is only a one-dimensional calculation. Thus, this optimization method reduces the problem in complexity by a factor of  $n/2$ , where  $n$  is the number of time steps (frames) in the spectrogram. More generally, the computational complexity of a sound type with  $k$  FM sections is reduced by  $n^{k-2}/(k-1)$ .

With the above optimization, this method for handling time variation takes only a fraction of the time required for calculating a spectrogram. Fundamentally, the method oper-

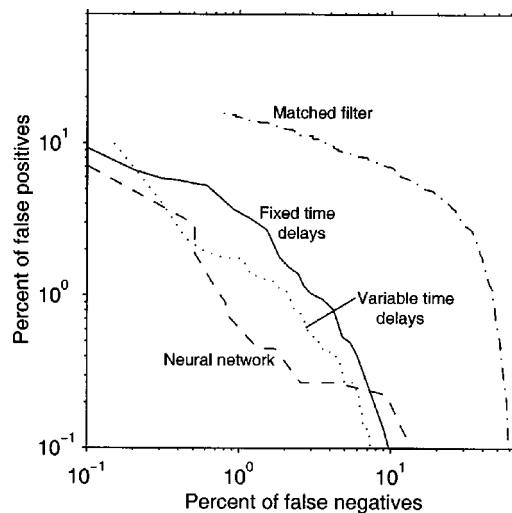


FIG. 8. Comparison of fixed-time method [Eq. (3), the same curve as the “spectrogram correlation” curve in Fig. 5] and variable time-delay method [Eq. (4)]. Curves lower and farther to the left represent lower error rates. Curves for the neural network and matched filter are plotted for reference.

ates on far fewer data points than are used in a spectrogram: it operates on only one data point per spectrogram frame, while the spectrogram calculation computes a many-point FFT for each spectrogram frame.

## C. Comparison

Does this new method, allowing timing variation among sections of a call, result in a recognizer that is any better at recognizing bowhead song end notes than the fixed-timing version of Eq. (3)?

Equations (5)–(8) were used on the same data set of end notes described in Sec. III. The false-positive versus false-negative curves for the time-variation method are shown in Fig. 8. Again, raising the threshold reduces the false-positive error rate, and lowering the threshold reduces the false-negative rate.

As the figure shows, the new method reduces the error rate moderately over most of the range of measurement. The combined error rate was 2.5% (1.9% false positive, 0.6% false negative). This represents an improvement of about one-third from the spectrogram-correlation error rate of 3.6%. The amount of improvement expected for other sound types depends on the amount of sound-to-sound variation, with sounds that vary more in timing between their constituent frequency sweeps likely to improve the most with this time-varying method.

## D. Other combination functions

Equation (4) combines the terms  $\lambda_i(t) \alpha_i(t)$  by adding them, but addition is not the only combination function possible. Any function that is monotonic and nondecreasing on positive numbers will work. Monotonicity is necessary for the optimization of Eqs. (5)–(8).

Two other combination functions were tried: multiplication and weighted addition, represented, respectively, by

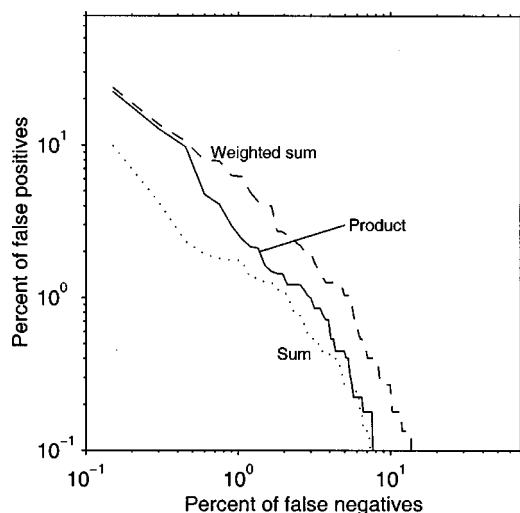


FIG. 9. Comparison of several combination rules for the time-varying recognition method. Results are shown for the sum method [Eq. (4)], the same curve as the “variable time-delay” curve of Fig. 8], the product method [Eq. (9)], and the weighted-sum method [Eq. (10)].

$$\alpha = \max_{t_1, t_2, t_3} (\alpha_1(t_1) \cdot \lambda_1(t_2 - t_1) \alpha_2(t_2) \cdot \lambda_2(t_3 - t_2) \alpha_3(t_3)), \quad (9)$$

and

$$\alpha = \max_{t_1, t_2, t_3} (\alpha_1(t_1)w_1 + \lambda_1(t_2 - t_1)\alpha_2(t_2)w_2 + \lambda_2(t_3 - t_2)\alpha_3(t_3)w_3), \quad (10)$$

where each  $w_i$  is a weighting value chosen for each time delay.

The weighted-sum method was devised in response to the perceived importance of different sections of bowhead song end notes, which varied from section to section. The initial FM downsweep and the FM upsweep were noticed to be nearly always present and relatively loud. The third final FM downsweep was sometimes weak or absent. Accordingly, weights of  $w_1 = 1.0$ ,  $w_2 = 1.0$ , and  $w_3 = 0.5$  were somewhat arbitrarily assigned to the three sections, respectively. These equations were then applied to the data set of bowhead song end notes.

Results are shown in Fig. 9, where the sum [Eq. (4)], product [Eq. (9)], and weighted-sum [Eq. (10)] combination methods are compared. The product and weighted-sum methods were worse than the sum method over almost all of their range. The best combined error rates for the product method was 3.1% (1.6% false positive+1.5% false negative). The weighted-sum method was worse than the fixed-offset method of Sec. II, with a combined error rate of 4.5% (2.7% false positive, 1.8% false negative).

## V. DISCUSSION

The spectrogram-correlation method presented here is effective at finding a large percentage of the bowhead sounds in a test database, much better than a simple matched filter. A neural-network method performs still better; however, there are several factors to consider in using one of these

methods. Here, we outline the advantages and disadvantages of the recognition methods presented [using the sum version of Eq. (4)] and the other methods used for the comparison in Sec. III.

**Matched filtering.** This method is known to be optimal in the case of white Gaussian noise and a known signal. If the sounds of a given species are stereotypic or nearly so, as in the case of blue whale calls (Stafford *et al.*, 1998), and if the noise environment approximates white Gaussian noise, matched filtering can be a quite effective method. In the presence of even a small amount of variation from sound to sound, or of harmonic interference such as from ship noise, matched filtering performs less well than the other methods.

**Neural network.** The advantage of this method is that it worked the best of the methods tried, with a best combined error rate of 1.6%. Its drawbacks are that (1) it requires a large training data set and associated operator time for data preparation, and (2) the functioning of a trained network is difficult to understand, making it uncertain under what conditions—e.g., what changes in the sound structures or noise—the network might fail. The network handled time variation in bowhead vocalizations well, probably due to its large training set.

**Spectrogram correlation.** This worked fairly well, with a best combined error rate of 2.5%, and should be suitable for many applications. It requires a relatively small set of training sounds—48 samples were used here, and as few as 5 samples have been used successfully in unpublished work. The shape of the kernel (center of Fig. 2) makes it easy to understand which characteristics of a sound the recognizer will and will not respond to.

## ACKNOWLEDGMENTS

Thanks to the field crew in Barrow for collection of sounds, and to Steve Mitchell and Beth Weisburn for insights into the analysis. This work was supported by Grant Number N00014-93-1-0431 from the U.S. Office of Naval Research, by Contract Number N00015-94-C-6016 from the Naval Research Laboratory, by the North Slope Borough (Box 69, Barrow, Alaska), and by the National Marine Fisheries Service (NMFS) through the Alaska Eskimo Whaling Commission (AEWC) to the North Slope Borough.

- Altes, R. A. (1980). “Detection, estimation, and classification with spectrograms,” *J. Acoust. Soc. Am.* **67**, 1232–1246.
- Bradbury, J. W., and Vehrencamp, S. L. (1998). *Principles of Animal Communication* (Sinauer, Sunderland, MA).
- Buck, J. R., and Tyack, P. L. (1993). “A quantitative measure of similarity for *Tursiops truncatus* signature whistles,” *J. Acoust. Soc. Am.* **94**, 2497–2506.
- Chabot, D. (1984). “Sound production of the humpback whale (*Megaptera novaeangliae*, Borowski) in Newfoundland waters,” Masters thesis, Memorial University of Newfoundland, St. John’s.
- Chabot, D. (1988). “A quantitative technique to compare and classify humpback whale (*Megaptera novaeangliae*) sounds,” *Ethology* **77**, 89–102.
- Charif, R. A., Mitchell, S. G., and Clark, C. W. (1995). *CANARY 1.2 User’s Manual* (Cornell Laboratory of Ornithology, Ithaca, NY).
- Clark, C. W. (1981). “Acoustic Communication and Behavior of the Southern Right Whale,” Ph.D. thesis, State University of New York at Stony Brook.

- Clark, C. W. (1982). "The acoustic repertoire of the southern right whale, a quantitative analysis," *Anim. Behav.* **30**, 1060–1071.
- Clark, C. W. (1983). "Acoustic communication and behavior of the Southern Right Whale (*Eubalaena australis*)," in *Communication and Behavior of Whales*, edited by R. Payne (Westview, Boulder), pp. 163–198.
- Clark, C. W. (1989). "Call tracks of bowhead whales based on call characteristics as an independent means of determining tracking parameters," *Rep. Intl. Whal. Commn.* **39**, 111–112.
- Clark, C. W. (1990). "Acoustic behavior of mysticete whales," in *Sensory Abilities of Cetaceans*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 571–583.
- Clark, C. W. (1991). *Ocean Voices of the Alaskan Arctic* [audio cassette] (Cornell Laboratory of Ornithology, Ithaca, NY).
- Clark, C. W., and Ellison, W. T. (1989). "Numbers and distributions of bowhead whales, *Balaena mysticetus*, based on the 1986 acoustic study off Pt. Barrow, Alaska," *Rep. Intl. Whal. Commn.* **39**, 297–303.
- Clark, C. W., and Ellison, W. T. (2000). "Calibration and comparison of the acoustic location methods used during the spring migration of the bowhead whale, *Balaena mysticetus*, off Pt. Barrow, Alaska, 1984–1993," *J. Acoust. Soc. Am.* **107**, 3509–3517.
- Clark, C. W., and Fristrup, K. M. (1996). "Whales '95: A combined visual and acoustic survey of blue and fin whales off Southern California," *Rep. Intl. Whal. Commn.* **47**, 583–600.
- Clark, C. W., and Johnson, J. H. (1984). "The sounds of the bowhead whale, *Balaena mysticetus*, during the spring migrations of 1979 and 1980," *Can. J. Zool.* **62**, 1436–1441.
- Clark, C. W., and Mellinger, D. K. (1994). "Application of Navy IUSS for whale research," *J. Acoust. Soc. Am.* **96**, 3315(A).
- Clark, C. W., Ellison, W. T., and Beeman, K. (1986). "Acoustic tracking of migrating bowhead whales," *Proc. IEEE Oceans '86*, pp. 341–346 (IEEE, New York).
- Clark, C. W., Marler, P., and Beeman, K. (1987). "Quantitative analysis of animal vocal phonology: An application to swamp sparrow song," *Ethology* **76**, 101–115.
- Clark, C. W., Bower, J. B., and Ellison, W. T. (1991). "Acoustic tracks of migrating bowhead whales, *Balaena mysticetus*, off Point Barrow, Alaska based on vocal characteristics," *Rep. Intl. Whal. Commn.* **40**, 596–597.
- Clark, C. W., Mitchell, S. G., and Charif, R. A. (1996). "Distribution and behavior of the bowhead whale, *Balaena mysticetus*, based on preliminary analysis of acoustic data collected during the 1993 spring migration off Point Barrow, Alaska," *Rep. Intl. Whal. Commn.* **46**, 541–554.
- Clark, C. W., Tyack, P. L., and Ellison, W. T. (1998). Quicklook, Low-Frequency Sound Scientific Research Program. Phase I: Responses of blue and fin whales to SURTASS LFA, Southern California Bight, 5 September–21 October 1997.
- Cummings, W. C., and Thompson, P. O. (1971). "Underwater sounds from the blue whale, *Balaenoptera musculus*," *J. Acoust. Soc. Am.* **50**, 1193–1198.
- deCharms, R. C., Blake, D. T., and Merzenich, M. M. (1998). "Optimizing sound features for cortical neurons," *Science* **280**, 1439–1443.
- Edds, P. L. (1982). "Vocalizations of the blue whale, *Balaenoptera physalus*, in the St. Lawrence River," *J. Mammal.* **63**, 345–347.
- Engineering Design. (1997). SIGNALRTS (Engineering Design, Belmont, MA).
- Frankel, A. S., Clark, C. W., Herman, L. M., and Gabriele, C. M. (1995). "Spatial distribution, habitat utilization, and social interactions of humpback whales, *Magaptera novaeangliae*, off Hawai'i, determined using acoustic and visual techniques," *Can. J. Zool.* **73**, 1134–1146.
- Fristrup, K. M. (1992). "Characterizing acoustic features of marine animal sounds," Woods Hole Oceanographic Institution, WHOI-92-04, Woods Hole, MA.
- Fristrup, K. M., and Watkins, W. A. (1994). "Marine animal sound classification," Woods Hole Oceanographic Institution, WHOI-94-13, Woods Hole, MA.
- Gaetz, W., Jantzen, K., Weinberg, H., Spong, P., and Symonds, H. (1993). "A neural network mechanism for recognition of individual *Orcinus orca* based on their acoustic behavior: Phase 1," *Proc. IEEE Oceans '93*, Vol. 1, pp. 455–457 (IEEE, New York).
- Ghosh, J., Deuser, L. M., and Beck, S. D. (1992). "A neural network-based hybrid system for detection, characterization, and classification of short-duration oceanic signals," *IEEE J. Ocean Eng.* **17**, 351–363.
- Goedeking, P. (1983). "A minicomputer-aided method for the detection of features from vocalizations of the cotton-top tamarin (*Saguinus oedipus*)," *Z. Tierpsychol.* **62**, 321–328.
- Graber, R. R. (1968). "Nocturnal migration in Illinois: Different points of view," *Wilson Bull.* **80**, 36–71.
- Griffin, D. R. (1964). *Bird Migration* (Natural History, Garden City, NY), pp. 11–13.
- Hebb, D. O. (1949). *The Organization of Behavior* (Wiley, New York).
- Herman, L. M., and Tavolga, W. N. (1980). "The communication systems of cetaceans," in *Cetacean Behavior: Mechanism and Function*, edited by L. M. Herman (Wiley, New York), pp. 149–209.
- Hubel, D. H., and Wiesel, T. N. (1962). "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *J. Physiol. (London)* **160**, 106–154.
- Hubel, D. H., and Wiesel, T. N. (1977). "Functional architecture of macaque monkey visual cortex," *Proc. R. Soc. London, Ser. B* **198**, 1–59.
- Jain, A. K., and Dubes, R. C. (1988). *Algorithms for Clustering Data* (Prentice-Hall, Englewood Cliffs, NJ).
- Kay, R. H., and Matthews, D. R. (1972). "On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency-modulated tones," *J. Physiol. (London)* **225**, 657–677.
- Larkin, R. P. (1978). "Radar observations of behavior of migrating birds in response to sounds broadcast from the ground," in *Animal Migration, Navigation, and Homing*, edited by K. Schmidt-Koenig and W. T. Keaton (Springer, New York).
- Leaper, R., Chappell, O., and Gordon, J. (1992). "The development of practical techniques for surveying sperm whale populations acoustically," *Rep. Intl. Whal. Commn.* **42**, 549–560.
- Lee, K.-F., Hon, H.-W., and Reddy, R. (1990). "An overview of the Sphinx speech-recognition system," *IEEE Trans. Acoust., Speech, Signal Process.* **38**, 35–45.
- Lippman, R. C. (1989). "Review of neural networks for speech recognition," *Neural Comput.* **1**, 1–38.
- McDonald, M. A., Hildebrand, J. A., and Webb, S. C. (1995). "Blue and fin whales observed on a seafloor array in the Northeast Pacific," *J. Acoust. Soc. Am.* **98**, 712–721.
- Mellinger, D. K. (1993). "Handling time variability in bioacoustic transient detection," *Proc. IEEE Oceans '93*, pp. 116–121 (IEEE, New York).
- Mellinger, D. K. (1995). *OSPREE 1.2 Guide*, Technical report (Cornell Laboratory of Ornithology, Ithaca, NY).
- Mellinger, D. K., and Clark, C. W. (1993). "A method for filtering bioacoustic transients by spectrogram image convolution," *Proc. IEEE Oceans '93*, pp. 122–127 (IEEE, New York).
- Mendelson, J. R. and Cynader, M. S. (1985). "Sensitivity of cat auditory primary cortex (AI) neurons to the direction and rate of frequency modulation," *Brain Res.* **327**, 331–335.
- Møller, A. R. (1977). "Coding of time-varying sounds in the cochlear nucleus," *Audiology* **17**, 446–468.
- Moody, D. B., and Stebbins, W. C. (1989). "Salience of frequency modulation in primate communication," in *The Comparative Psychology of Audition: Perceiving Complex Sounds*, edited by R. J. Dooling and S. H. Hulse (Erlbaum, Hillsdale, NJ).
- Moore, P. W. B., Roitblat, H. L., Penner, R. H., and Nachtigall, P. E. (1991). "Recognizing successive dolphin echoes with an integrator gateway network," *Neural Networks* **4**, 701–709.
- Moore, S. E., Stafford, K. M., Dahlheim, M. E., Fox, C. G., Braham, H. W., and Polovina, J. J. (1998). "Seasonal variation in reception of fin whale calls at five geographic areas in the North Pacific," *Mar. Mammal Sci.* **14**, 617–627.
- Newman, J. D., Lieblich, A. K., Talmage-Riggs, G., and Symmes, D. (1978). "Syllable classification and sequencing in twitter calls of squirrel monkeys (*Saimiri sciureus*)," *Z. Tierpsychol.* **47**, 77–88.
- Nyamsi, R. G. M., Aubin, T., and Bremond, J. C. (1994). "On the extraction of some time-dependent parameters of an acoustic signal by means of the analytic signal concept: its application to animal sound study," *Bioacoustics* **5**, 187–203.
- Oppenheim, A. V., and Schaffer, R. W. (1975). *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ).
- Patrick, P. H., Ramani, N., Sheehan, R. W., and Hanson, W. (1994). "Listening to and identifying wildlife using computers," *Global Biodiversity* **3**(3), 12–16.
- Payne, R. S., and Payne, K. B. (1971). "Underwater sounds of southern right whales," *Zoologica* **58**, 159–165.
- Payne, K., Tyack, P., and Payne, R. (1983). "Progressive changes in the songs of humpback whales (*Megaptera novaeangliae*): A detailed analysis of two seasons in Hawaii," in *Communication and Behavior of Whales*, edited by R. S. Payne (Westview, Boulder), pp. 9–57.

- Payne, R. S., and McVay, S. (1971). "Songs of humpback whales," *Science* **173**, 587–597.
- Pinkowski, B. (1994). "Robust Fourier descriptors for characterizing amplitude-modulated waveform shapes," *J. Acoust. Soc. Am.* **95**, 3419–3423.
- Potter, J. R., Mellinger, D. K., and Clark, C. W. (1994). "Marine mammal call discrimination using artificial neural networks," *J. Acoust. Soc. Am.* **96**, 1255–1262.
- Rabiner, L. R., and Juang, B.-H. (1986). "An introduction to hidden Markov models," *IEEE ASSP Magazine*, January 1986, Vol. 3, pp. 4–15.
- Rabiner, L. R., and Juang, B.-H. (1993). *Fundamentals of Speech Recognition* (Prentice-Hall, Englewood Cliffs, NJ).
- Raftery, A. E., and Zeh, J. E. (1998). "Estimating bowhead whale population size and rate of increase from the 1993 census," *J. Am. Stat. Assoc.* **93**, 451–463.
- Ramani, N., and Patrick, P. H. (1992). "Fish detection and identification using neural networks," *IEEE J. Ocean Eng.* **17**, 364–368.
- Rivers, J. A. (1997). "Blue whale, *Balaenoptera musculus*, vocalizations from the waters off central California," *Mar. Mammal Sci.* **13**, 186–195.
- Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1987). *Parallel Distributed Processing* (MIT, Cambridge, MA).
- Schevill, W. E. (1964). "Underwater sounds of cetaceans," in *Marine Bio-acoustics*, edited by W. N. Tavolga (Pergamon, New York), pp. 307–316.
- Schevill, W. E., and Watkins, W. A. (1962). "Whale and Porpoise Voices: A Phonograph Record," Contribution #1320 from Woods Hole Oceanogr. Inst., Woods Hole.
- Schevill, W. E., Watkins, W. A., and Backus, R. H. (1964). "The 20-cycle signals and *Balaenoptera* (fin whales)," in *Marine Bio-acoustics*, edited by W. N. Tavolga (Pergamon, New York), pp. 147–152.
- Silber, G. K. (1986). "The relationship of social vocalizations to surface behavior and aggression in the Hawaiian humpback whale *Megaptera novaeangliae*," *Can. J. Zool.* **64**, 2075–2080.
- Stafford, K. M., Fox, C. G., and Clark, D. S. (1998). "Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean," *J. Acoust. Soc. Am.* **104**, 3616–3625.
- Stafford, K. M., Nieuwkerk, S. L., and Fox, C. G. (1999). "An acoustic link between blue whales in the eastern tropical Pacific and the northeast Pacific," *J. Acoust. Soc. Am.* **15**, 1258–1268.
- Sturtivant, C., and Datta, S. (1997). "Automatic dolphin whistle detection, extraction, encoding, and classification," *Proc. Inst. Acoust.* **19(9)**, 259–266.
- Taylor, A. (1995). "Bird flight call discrimination using machine learning," *J. Acoust. Soc. Am.* **97**, 3370(A).
- Taylor, A., Watson, G., Grigg, G., and McCallum, H. (1996). "Monitoring frog communities: an application of machine learning," in *Innovative Applications Artificial Intelligence Conference* (AAAI Press, Menlo Park, CA), pp. 1564–1569.
- Terres, J. K. (1980). *The Audubon Society Encyclopedia of North American Birds* (Knopf, New York), pp. 604–605.
- Thompson, P. O., Findley, L. T., Vidal, O., and Cummings, W. C. (1996). "Underwater sounds of blue whales, *Balaenoptera musculus*, in the Gulf of California, Mexico," *Mar. Mammal Sci.* **12**, 288–293.
- Thompson, T. J., Winn, H. E., and Perkins, P. J. (1979). "Mysticete sounds," in *Behavior of Marine Mammals: Current Perspectives in Research*, edited by H. E. Winn and B. L. Olla (Plenum, New York), pp. 403–431.
- van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory* (Wiley, New York), Vol. I.
- Watkins, W. A., and Schevill, W. E. (1972). "Sound source location by arrival-times on a non-rigid three-dimensional hydrophone array," *Deep-Sea Res.* **19**, 691–706.
- Watkins, W. A., Tyack, P., Moore, K. E., and Bird, J. E. (1987). "The 20-Hz signals of finback whales (*Balaenoptera physalus*)," *J. Acoust. Soc. Am.* **82**, 1901–1912.
- Weisburn, B. A., Mitchell, S. G., Clark, C. W., and Parks, T. W. (1993). "Isolating biological acoustic transient signals," *Proc. IEEE Intl. Conf. Acoust., Speech, Sig. Process.*, Vol. 1, pp. 269–272 (IEEE, New York).
- Whitehead, H., and Weilgart, L. (1990). "Click rates from sperm whales," *J. Acoust. Soc. Am.* **87**, 1798–1806.
- Whitfield, I. C., and Evans, E. F. (1965). "Responses of auditory cortical neurons to stimuli of changing frequency," *J. Neurophysiol.* **28**, 655–672.
- Wiley, R. H., and Richards, D. C. (1982). "Adaptations for acoustic communication in birds: Sound transmission and signal detection," in *Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Academic, London), Vol. I, pp. 131–181.
- Winn, H. E., and Perkins, P. J. (1976). "Distribution and sounds of the minke whale, with a review of mysticete sounds," *Cetology* **19**, 1–12.
- Würsig, B., and Clark, C. W. (1993). "Behavior," in *The Bowhead Whale*, edited by J. J. Burns, J. J. Montague, and C. J. Cowles (Allen, Lawrence, KS), pp. 157–199.
- Zeh, J. E., Clark, C. W., George, J. C., Withrow, D., Carroll, G. M., and Koski, W. R. (1993). "Current population size and dynamics," in *The Bowhead Whale*, edited by J. J. Burns, J. J. Montague, and C. J. Cowles (Allen, Lawrence, KS), pp. 409–489.

# Sound scattering and localized heat deposition of pulse-driven microbubbles

Sascha Hilgenfeldt

*Division of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, Massachusetts 02138*

Detlef Lohse

*Department of Applied Physics and J. M. Burgers Centre for Fluid Dynamics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*

Michael Zomack

*Schering AG, Clinical Development, Müllerstrasse 178, D-13342 Berlin, Germany*

(Received 13 July 1999; accepted for publication 10 February 2000)

The sound scattering of free microbubbles released from strongly driven ultrasound contrast agents with brittle shell (e.g., Sonovist<sup>®</sup>) is studied numerically. At high peak pressure of the driving pulses, the bubbles respond nonlinearly with cross sections pronouncedly larger than in the linear case; a large portion of the energy is radiated into high frequency ultrasound. Subsequent absorption of these high frequencies in the surrounding liquid (blood) diminishes the effective scattering cross section drastically. The absorption results in highly localized heating, with a substantial temperature rise within the first few  $\mu\text{m}$  from the bubble surface. The maximum heating in 1  $\mu\text{m}$  distance is strongly dependent on driving pressure. Temperature elevations of more than 100 K can be achieved for amplitudes of  $P_a \approx 30$  atm, which coincides with the highest pressures used in ultrasound diagnostics. The perfectly spherical collapses assumed here occur rarely, and the heating is highly localized and transient ( $\sim 10$   $\mu\text{s}$ ). Therefore, a thermal hazard would only be expected at driving pressures beyond the diagnostic range. © 2000 Acoustical Society of America.

[S0001-4966(00)02905-2]

PACS numbers: 43.80.Qf, 43.25.Yw [FD]

## INTRODUCTION

Ultrasound as a diagnostic tool in medicine has many advantages: it is an easy-to-use, real-time, noninvasive, relatively inexpensive means of imaging inner organs or the bloodstream. Furthermore, it does not expose the patient to potentially harmful high-energy radiation. The quality of the images, however, cannot (yet) compete with the results of tomographic devices (X-ray CT, NMR), where contrast agents are in use. Improving the brightness and contrast of ultrasonograms has therefore been a vital field of research in the past decades. To this end, a special kind of contrast agent has been proposed:<sup>1</sup> a suspension of microbubbles ( $\mu\text{m}$  sized gas bubbles) with tremendous scattering efficiency for incident ultrasound.<sup>2</sup> These agents have been successful in yielding brighter images, in particular when the bubbles are driven strongly and excited to nonlinear oscillations, leading to active sound emission both at the center driving frequency (fundamental) and its higher harmonics.<sup>3-6</sup>

Higher harmonics with frequencies beyond 10 MHz are strongly damped while passing through liquids. This energy deposition leads to a temperature increase. The aim of this paper is to calculate the absorbed energy and the degree of heating. The main result is that temperature rises stay moderate ( $\sim 10$  K or less) even at a distance of 1  $\mu\text{m}$  from the bubble surface, provided that pressure amplitudes of  $\approx 15$  atm are not exceeded. If the driving amplitude is increased to the upper limit of the diagnostic range at  $\approx 30$  atm, local temperatures can rise by more than 100 K for a brief time.

The paper is organized as follows: In Secs. I and II we introduce the computational scheme and calculate the bubble's response to acoustical driving. In Sec. III we investigate the energy balance of the driven bubble system and present results for both the scattering and the absorption of the emitted sound. Based on these findings we compare possible sources of heating; the heating through sound absorption is found to be dominant, leading to the abovementioned temperature rises (Sec. IV). Section V contains further discussions and presents conclusions.

## I. COMPUTATIONAL SCHEME

In order to calculate the intensity and spectrum of the emitted sound, we first model the dynamics of the bubble volume oscillations. The resulting time series for the bubble radius is then used to evaluate the emitted sound pressure and intensity, and subsequently the other quantities of our interest, namely, cross sections and dissipation rates of sound energy.

### A. Radius dynamics

All results presented here are obtained for the case of a single spherically oscillating bubble. The dynamics of the bubble radius  $R(t)$  is well described by the (modified) Keller equation<sup>7-10</sup>

$$\begin{aligned}
& \left(1 - \frac{\dot{R}}{c_l}\right) R \ddot{R} + \left(1 - \frac{\dot{R}}{3c_l}\right) \frac{3}{2} \dot{R}^2 \\
& = \left(1 + \frac{\dot{R}}{c_l}\right) \frac{1}{\rho_l} (p(R,t) - P(t) - P_0) \\
& \quad + \frac{R}{\rho_l c_l} \frac{d}{dt} p(R,t) - 4\nu_l \frac{\dot{R}}{R} - \frac{2\zeta}{\rho_l R}, \quad (1)
\end{aligned}$$

where  $\nu_l$ ,  $\rho_l$  and  $c_l$  are the kinematic viscosity, density, and speed of sound of the liquid, respectively, and  $\zeta$  is the surface tension of the gas/liquid system. We set all these parameters except the viscosity to their values for the water/air system, i.e.,  $\rho_l = 1000 \text{ kg m}^{-3}$ ,  $c_l = 1481 \text{ m s}^{-1}$  and  $\zeta = 0.072 \text{ kg s}^{-2}$ . As we want to model the oscillation of air bubbles in blood, we have to take into account the higher viscosity of blood and choose  $\nu_l = 3 \cdot 10^{-6} \text{ m}^2 \text{ s}^{-1}$  as a convenient mean value (the actual viscosity being dependent on shear rate).  $P_0$  is the static ambient pressure ( $P_0 = 1 \text{ atm}$  throughout this study),  $P(t)$  represents the pressure of the driving (incident) sound wave. Terms  $\propto 1/c_l$  describe effects of liquid compressibility on the bubble oscillations, in particular the influence of sound emission (radiation damping). The terms in parentheses containing  $\dot{R}/c_l$ , while not important for calculating the bubble dynamics itself, can be significant for determining the value of  $\ddot{R}$  and the sound emission from the bubble, if driving pressures are very high.

Finally,  $p(R,t)$  gives the interior pressure of the gas-filled bubble, modeled by a polytropic process equation with a van der Waals hard core radius  $h$ :

$$p(R,t) = p(R(t)) = \left( P_0 + \frac{2\zeta}{R_0} \right) \left( \frac{R_0^3 - h^3}{R^3(t) - h^3} \right)^\kappa. \quad (2)$$

The parameter  $R_0$  is the ambient radius, i.e., the radius of the undriven bubble at the static pressure  $P_0$ . The hard core radius is tabulated; for air we have  $h \approx R_0/8.54$ .<sup>11</sup> As a good approximation, the polytropic index  $\kappa$  can be set equal to one (isothermal behavior), because the bubbles treated in this work are much smaller than the thermal diffusion length on typical time scales of the oscillation.<sup>12</sup> An exception to this occurs if the velocity  $|\dot{R}|$  is too large (very small time scales), as is the case when the bubble undergoes a violent collapse (see Sec. II). Because the duration of the collapse is much smaller than the time scales of the oscillation, an isothermal description throughout will not lead to large errors (cf. Ref. 13).

The Keller equation (1) has proved to be an accurate description of bubble wall dynamics even in the case of strongly nonlinear oscillations.<sup>9,14</sup> Some terms traditionally included in the Keller equation, e.g., time derivative terms of surface tension and viscous contribution, have been omitted from Eq. (1) for computational simplicity, as they are negligible throughout the whole oscillation in the present case. A necessary (though not sufficient) condition for strong collapse behavior is that  $P_a > P_0$ , where  $P_a$  is the amplitude of the driving  $P(t)$ . This allows for the bubble to experience a phase of negative effective pressure during its oscillation (cf. Ref. 15). In ultrasound diagnostics, peak driving pressures

can be as high as 30 atm, i.e., strong collapses, accompanied by intense sound emission, are to be expected.

The microbubbles used in ultrasonic contrast agents are often coated to make them more stable against diffusive dissolution as well as against agglomeration. If the coating is an elastic material, modifications to the Keller equation can be derived, cf. Ref. 16. In this study, we focus on modeling the properties of Sonovist®, a contrast agent with a polycyanacrylate coating that forms a rather brittle shell around the bubble. Experimentally, this kind of contrast agent shows an abrupt onset of strong signals as the driving amplitude is increased. This is interpreted as the sound emission from free gas bubbles (which is much stronger than that of coated bubbles), released as the shells break under the high driving force. Therefore, we use the Keller equation for uncoated gas bubbles in the present work.

## B. The incident pulse

Next we specify the pressure of the driving sound wave. In an earlier study,<sup>5</sup> we assumed monochromatic driving with a standing wave to simplify our analysis. In reality, bubbles in an ultrasound diagnostics application are driven by the signal of a diagnostic transducer, which is almost always *pulsed*, and quite often the pulses are only a few wavelengths long. Note that the wavelength for diagnostic ultrasound of 3 MHz frequency is  $\lambda \approx 500 \text{ } \mu\text{m}$ , which is much larger than the bubble ambient radii of  $R_0 \approx 0.5\text{--}5 \text{ } \mu\text{m}$  we are interested in here. Therefore, we can assume that an individual bubble experiences a spatially uniform driving pressure.

We model the pressure pulse  $P(t)$  of a diagnostic transducer as

$$P(t) = P_a \cos(2\pi f_d(t - t_c)) \exp\left(-\frac{h_d^2(2\pi f_d)^2}{4}(t - t_c)^2\right), \quad (3)$$

centered around  $t_c$  with relative width  $h_d$ . We choose  $h_d = 1/3$  here. The center frequency is  $f_d = 3 \text{ MHz}$  throughout this study; after Fourier transform, the frequency space representation is

$$\begin{aligned}
P(f) = \frac{P_a}{2\pi f_d h_d} \left[ \exp\left(-\frac{(f - f_d)^2}{h_d^2 f_d^2}\right) \right. \\
\left. + \exp\left(-\frac{(f + f_d)^2}{h_d^2 f_d^2}\right) \right], \quad (4)
\end{aligned}$$

which is (almost) a Gaussian spectrum; the corresponding power  $\propto |P(f)|^2$  decays to  $1/e^2$  of its maximum value within a distance of  $\pm h_d f_d$  around  $f_d$  (cf. Fig. 1). Thus,  $1/h_d$  is a measure for the spatial extension of the pulse (in units of wavelengths), i.e., our pulse is about three wavelengths long, corresponding to a short pulse routinely available in medical applications of diagnostic ultrasound.

As the peak pressures of the driving pulse are quite high, the nonlinearity in the Navier–Stokes equations will lead to a distortion of the wave profile. For a plane harmonic wave, the initial steepening of the profile can be described analytically by Fubini's method.<sup>17</sup> After traveling a certain distance,

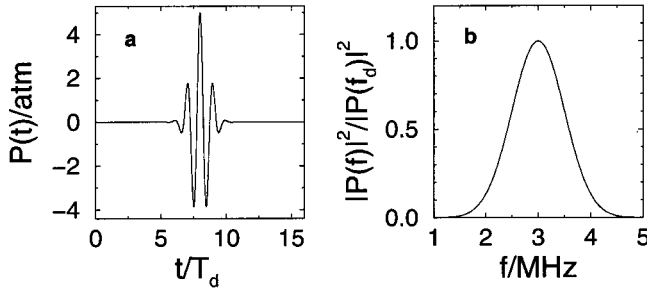


FIG. 1. Modeled ultrasound diagnostics driving pulse. Left: time series  $P(t)$  ( $t$  normalized to  $T_d \equiv 1/f_d$ ); right: normalized spectral acoustic energy density  $\propto |P(f)|^2$ . It is centered around the mean frequency  $f_d = 3$  MHz and decays to  $1/e^2$  of its maximum value within  $h_{df_d} = f_d/3 = 1$  MHz above and below  $f_d$ . This corresponds to the effective spatial extent of the pulse of  $\approx 3$  wavelengths.

a shock front develops and the wave profile becomes more and more similar to a sawtooth, with an amplitude decaying roughly as  $1/(\text{distance})$ . As mentioned above, the wavelengths in our example are large enough to regard the pulse as a plane wave. In addition, we assume that the Gaussian envelope of the pulse Eq. (3) is largely unaffected by the distortion of the sinusoidal part of the wave. Thus we obtain a nonlinearly distorted driving pulse (cf. Ref. 18) whose effects we can compare with the “linear” formula Eq. (3). This comparison and the transition from Fubini to sawtooth profiles is elaborated further in Ref. 19. We only mention here that the results shown in the present paper for cross sections and liquid heating were practically unaltered by the changes in the pulse profile, and we use the undistorted driving Eq. (3) everywhere. The selective damping of high frequencies, discussed below, also helps to limit nonlinear distortion.

The ultrasound pulses in diagnostic applications are commonly characterized by their peak rarefaction pressure  $P_-$ , whose magnitude is usually a factor of 2–3 lower than the peak compression pressure. It is used to define the *mechanical index*<sup>20</sup>  $MI = (P_-/\text{MPa})/\sqrt{f_d/\text{MHz}}$ . We observe that in our case  $P_- \approx P_a$  after distortion, so that  $P_a$  can be used to evaluate the  $MI$ . For  $P_a \approx 33$  atm, we achieve  $MI \approx 1.9$ , which is the recommended upper limit for diagnostic applications.

### C. Emitted sound pressure

In general, the pressure of sound emitted from a volume oscillator (i.e., a bubble) consists of two contributions: (i) the *active emission* due to the volume oscillation, and (ii) the *passive* disturbance of the incident sound field caused by the mere presence of the (not necessarily oscillating) body in the liquid (cf. Ref. 21). We have shown in Ref. 6 that for the frequencies, driving pressures, and bubble radii of our interest the passive contribution is always very small; we therefore restrict ourselves to the active contribution. Passive effects may become important in the presence of very large scatterers, but we will not consider this case here.

From the general formulas for far-field ( $r \gg \lambda$ ) sound emission from an oscillating body of given volume it is easy to find, for the special case of a spherical bubble,

$$P_s(r, t) = \rho_l \frac{R(t')}{r} (2\dot{R}(t')^2 + R(t')\ddot{R}(t')) = \frac{1}{r} q_s(t), \quad (5)$$

where we have introduced  $q_s(t) \equiv rP_s(r, t)$ , thus separating the trivial (geometrical)  $1/r$  spatial dependence of the radiation from the time-dependent part;  $t' = t - r/c_l$  is the retarded time. There is no angular dependence because we have assumed the long wavelength limit  $\lambda \gg R$  (S wave scattering, cf. Sec. I B).

The far-field limit certainly holds in all situations studied here where the detected sound energy in a few cm distance from the bubble is concerned. However, the condition  $r \gg \lambda$  will be violated when we need to determine the energy density of emitted sound very near the bubble surface, as will be the case in Sec. IV C, where we discuss the temperatures induced by the absorption of the sound wave at distances of only a few  $\mu\text{m}$  from the bubble surface. Then, we will take into account an additional term for sound emission in the *intermediate range*  $R \ll r \leq \lambda$  (Ref. 21) and define the effective pressure

$$P_s^{\text{eff}}(r, t) = \rho_l \frac{R(t')}{r} \left( 2\dot{R}(t')^2 + R(t')\ddot{R}(t') + \frac{c_l}{r} R(t')\dot{R}(t') \right), \quad (6)$$

whose square is a measure of the energy contained in the displacement field close to the bubble. The localized correction term decays more rapidly with  $r$  than the propagating wave term, and becomes negligible for  $r \gg \lambda$ . It will turn out that, paradoxically, even at distances as small as  $1 \mu\text{m}$  from the bubble the intermediate-field correction is small, so that using Eq. (5) is sufficient for all of our computations (cf. Sec. IV C and Ref. 19). We will drop the prime of  $t'$  for convenience in subsequent formulas and figures.

### D. Energy spectrum and damping of emitted sound

From Eq. (5), we obtain the total emitted sound energy as

$$E_s = \frac{4\pi}{\rho_l c_l} \int_{-\infty}^{+\infty} q_s(t)^2 dt = \frac{4\pi}{\rho_l c_l} \int_{-\infty}^{+\infty} |q_s(f)|^2 df \equiv \int_{-\infty}^{+\infty} \varepsilon_0(f) df, \quad (7)$$

where  $q_s(f)$  is the Fourier transform of  $q_s(t)$ , and the second equality is Parseval’s theorem. We have also defined the spectral energy density or *spectral power* of the (undamped) emitted sound here,  $\varepsilon_0(f) = 4\pi |q_s(f)|^2 / \rho_l c_l$ .

Every frequency component  $\varepsilon_0(f)$  of emitted sound energy is attenuated as the wave passes through a medium. Having traveled a distance  $r$ , the remaining spectral power is

$$\varepsilon(r, f) = \varepsilon_0(f) \exp(-\alpha(f)r). \quad (8)$$

Here,  $\alpha(f)$  is the characteristic attenuation coefficient for the *spectral power* of sound in blood (twice the attenuation coefficient for the pressure field amplitude). Its functional form can be approximated by<sup>22</sup>

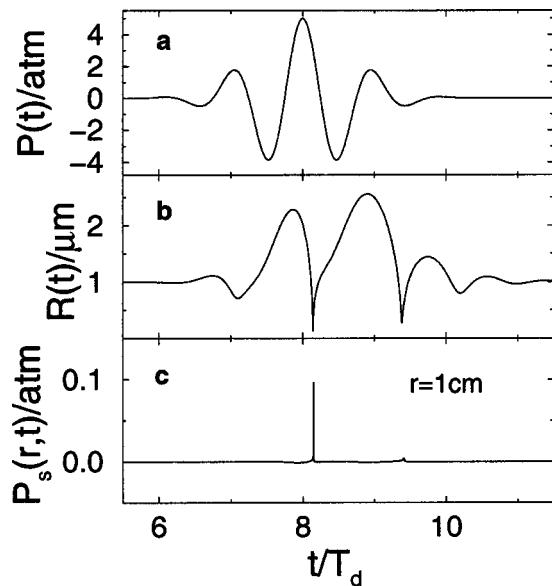


FIG. 2. Time series of (a) incident sound pressure  $P(t)$  ( $P_a=5$  atm), (b) bubble radius  $R(t)$  ( $R_0=1$   $\mu\text{m}$ ), and (c) emitted sound pressure  $P_s(r,t)$  at a fixed distance  $r=1$  cm from the center of the bubble (*without* taking into account sound damping in the medium).  $P_s$  responds to the main bubble collapse with a sharp pressure spike.

$$\alpha(f) \approx \alpha_0 \left( \frac{f}{\text{MHz}} \right)^\beta, \quad (9)$$

with  $\alpha_0 \approx 3.8 \text{ m}^{-1}$  and  $\beta \approx 1.5$ . This leads to noticeable damping in blood even at low multiples of  $f_d$ , and very strong attenuation for higher frequencies. Damping is even more pronounced in most tissues, but can be described by Eq. (9) as well, with similar values for  $\beta$  and generally larger  $\alpha_0$ .<sup>22</sup>

## II. BUBBLE RESPONSE

### A. Nonlinear oscillations and emitted sound

Let us first look at a typical response of the bubble radius  $R(t)$  to the incident pulse  $P(t)$  given by Eq. (3), computed via Eq. (1). Figure 2 shows time series of driving pressure and bubble oscillation near resonance ( $R_0=1.0$   $\mu\text{m}$ ) at a driving level of  $P_a=5$  atm, where the collapses are already quite violent. More specifically, there is one single collapse within the  $\sim 3$  completed oscillation cycles (the middle one) which is particularly strong; see Fig. 2(b). When we compute the resulting emitted sound pressure from Eq. (5), we observe that the radiated signal concentrates right at the instant of this collapse: a very short sound *spike* with high amplitude is emitted.

In Fig. 3, a close-up of the time interval around the violent collapse is shown (about 3 ns in this example). We see that the typical duration of the pressure spike [a few  $10^{-10}$  s, Fig. 3(c)] matches the turnaround time scale of  $R(t)$  at the minimum radius  $R_{\min} \approx h$ , cf. Fig. 3(a). This is not surprising, as this sets the time scale of high acceleration of the bubble wall; from Eq. (5) it is clear that a high  $|\ddot{R}|$  will cause a large pressure of the emitted wave. Indeed, bubbles driven at high pressure amplitudes reach maximum accelerations in excess of  $10^{12}$  g, see Fig. 3(b).

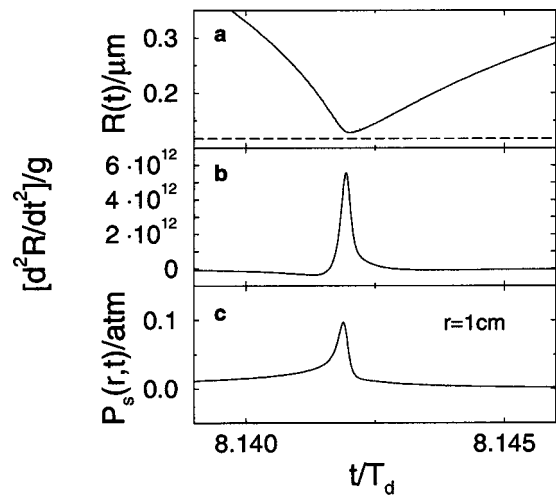


FIG. 3.  $R(t)$  (a),  $\ddot{R}(t)$  (b), and  $P_s(r,t)$  (c) around the main collapse dynamics of Fig. 2. Note the intimate connection of the time scales and the immense magnitude of the acceleration. The dashed line in (a) indicates the hard core radius  $h$ , which is almost reached at collapse.

These results are of course different for small  $P_a$  (about 1 atm or lower) or bubble radii far away from resonance. In these cases, the bubble cannot respond very strongly to the driving; its oscillation is more or less sinusoidal around  $R_0$ , and so is the shape of the corresponding emitted wave  $P_s(t)$ .

### B. Propagation of the radiated sound wave

The more violent the collapse gets, the smaller the time scales of the emitted sound pulse become. As the pulse gets closer and closer to a delta spike, more and more energy is emitted into a spectral region of high frequencies, where it is dissipated rapidly, because the damping of ultrasound becomes much more pronounced as the frequency rises, cf. Eq. (9). Thus, traveling through a layer of blood effectively fil-

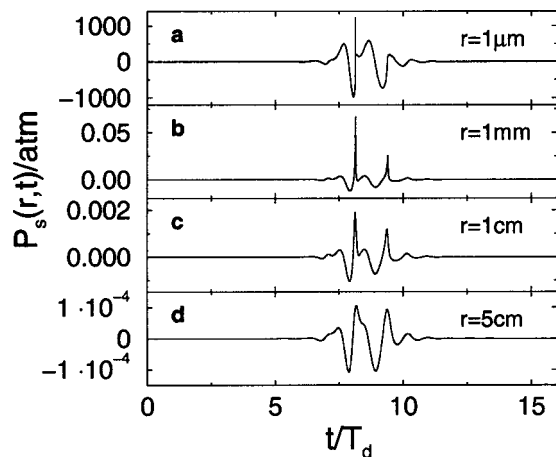


FIG. 4. Emitted sound pressure as a function of time for  $P_a=5$  atm,  $R_0=1$   $\mu\text{m}$  after traveling distances of (a) 1  $\mu\text{m}$ , (b) 1 mm, (c) 1 cm, and (d) 5 cm from the bubble's center. In (a), the pressure is the full effective pressure  $P_s^{\text{eff}}$  from Eq. (6). The nonpropagating part of  $P_s^{\text{eff}}$  decays rapidly with  $r$ , so that in (b)–(d) only the far field wave part  $P_s$  is present. The stronger damping of the higher frequency components reduces the “spikes” with traveled distance from (b) to (d) and establishes a smoother response. Also, the peak pressure drops considerably faster than the  $1/r$  behavior of an undamped spherical wave.



ters the high frequency components out of the signal  $P_s(t)$ . In Fig. 4, the emitted pulse is displayed after having traveled distances of 1  $\mu\text{m}$ , 1 mm, 1 cm, and 5 cm from the bubble. The result for 1  $\mu\text{m}$  distance necessitates the calculation of the full effective pressure  $P_s^{\text{eff}}(r)$  from Eq. (6), whereas for the other distances ( $>\lambda$ ) only the propagating far-field wave term  $P_s$  remains. Notice that the nonpropagating intermediate-field correction of the pressure introduces a strong low-frequency contribution in Fig. 4(a). When the pulse has entered the far-field region [Fig. 4(b)], the intermediate-field part has decayed. After propagating further for a few cm, the spike in the pulse (the indication of high frequencies) almost disappears. Also note the huge effect of dissipation evident in the comparison of the undamped and damped pulses in Fig. 2(c) and Fig. 4(c), respectively. The peak pressure of the damped pulse is only  $\sim 1/1000$  of that of the undamped pulse.

In diagnostic applications, the pulse usually travels through a variety of different tissues (blood, muscle, skin, etc.) with different absorption coefficients. In this case, it is straightforward to model an effective damping by modifying the parameters of Eq. (9). An exception to this is transcranial ultrasound diagnostics, because the frequency dependence of absorption in bone is more complicated.<sup>22</sup>

### III. ENERGY BALANCE

#### A. Total and effective scattering cross sections

Having computed a few examples of sound emission and subsequent absorption, we would now like to give a more complete account of the energy balance of the pulse-driven bubble. In particular, we would like to know what fraction of the initially emitted energy can actually reach an ultrasound detector, and which fraction gets absorbed, leading to a heating of the liquid.

It is common practice in studies of responses of bubbles to ultrasound to determine the effectiveness of scattering and/or absorption by giving cross sections.<sup>2,14,16,23</sup> To evaluate a scattering cross section  $\sigma_{sc}$ , we need the incident intensity (energy/area/time)  $I_{\text{inc}}$  and the scattered power  $W_{sc}$  (energy/time) to get

$$\sigma_{sc} = W_{sc} / I_{\text{inc}}. \quad (10)$$

$I_{\text{inc}}$  is determined by the incident pressure wave  $P(t)$ , namely (see, e.g., Ref. 21, Sec. 65)

$$I_{\text{inc}} = \frac{1}{\rho_l c_l} \langle P^2(t) \rangle_t, \quad (11)$$

and  $W_{sc}$  follows from  $P_s(r, t)$  through

$$W_{sc} = \frac{4\pi}{\rho_l c_l} \langle r^2 P_s^2(r, t) \rangle_t = \frac{4\pi}{\rho_l c_l} \langle q_s^2(t) \rangle_t. \quad (12)$$

Here,  $\langle \cdot \rangle_t$  denotes a time mean; we have exploited the virial theorem to express intensities and powers solely in terms of pressures (the ‘‘potential’’ part of the energy of the wave). Thus, with Eqs. (10)–(12) we have

$$\sigma_{sc} = \frac{4\pi}{\langle P^2(t) \rangle_t} \langle q_s^2(t) \rangle_t. \quad (13)$$

This formula, however, does not take acoustic damping into account, and therefore does not represent the cross section one would actually measure in experiment, with the detector at a distance  $r$ . After Fourier transformation of  $q_s(t)$  and using the equalities in Eqs. (7) and (8), we obtain the *effective* scattering cross section

$$\sigma_{sc}^{\text{eff}}(r) = \frac{4\pi\rho_l c_l}{\int_{-\infty}^{+\infty} P^2(t) dt} \int_{-\infty}^{+\infty} \varepsilon(r, f) df. \quad (14)$$

A number of other contributions to  $P_s(r, t)$  can occur and modify the scattering cross sections. Apart from the passive contribution mentioned above, direct and indirect sound emissions from shape oscillations of the bubble are possible. They are, however, not part of our model as we treat spherically symmetric oscillations only. Also, there is a contribution if the bubble (or scattering body in general) can be translated as a whole by the incident sound. This term results in an *additional* scattering cross section (cf. Ref. 21 Sec. 78 or Ref. 24) with a characteristic angular dependence. For a gas bubble in a liquid, it is always of the same order of magnitude as the passive contribution, and is therefore equally negligible. Finally, bubble *interactions* can modify the emitted sound field, either directly by scattering of the sound emitted by one bubble from another, or indirectly by modifying their modes of oscillation (e.g., via secondary Bjerknes forces<sup>25,26</sup>). The treatment of these effects for a large number of bubbles is extremely difficult and no theory of general applicability is known so far (cf. Ref. 14). We will therefore not try to incorporate these phenomena, but restrict ourselves to the case of a single bubble.

#### B. Viscous dissipation: absorption cross section

In analogy to Eq. (10), we can define the absorption cross section  $\sigma_{\text{abs}}^v$  due to viscous drag during the bubble oscillations by replacing  $W_{sc}$  by the dissipated energy  $W_{\text{dis}}^v$ . We want to evaluate this quantity to keep track of the energy balance of the bubble oscillator as a whole.

The liquid viscosity exerts a stress (drag force per area) of

$$p_{\text{vis}} = \frac{4\nu_l \rho_l \dot{R}}{R} \quad (15)$$

over the bubble surface of size  $A = 4\pi R^2$  [Eq. (15) can be read off directly from the Keller equation]. As the bubble wall moves with velocity  $\dot{R}$ , the dissipated power is

$$W_{\text{dis}}^v = p_{\text{vis}} A \dot{R} = 16\pi\nu_l \rho_l R \dot{R}^2, \quad (16)$$

which yields  $\sigma_{\text{abs}}^v$ .

#### C. Results: Cross sections for nonlinearly responding bubbles

Figure 5 compares numerical calculations of  $\sigma_{sc}$  and  $\sigma_{sc}^{\text{eff}}$  (5 cm) for driving pressure amplitudes from 1 to 33 atm (the latter corresponding to  $MI \approx 1.9$ ). The cross sections are normalized to the geometric cross section  $\pi R_0^2$ . If the oscillator responded linearly, the cross sections would be independent of  $P_a$ . In stark contrast to this, a tremendous in-

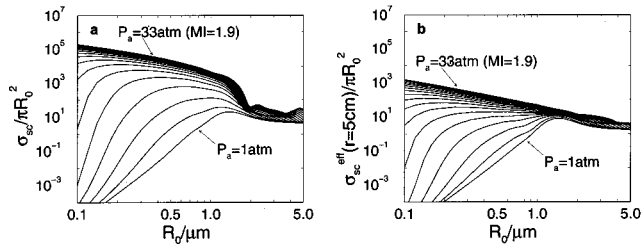


FIG. 5. Scattering cross sections as a function of  $R_0$ , parametrized with  $P_a = 1$  atm to  $P_a = 33$  atm ( $MI \approx 1.9$ ) in 2 atm steps (bottom to top). Cross sections are normalized to the geometric cross section  $\pi R_0^2$ . (a)  $\sigma_{sc}$ , computed from the total undamped emitted sound, increases dramatically with growing  $P_a$ . (b) Effective cross section  $\sigma_{sc}^{\text{eff}}$ , computed from the detectable sound energy after traveling through a blood layer of thickness 5 cm. Sound attenuation leads to severe decreases in cross section compared to  $\sigma_{sc}$ , especially for strongly emitting bubbles near nonlinear resonance.

crease in  $\sigma_{sc}$  is seen as  $P_a$  gets larger, the highest values being  $\sim 10^4$  times larger than the linear cross sections at low  $P_a$ . The violent collapses of strongly driven bubbles generate much more sound than would be expected in a linear study (cf. e.g., Ref. 23). However, the *effective* cross section shows that, in 5 cm distance, much of the initially emitted energy is lost again because of the dissipation in the medium.

It is precisely for the parameter combinations with a large nonlinear increase in the total scattered intensity [cf. Fig. 5(a)] that much of this radiated energy is emitted at very high frequencies and is thus not useful for detection in either scientific or diagnostic applications. Nevertheless, the total scattering cross sections that remain [see Fig. 5(b)] are still larger, in many cases considerably larger, than what would be expected from linear theory at small  $P_a$ .

The resonance peak  $R_0^{\text{res}}$  of the  $\sigma_{sc}(R_0)$  curve is shifted toward smaller  $R_0$  upon increasing  $P_a$  (cf. Ref. 27). This is due to the nonlinearity of the Keller equation: The resonance frequency decreases with increasing amplitude (cf. Ref. 6), and therefore a smaller  $R_0$  is needed in order to match the eigenfrequency of the bubble to the constant frequency of the driving (or its center frequency  $f_d = 3$  MHz in our case). There are also peak shifts due to viscous damping and the polychromaticity of the driving, the latter tending to favor larger  $R_0^{\text{res}}$ .<sup>6</sup> The nonlinear effect is, however, clearly dominant at large  $P_a$ . Yet again, when the dissipation of high frequency sound in the medium is taken into account, this effect is not as prominent [Fig. 5(b)].

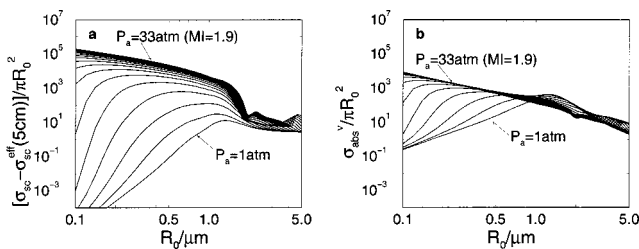


FIG. 6. (a) Difference of  $\sigma_{sc}$  and  $\sigma_{sc}^{\text{eff}}$  (5 cm) from Fig. 5. As  $\sigma_{sc} \gg \sigma_{sc}^{\text{eff}}$  (5 cm), the displayed difference is almost indistinguishable from  $\sigma_{sc}$  itself. (b) Normalized absorption cross section due to viscous damping  $\sigma_{\text{abs}}^{\nu}$ . Curves are parametrized with  $P_a = 1$  atm to  $P_a = 33$  atm in 2 atm steps. For low  $P_a$ , energy is dissipated primarily through viscous damping, but for high  $P_a$ , sound damping is much more effective.

The power dissipated due to sound absorption is proportional to the *difference* of  $\sigma_{sc}$  and  $\sigma_{sc}^{\text{eff}}$ . In Fig. 6(a), (b), we compare this quantity (the sound absorption cross section  $\sigma_{\text{abs}}^s$ ) to the viscous absorption cross section  $\sigma_{\text{abs}}^{\nu}$ . While the latter is larger (for small  $P_a$  (approximately linear responses), sound absorption is the much stronger cause of energy loss at high driving amplitudes. A closer look at Eqs. (5) and (16) reveals that the viscous term depends on the velocity of the bubble movement, while for the radiation process the acceleration of  $R$  is important. We have seen above that  $|\ddot{R}|$  can take on tremendous values when the bubble collapse is very strong, and this is why  $\sigma_{\text{abs}}^s$  can grow to be much larger than  $\sigma_{\text{abs}}^{\nu}$  for high  $P_a$ .

#### IV. MECHANISMS OF LIQUID HEATING

Loss of energy to the disordered molecular movements in the liquid is tantamount to heating the medium. For a risk estimate of bubble diagnostics, it is very desirable to have an approximation for the temperature rise that can be expected in the vicinity of a collapsing bubble. While a continuously driven cavitation activity will certainly heat up the bulk of the liquid considerably after some time, we are concerned here with the *local* effects, i.e., possibly very high temperatures very near an individual collapsing bubble. Heating can, in this context, be expected from three different processes: sound damping and viscous damping, as described above, and direct heating from the gas inside the bubble as its interior gets hot during the collapse, and conduction of heat into the liquid occurs.

##### A. Gas heating

Let us first estimate the order of magnitude of the temperature rise at a given distance from the bubble caused by the heating of the gas inside the bubble at collapse. From experiments in multi bubble sonoluminescence<sup>28</sup> and sonochemistry<sup>29</sup> we know that the maximum temperatures of the gas do not exceed  $\sim 6000$  K. It is only in well-controlled single bubble sonoluminescence experiments that more than  $10^4$  K can be reached, accompanied by intense light emission.<sup>30,31</sup> As no strong light emission is reported for diagnostic microbubble suspensions, it is safe to say that the temperature is below  $10^4$  K even for the strongest collapses in our study. This is in accord with previous work of Kamath *et al.*<sup>32</sup> for larger bubbles and smaller driving frequencies, where even more violent collapses would be expected, but only  $\sim 3000$  K are achieved. To estimate an upper bound for the maximum temperature outside the bubble, we assume that the *whole* bubble is heated to  $10^4$  K and that *all* of this heat is then expelled instantaneously into an infinitely thin shell at radius  $r = R_0$ .

This heat is then propagated according to the equation of thermal conduction. For this specific case, an analytical solution can be derived for the space- and time-dependent temperature rise  $\Delta T$  (cf. e.g., Ref. 33):

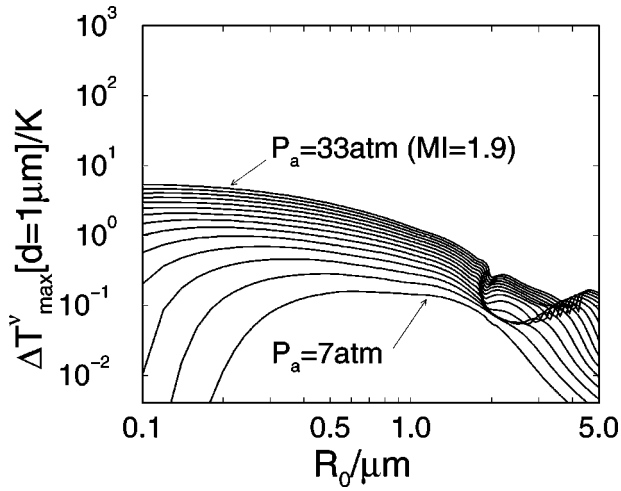


FIG. 7. Maximum temperature rises due to viscous heating  $\Delta T_{\max}^v$  according to Eq. (17) as a function of  $R_0$  and  $P_a$  (curves represent  $P_a = 7$  atm–33 atm in steps of 2 atm). These values are valid for a distance from the bubble radius at rest of  $d = r - R_0 = 1 \mu\text{m}$ .  $\Delta T^v$  stays below  $\approx 5$  K everywhere.

$$\Delta T(r, t) = \frac{E}{8\pi^{3/2}c_p^l\rho_l(\chi_l t)^{1/2}R_0} \frac{1}{r} \left[ \exp\left(-\frac{(r-R_0)^2}{4\chi_l t}\right) - \exp\left(-\frac{(r+R_0)^2}{4\chi_l t}\right) \right], \quad (17)$$

where  $E$  is the amount of energy released, and  $c_p^l, \chi_l$  are the specific heat and the thermal diffusivity of the liquid, respectively. For the estimates presented here, we take the values for water, i.e.,  $c_p^l = 4190$  J/(kg K) and  $\chi_l \approx 1.4 \cdot 10^{-7}$  m<sup>2</sup> s<sup>-1</sup>.

The total energy of the heated gas is readily approximated by

$$E_{\text{gas}} \approx \frac{4\pi R_0^3}{3V_m} C_v T_{\text{gas}}, \quad (18)$$

with the ideal gas mole volume  $V_m = 0.0224$  m<sup>3</sup> mol<sup>-1</sup>, the molar heat capacity at constant volume  $C_v = 2.5$  G (G = universal gas constant), and  $T_{\text{gas}} = 10^4$  K. Inserting  $E = E_{\text{gas}}$  in Eq. (17), it is now straightforward to compute the maximum temperature experienced at a given distance  $d \equiv r - R_0$  as the temperature pulse  $T(r, t)$  travels through the liquid. For near-resonant bubbles with  $R_0 \approx 1 \mu\text{m}$ , this yields a maximum temperature rise of less than 0.2 K in the liquid even at a distance as small as  $d = 1 \mu\text{m}$ . This can certainly be regarded as safe for clinical use: Note that we have most probably overestimated this value, and that these temperature rises last for only  $\approx 10^{-5}$  s (the time needed for the temperature pulse to travel the distance 1  $\mu\text{m}$  with the  $\chi_l$  given above). Farther away from the bubble, the temperature rises drop off rapidly to even smaller values. The rather crude estimate presented here has mainly served to introduce Eq. (17), of which we will make use in subsequent sections.

## B. Viscous heating

The viscous contribution to dissipation is only pronounced for high bubble wall velocities  $|\dot{R}|$  [cf. Eq. (16)], i.e., near the violent collapses. The whole energy dissipated during the pulse  $E_{\text{dis}}^v = \int_{-\infty}^{\infty} W_{\text{dis}}^v dt$  is therefore released over

a time scale of less than a microsecond (distance from first to last collapse). Within this time, a temperature pulse can travel in water (with the thermal diffusivity given above) only about 0.3  $\mu\text{m}$ . Therefore, we can assume the heat to be “frozen” at the bubble wall until the oscillation has ceased and again treat the dissipated energy as a heat source concentrated in an infinitely thin shell at  $r = R_0$ . Substituting  $E = E_{\text{dis}}^v$  in Eq. (17), and solving for the maximum temperature rise at  $d = 1 \mu\text{m}$ , yields the values  $\Delta T_{\max}^v$  shown in Fig. 7. Note that we do not display the very small values of  $\Delta T_{\max}^v$  for  $P_a < 7$  atm. Even for the highest drivings ( $P_a \approx 30$  atm), the temperature elevations at  $d$  due to viscous losses stay below  $\approx 5$  K. The same comments regarding the short duration and rapid decay of the temperature pulses given in the previous section apply here. Thus, viscous heating around the bubble should not be a major risk factor, either.

## C. Secondary acoustic heating

We have shown above that for small, nonlinearly driven bubbles the energy dissipated through absorption of the emitted sound can be far greater than the energy dissipated by viscosity. We now determine the impact of this effect on the liquid temperature very near the bubble, i.e., at radii  $r$  of a few  $\mu\text{m}$ . In order to give analytical estimates for the expected temperature rise, we use a separation of three time scales inherent in our system on these length scales: (i) the *sound* time scale  $\tau_s \equiv r/c_l \sim 1$  ns in which the emitted sound pulse travels a distance  $r$ ; (ii) the *dynamical* time scale  $\tau_d \equiv f_d^{-1} \sim 1 \mu\text{s}$  on which the bubble oscillates (comparable to the duration of the short pulse of driving); and (iii) the *thermal diffusive* time scale  $\tau_{td} \equiv r^2/\chi_l \sim 10 \mu\text{s}$  over which heat spreads in a quiescent liquid over a radius  $r$ .

The separation of time scales  $\tau_s \ll \tau_d \ll \tau_{td}$ , which for our values of  $c_l, f_d$  is valid for  $0.2 \mu\text{m} \ll r \ll \lambda$ , suggests the following three-stage model of liquid heating: (i) The emitted sound wave imprints a temperature profile outside the bubble, depositing heat almost instantaneously (on a  $\tau_s$  scale). (ii) The bubble oscillation ceases after  $\tau_d$ , which results in a quiescent bubble with a hot shell. (iii) Finally, the heat from the shell spreads outward according to the laws of heat conduction.

Let us first compute the initial profile induced by the sound pulse: From Eq. (8) it follows that, upon traveling to radius  $r$ , the scattered power is diminished by  $\varepsilon_{\text{dis}}^s(r, f) \equiv \varepsilon_0(f)[1 - \exp(-\alpha(f)r)]$ , so that the dissipated energy across a line element  $dr$  at  $r$  is

$$d\varepsilon_{\text{dis}}^s(r, f) = \alpha(f)\varepsilon_0(f)\exp[-\alpha(f)r]dr. \quad (19)$$

This energy is deposited into a spherical shell of mass  $dm = 4\pi\rho_l r^2 dr$ . Taking all frequency components into account, the resulting temperature rise at radius  $r$  is

$$\begin{aligned} \Delta T^s(r, 0) &= \frac{dr}{c_p^l dm} \int_{-\infty}^{+\infty} \alpha(f)\varepsilon_0(f)\exp[-\alpha(f)r]df \\ &= \frac{\alpha_0}{c_p^l \rho_l^2 c_l} \frac{1}{r^2} \int_{-\infty}^{+\infty} \left(\frac{f}{\text{MHz}}\right)^\beta |q_s(f)|^2 \\ &\quad \times \exp[-\alpha_0(f/\text{MHz})^\beta r] df. \end{aligned} \quad (20)$$

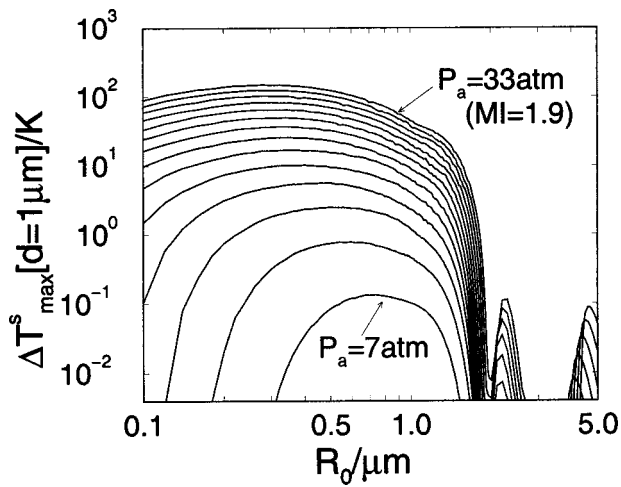


FIG. 8. Maximum temperature rises due to sound heating  $\Delta T_{\max}^s$  as a function of  $R_0$  and  $P_a$  (curves signify  $P_a=7$  atm–33 atm in 2 atm steps) at a distance  $d=1$   $\mu\text{m}$  from the bubble wall. For  $R_0$  near resonance and  $P_a \gtrsim 30$  atm,  $\Delta T_{\max}^s$  reaches more than 100 K.

The initial temperature profile calculated in Eq. (20) decays extremely rapidly with  $r$ . In fact, at a distance as near as  $d=1$   $\mu\text{m}$ ,  $\Delta T^s(r,0) < 0.1$  K even for the highest  $P_a$ . Therefore, we can introduce the same approximation as above to evaluate the maximum temperatures at a distance  $d$ : We assume that all of the absorbed heat from the emitted sound pulse up to distance  $d$  is delta-distributed in a shell at radius  $R_0$  (as the bubble motion has ceased after  $\sim \tau_d$ ). Thus, we replace  $E$  in Eq. (17) by the heat from the absorbed sound, namely,

$$E_{\text{dis}}^s(d) = \int_{R_0}^{R_0+d} dr \int_{-\infty}^{+\infty} df \alpha(f) \varepsilon_0(f) \exp[-\alpha(f)r]. \quad (21)$$

The resulting maximum temperature rises  $\Delta T_{\max}^s$  ( $d=1$   $\mu\text{m}$ ) of this computation are displayed in Fig. 8. While for moderate  $P_a \lesssim 15$  atm,  $\Delta T_{\max}^s$  stays below 10 K, the heating is quite substantial at the highest  $P_a$ . The temperature rises for very small bubbles near the nonlinear resonance are over 100 K for driving pressures  $P_a \approx 30$  atm ( $MI \approx 1.9$ ) at the upper limit of the diagnostic range. We have to caution against these values in two ways: first, as mentioned above, this temperature can only last for about 10 microseconds; it is not clear (to us) what kind of biological damage can result from such short temperature pulses. Second, we have calculated the effects of a single undisturbed, spherical bubble. It is well-known that—except in the well-controlled environment of a single-bubble sonoluminescence experiment<sup>30,34</sup>—bubble collapses tend to be asymmetrical, i.e., the bubble surface gets distorted. This weakens the collapses considerably, which will affect the sound emission as well. Thus, these calculations must be seen as an upper limit or worst-case estimate for the actual heating.

If  $q_s(f)$  in Eq. (20) is replaced by the corresponding Fourier transform of  $rP_s^{\text{eff}}(r,t)$  from Eq. (6), we would expect more accurate values for  $\Delta T_{\max}^s$ . Indeed, one would think that using the far-field pressure Eq. (5) at a distance of 1  $\mu\text{m}$  must be plain wrong. Instead, a calculation with the

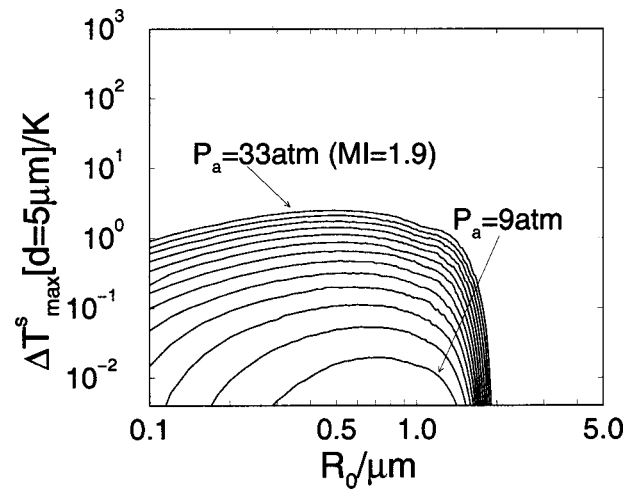


FIG. 9. Same as Fig. 8, but the sound heating is now evaluated at a distance  $d=5$   $\mu\text{m}$  from the bubble (curves represent  $P_a=9$  atm–33 atm in steps of 2 atm). Even the most extreme heating for  $P_a=33$  atm and resonant bubble sizes stays below 3 K.

intermediate-field pressure correction shows no substantial changes in the range of high  $P_a$  which lead to appreciable temperature rises.<sup>19</sup> How is that possible? Figure 3 has shown the ultra-short (sub-ns) time scales of the bubble acceleration  $\ddot{R}$ . In contrast, the bubble wall velocity  $\dot{R}$  changes on much slower time scales. As the intermediate-pressure correction in Eq. (6) is proportional to  $\dot{R}$ , while the far-field pressure contains a term proportional to  $\ddot{R}$ , ultra-high frequencies ( $f \gtrsim 1$  GHz) are almost exclusively emitted from the far-field part of the pressure. But only these high frequencies are responsible for heating at  $\mu\text{m}$  distances from the bubble wall, as lower frequencies are not damped as strongly [cf. Eq. (9)]. Thus, the far-field term is, paradoxically, the dominant factor if one wants to compute the heating very close to the bubble.<sup>19</sup> The influence of the intermediate-pressure term becomes larger when the driving pressure amplitudes are smaller ( $\lesssim 5$  atm), because then the collapses are less violent and the time scales of  $\ddot{R}$  are longer. In this regime, however, the temperature rises are well below 1 K for  $d=1$   $\mu\text{m}$ .

If the distance from the bubble is even slightly larger, the values of  $\Delta T_{\max}^s$  drop quickly. Already at  $d=5$   $\mu\text{m}$ , we obtain no value above 3 K even for  $P_a=33$  atm (see Fig. 9). Our conclusion is therefore that any damage which could be done to biomolecules by the passage of the temperature pulses through the liquid would be limited to extremely small fluid volumes around the bubbles.

We stress that the risk evaluated here is the *local* thermal heating due to the pressure wave emitted at collapse. After a sufficient number of collapses from many bubbles, the heating of the bulk liquid will obviously become pronounced, as every bubble collapse deposits an energy on the order of  $10^{-10}$  J in the liquid. Driven by pulses described by Eq. (3) at a pulse repetition rate of 10 kHz and at a bubble volume fraction of  $10^{-6}$ , a constant insonation of 1 min duration would, for example, result in a bulk temperature rise of  $\approx 3$  K, assuming the heat capacity of water to be valid for blood. This rather high value does not take into account ac-

tive cooling of the insonated volume or the transport of liquid out of the transducer focus, as it would occur with circulating blood. Moreover, as stated above, most bubbles will not converge spherically and therefore will give off less heat to the liquid, so that this value for bulk heating is, again, an upper limit. Similar numbers have been obtained, however, with a very different model by Wu,<sup>35</sup> which supports the validity of these estimates.

It is difficult to distinguish experimentally between localized heating and bulk heating. We propose to insonate a suspension of microbubbles containing a protein (or other temperature-sensitive molecules), and ensure by efficient cooling that the bulk temperature does not rise above the denaturation temperature of the protein. After sufficient exposure, the localized ‘‘hot spots’’ resulting from bubble collapses should generate a detectable amount of denatured protein.

## V. CONCLUSIONS AND OUTLOOK

We have shown that the full treatment of nonlinear bubble oscillations is vital for the assessment of effectiveness of microbubble contrast agents for ultrasound diagnostics. The scattering cross sections of the bubbles are considerably enhanced as compared to the case of linear response, and would be severely underestimated if modeled by a linear theory. The sound emission occurs as a sharp pressure spike at the time of bubble collapse. If very short incident (driving) sound pulses of  $\sim 3$  wavelengths of spatial extension are used, there is often exactly one pressure spike response of the bubble.

For large  $P_a$ , most of the power is emitted at very high or even hypersonic frequencies. Not only is this radiation useless for detection with diagnostic devices, but the high frequency sound is *damped out* already at very short distances from the bubble. Although the remaining intensity at smaller frequencies that can reach a detector is still larger for larger  $P_a$ , the sound absorption makes the effective cross sections much smaller than what would be expected from the energy emitted into the whole spectrum. The absorption of sound is the primary reason for heating of the liquid near the bubble. The temperature rises induced by heat dissipation are tiny for very small and large bubbles far from resonance. However, the violently collapsing bubbles near resonance, when driven by diagnostic pressures near the maximum allowed mechanical index ( $MI=1.9$ ,  $P_a \approx 33$  atm), induce considerable ( $\sim 100$  K) heating over  $\mu\text{m}$  distances. The temperatures remain this high only for  $\leq 10$   $\mu\text{s}$  and fall off rapidly with growing distance from the bubble. The short duration of high temperatures probably alleviates adverse biological effects, as it may rather be the total energy absorbed by a macromolecule or a blood cell that is a measure for potential thermal damage.

A major problem to treat in future studies is the question of *bubble shape stability*. We have investigated the *parametric* instability of the bubble wall against perturbations of its spherical shape in Ref. 5. This mechanism relies on amplification of the perturbations over several oscillation cycles. This is possible with monochromatic driving, as in Ref. 5, but not to the same extent with a pulsed  $P(t)$ . In this latter

case, one has to concentrate on the *Rayleigh–Taylor* (RT) instability, which acts on much smaller time scales and is able to disrupt the bubble within a fraction of the oscillation cycle, provided the collapse is strong enough.<sup>13</sup> This mechanism is likely to be relevant in bubble contrast agent experiments for two reasons. (i) In the case we have focused on, *coated* bubbles are used, with a fairly rigid shell that must first be broken by the ultrasound pulse before the gas can start to respond to it. The coated bubbles are not exactly spherical, so that the bubble starts its oscillation with a substantial deviation from the spherical form. This facilitates the destruction of the bubble via RT instability. (ii) The peak pressures used in diagnostic applications can be much higher than the threshold pressures for the onset of shape instability identified in a perturbation analysis.<sup>5</sup>

Another important extension of the present theory concerns the treatment of bubble–bubble interactions. Secondary Bjerknes forces, streaming, or shadowing modify the dynamics of the bubbles and thus the emitted sound. In particular, bubble–bubble interactions can provide perturbations which lead to the aforementioned nonspherical collapses. Compared to perfect radial collapses, the energy densities achieved in the bubble are much smaller in this case. While some bubbles will still approximate the ideal, spherical case, and have cross sections and temperature rises comparable to the ones calculated here, the majority of bubbles will show substantially reduced sound emission and heating. Another factor contributing to smaller signal intensities is shadowing, i.e., the increased attenuation of the incoming sound wave by a dense cloud of microbubbles. Marsh *et al.*<sup>36</sup> have shown that the measured backscatter from a dense suspension of bubbles can be *less* than that from a dilute suspension, despite the much larger number of oscillators present. We therefore expect that a comprehensive many-bubble description of the system will compute detected signals much smaller than the sum of the emissions that would result from the individual bubbles. However, cooperative effects can also lead to more intense emission: Kuttruff<sup>37</sup> has demonstrated that the collapses of several bubbles tend to synchronize, and thus increase the intensity of the response.

We hope that with the present work, we have provided useful predictions for researchers in bubble dynamics and ultrasound diagnostics to estimate parameter regions suitable for their purposes. The amount of sound energy scattered from a bubble suspension, as well as the energy dissipated into heat, cannot be evaluated without incorporating nonlinear bubble dynamics, sound absorption, and sound heating, as it was done in this study.

<sup>1</sup>R. Gramiak and P. M. Shah, *Invest. Radiol.* **3**, 356–366 (1968).

<sup>2</sup>See the articles in *Advances in Echo Imaging Using Contrast Enhancement*, edited by N. C. Nanda and R. Schlieff (Kluwer, Dordrecht, 1993).

<sup>3</sup>D. W. Schwendeman, *Proc. R. Soc. London, Ser. A* **441**, 331–341 (1993); P. N. Burns, *Clinical Radiology Suppl.* **1** **51**, 50–55 (1996); B. Ward, A. C. Baker, and V. F. Humphrey, *J. Acoust. Soc. Am.* **101**, 143–154 (1997).

<sup>4</sup>N. deJong, R. Cornet, and C. T. Lancee, *Ultrasonics* **32**, 447–453 (1994); *ibid.* **34**, 587–590 (1996); P. J. A. Frinking, N. de Jong, and E. I. Cepedes, *J. Acoust. Soc. Am.* **105**, 1989–1996 (1999).

<sup>5</sup>S. Grossmann, S. Hilgenfeldt, D. Lohse, and M. Zomack, *J. Acoust. Soc. Am.* **102**, 1223–1230 (1997).

- <sup>6</sup>S. Hilgenfeldt, D. Lohse, and M. Zomack, *Eur. Phys. J. B* **4**, 247–255 (1998).
- <sup>7</sup>J. B. Keller and I. I. Kolodner, *J. Appl. Phys.* **27**, 1152–1161 (1956).
- <sup>8</sup>J. B. Keller and M. J. Miksis, *J. Acoust. Soc. Am.* **68**, 628–633 (1980).
- <sup>9</sup>A. Prosperetti and A. Lezzi, *J. Fluid Mech.* **168**, 457–478 (1986).
- <sup>10</sup>C. E. Brennen, *Cavitation and Bubble Dynamics* (Oxford University Press, Oxford, 1995).
- <sup>11</sup>R. Löfstedt, B. P. Barber, and S. J. Putterman, *Phys. Fluids A* **5**, 2911–2928 (1993).
- <sup>12</sup>M. S. Plesset and A. Prosperetti, *Annu. Rev. Fluid Mech.* **9**, 145–185 (1977).
- <sup>13</sup>S. Hilgenfeldt, D. Lohse, and M. P. Brenner, *Phys. Fluids* **8**, 2808–2825 (1996).
- <sup>14</sup>T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1996).
- <sup>15</sup>S. Hilgenfeldt, M. P. Brenner, S. Grossmann, and D. Lohse, *J. Fluid Mech.* **365**, 171–204 (1998).
- <sup>16</sup>C. C. Church, *J. Acoust. Soc. Am.* **97**, 1510–1521 (1995).
- <sup>17</sup>R. T. Beyer, *Nonlinear Acoustics* (Ed. Naval Sea Systems Command, 1974).
- <sup>18</sup>E. J. Aymé-Bellegarda, *J. Acoust. Soc. Am.* **88**, 1054–1060 (1990).
- <sup>19</sup>S. Hilgenfeldt and D. Lohse, “The acoustics of diagnostic microbubbles: Dissipative effects and heat deposition,” *Ultrasonics* (to appear).
- <sup>20</sup>R. E. Apfel and C. K. Holland, *Ultrasound Med. Biol.* **17**, 179–185 (1991).
- <sup>21</sup>L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1987).
- <sup>22</sup>*Physical Principles of Medical Ultrasonics*, edited by C. R. Hill (Wiley, New York, 1986).
- <sup>23</sup>R. Y. Nishi, *Acustica* **33**, 65–74 (1975).
- <sup>24</sup>S. Meerbaum, in *Advances in Echo Imaging Using Contrast Enhancement*, edited by N. C. Nanda and R. Schlieff (Kluwer, Dordrecht, 1993), p. 9–42.
- <sup>25</sup>V. Bjerknes, *Die Kraftfelder* (Friedrich Vieweg, Braunschweig, 1909).
- <sup>26</sup>A. Prosperetti, *Ultrasonics* **22**, 115–124 (1984).
- <sup>27</sup>J. S. Allen, R. A. Roy, and C. C. Church, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 743–751 (1997).
- <sup>28</sup>E. B. Flint and K. S. Suslick, *J. Am. Chem. Soc.* **111**, 6987–6992 (1989).
- <sup>29</sup>K. S. Suslick, *Science* **247**, 1439–1445 (1990); K. S. Suslick, “The chemistry of ultrasound,” in *Encyclopaedia Britannica, Yearbook of Science and the Future* (London, 1994), p. 138.
- <sup>30</sup>B. P. Barber, R. A. Hiller, R. Löfstedt, S. J. Putterman, and K. R. Weninger, *Phys. Rep.* **281**, 65–143 (1997); R. Hiller, S. J. Putterman, and B. P. Barber, *Phys. Rev. Lett.* **69**, 1182–1184 (1992).
- <sup>31</sup>S. Hilgenfeldt, S. Grossmann, and D. Lohse, *Nature (London)* **398**, 402–405 (1999); *Phys. Fluids* **11**, 1318–1330 (1999).
- <sup>32</sup>V. Kamath, A. Prosperetti, and F. N. Egofoopoulos, *J. Acoust. Soc. Am.* **94**, 248–260 (1993).
- <sup>33</sup>H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon, Oxford, 1989).
- <sup>34</sup>D. F. Gaitan, L. A. Crum, R. A. Roy, and C. C. Church, *J. Acoust. Soc. Am.* **91**, 3166–3183 (1992).
- <sup>35</sup>J. Wu, *Ultrasound Med. Biol.* **24**, 267–274 (1998).
- <sup>36</sup>J. N. Marsh, M. S. Hughes, C. S. Hall, S. H. Lewis, R. L. Trousil, G. H. Brandenburger, H. Levene, and J. G. Miller, *J. Acoust. Soc. Am.* **104**, 1654–1666 (1998).
- <sup>37</sup>H. Kuttruff, *J. Acoust. Soc. Am.* **106**, 190–194 (1999).

# The impact of sound speed errors on medical ultrasound imaging

M. E. Anderson

*Department of Biomedical Engineering, University of Rochester, Box 270126, Rochester, New York 14627-0126*

M. S. McKeag, and G. E. Trahey

*Department of Biomedical Engineering, Duke University, Box 90281, Durham, North Carolina 27708*

(Received 23 November 1999; accepted for publication 15 February 2000)

The results of a quantitative study of the impact of sound speed errors on the spatial resolution and amplitude sensitivity of a commercial medical ultrasound scanner are presented in the context of their clinical significance. The beamforming parameters of the scanner were manipulated to produce sound speed errors ranging over  $\pm 8\%$  while imaging a wire target and an attenuating, speckle-generating phantom. For the wire target, these errors produced increases in lateral beam width of up to 320% and reductions in peak echo amplitude of up to 10.5 dB. In the speckle-generating phantom, these errors produced increases in speckle intensity correlation cell area of up to 92% and reductions in mean speckle brightness of up to 5.6 dB. These results are applied in statistical analyses of two detection tasks of clinical relevance. The first is of low contrast lesion detectability, predicting the changes in the correct decision probability as a function of lesion size, contrast, and sound speed error. The second is of point target detectability, predicting the changes in the correct decision probability as function of point target reflectivity and sound speed error. Representative results of these analyses are presented and their implications for clinical imaging are discussed. In general, sound speed errors have a more significant impact on point target detectability over lesion detectability by these analyses, producing up to a 22% reduction in correct decisions for a typical error. © 2000 Acoustical Society of America. [S0001-4966(00)01206-6]

PACS numbers: 43.80.Vj, 43.80.Ev, 43.58.Dj [FD]

## INTRODUCTION

The beamforming process underlying image formation in medical ultrasound relies on an assumed constant speed of sound *in vivo*. The speed typically used is 1540 m/s, despite the fact that in soft tissues this property is known to range from approximately 1450 m/s in fat to over 1600 m/s in muscle.<sup>1-3</sup> Ultrasound is thus routinely used under conditions producing sound speed errors on the order of 5%. These sound speed errors significantly reduce the contrast and spatial resolution of ultrasound. Dynamic receive focusing compounds the impact of these errors: the focusing delays applied over range are based not only on an erroneous speed but also on erroneous range calculations. While the presence of these errors and their general effects are well known in the clinical and ultrasound research communities, the magnitude of their impact on contrast and spatial resolution has not been well characterized.

We describe two studies of the impact of sound speed errors on these image parameters. By programming the beamformer of a commercial scanner, we captured echo data from both a wire target and a speckle-generating phantom across a range of assumed sound speeds above and below the actual speed. We then analyzed these data and compared them to control data acquired at the correct speed to quantify the effect of these errors on the system spatial resolution and amplitude sensitivity as well as the observed speckle statistics.

In discussing the impact of these errors on clinical im-

aging tasks, we consider two statistical detection problems. The first is the detection of circular low contrast lesions by a human observer. This task is thoroughly characterized in the literature, and has been discussed in the context of other types of beamforming distortion. The second is the detection of a point target within a speckle field on the basis of its amplitude by an ideal observer. This analysis has been applied in the context of microcalcification detection during breast imaging with medical ultrasound, although it is also relevant to other diagnostic aims.

## I. EXPERIMENTAL SETUP AND DATA ACQUISITION

For all experiments we used the 7.5-MHz linear phased array (model 7.5240) of a Siemens Elegra scanner (Siemens Medical Systems, Inc., Issaquah, WA) mounted on a 3-axis computer-controlled translation stage ( $2 \times$ UTM100CC.1 and  $1 \times$ UTM50CC.1 with MM3000-OPT controller, Newport Corporation, Irvine, CA) and recorded summed RF echo data sampled at 36 MHz. This array has a 28.2-mm receive aperture and 60% relative  $-6$ -dB bandwidth. For the speckle statistics study we acoustically coupled this array to a speckle-generating, attenuating, phantom using a thin layer of deionized water. This custom-built phantom is agar-graphite in 10% vol/vol *n*-propanol in water, with an attenuation of  $0.53 \times f^{1.04}$  dB/cm-MHz,  $f$  in MHz, and sound speed of 1546.8 at 22 °C. The phantom's acoustic window is 13- $\mu$ m polyvinylidene chloride film (Saran Wrap). We assumed the temperature coefficient of this phantom to be

equal to that of water, i.e.,  $3.05 \text{ m/s/}^\circ\text{C}$ .<sup>4,5</sup> To estimate the temperature within the phantom, we monitored the temperature of the water layer with a calibrated digital thermometer with  $0.015^\circ\text{C}$  accuracy (model HH42 with ON-402-PP probe, NIST traceable calibration, Omega Engineering, Stamford, CT). We used the stage to move the transducer along the elevation dimension between acquisitions in order to capture multiple images having independent speckle patterns. Before collecting data, we set the scanner gain to maximize the signal-to-noise ratio without producing saturation. To provide relatively constant receive amplitude over range, we set a smooth time-gain compensation (TGC) curve, increasing with range. Both the overall gain and TGC curve were set before the beginning of the phantom study and were not altered thereafter.

Before beginning the phantom study, we performed a preliminary experiment to ensure that the images acquired in the phantom would be statistically independent. We captured and stored the summed RF echo data for a series of images (“frames”) of the phantom separated by  $50\text{-}\mu\text{m}$  translations in the elevation dimension. We then calculated the normalized correlation coefficient ( $\rho$ ) of our region of interest (ROI) in each frame relative to that in the first, thus indicating the degree of correlation as a function of translation. This analysis, equivalent to estimating the normalized autocovariance function of the observed RF speckle in the elevation dimension, indicated that elevation translations of  $\geq 2 \text{ mm}$  would produce frames that are decorrelated to  $\rho \approx 0.01$ . Our selection of the ROI is described in Sec. II.

For the phantom study, we initially set the scanner to the phantom’s sound speed and collected 20 summed RF frames separated in elevation by the previously proscribed translation of  $2 \text{ mm}$ . (Throughout this paper, we will refer to data collected at the correct sound speed as the control.) We then repeated this process with the scanner beamformer set to apply each of 8 sound speed errors:  $\pm 1, 2, 4, \text{ and } 8\%$ . Finally, we returned the stage to the starting position and acquired one final frame at the original, correct sound speed to be used as an experimental control against positioning error.

Given that our procedure relied on the spatial reproducibility of the translation stage to return repeatedly to the same 20 positions, we included an experimental control against spatial registration error. Using the normalized correlation coefficient, we compared the very first data frame to the control frame acquired at the very end, and found the RF data in these frames were correlated to  $\rho = 0.99$ . This correlation coefficient is in fact the maximum correlation value found after axial interpolation of the data by a factor of eight and a limited axial search, such as one would perform to axially align RF data.<sup>6</sup> We found an apparent axial translation of  $19 \mu\text{m}$ . This apparent axial translation was most likely due to the temperature drift of the phantom arising from acoustic energy absorption over the duration of the experiment (59 minutes), which would produce slight changes in the phantom’s sound speed over time. The temperature readings made during the study showed an increase of  $0.18^\circ\text{C}$ . Given a mean temperature of  $20.6^\circ\text{C}$  and an assumed sound speed temperature coefficient of  $3.03 \text{ m/s/}^\circ\text{C}$  and a focal range of  $30 \text{ mm}$ , this temperature drift would

correspond to an apparent  $11\text{-}\mu\text{m}$  axial shift. Given that we could only place the temperature probe in the surface water layer, we expect some discrepancy between our temperature readings and the true interior temperature of the phantom in the region of insonification.

For the wire target study, we suspended a  $100\text{-}\mu\text{m}$  steel wire in a small tank of deionized water and imaged this target with the same transducer and stage. We placed the wire target and the transmit focus at  $30\text{-mm}$  range, with the wire aligned perpendicular to the imaging plane. (We positioned the wire target using the scanner’s range calipers, which assumed a sound speed of  $1540 \text{ m/s}$ . We subsequently examined the echo data and found the actual wire target range to be  $28.9 \text{ mm}$ . This agrees with an expected sound speed in the water of  $1484 \text{ m/s}$ , given a mean temperature of  $21.3^\circ\text{C}$ .) The bottom of the tank was lined with an absorbent rubber that is acoustically matched to water. We fixed the tank over a magnetic stirrer and placed a stir-bar in the measurement tank to ensure a homogeneous temperature throughout. To measure the system pulse-echo point spread function (PSF), we translated the transducer laterally over a distance of  $6 \text{ mm}$  in  $50\text{-}\mu\text{m}$  steps. After each step we acquired a summed RF echo frame. We selected a single image line and extracted it from each of this series of frames to construct the PSF. This procedure was repeated for each beamformer sound speed setting ( $\pm 1, 2, 4, \text{ and } 8\%$  error). We measured the water temperature immediately before each PSF acquisition sequence using the digital thermometer described above. We used each temperature reading to calculate the current sound speed in the water tank and programmed the beamformer accordingly for each sound speed error. We based this calculation on a 5th order polynomial fit of sound speed in water versus temperature.<sup>4</sup> The maximum temperature drift recorded during any PSF acquisition was  $0.15^\circ\text{C}$ , corresponding to an unintentional relative sound speed error of  $0.03\%$ .

## II. DATA ANALYSIS AND MATCHED SIMULATIONS

We used the data collected in the speckle-generating phantom to calculate first and second order speckle statistics as a function of sound speed error. We restricted our statistical analysis to a region of interest (ROI) within each frame at the center of the field of view, centered at the transmit focus and within the predicted transmit depth of field given Hamming apodization across both the  $15\text{-mm}$  transmit aperture the  $28\text{-mm}$  receive aperture.<sup>7</sup> This corresponded to a region  $14.1 \text{ mm}$  laterally and  $5.8 \text{ mm}$  axially. Our subsequent calculation of correlation cell areas indicate that this ROI contained 1112 independent speckle cells in the control case and 586 cells in the most degraded experimental case. All of the RF echo data frames were envelope-detected using the Hilbert transform.<sup>8</sup> The average speckle signal-to-noise ratio (mean magnitude/standard deviation of magnitude) for all the data sets was  $1.87$  with a standard error of  $\pm 0.002$ , which is in good agreement with the expected value of  $1.91$ .<sup>9</sup>

While a typical B-mode image presents the envelope-detected echo data (i.e., speckle magnitude), much of the literature of speckle statistics and lesion detectability has applied the statistics of echo intensity for reasons of analytical



TABLE I. Comprehensive table of experimental results, all normalized to the corresponding control case ( $c_{as}/c=1$ ). The columns are the ratio of assumed sound speed to actual sound speed  $c_{as}/c$ , the FWHM beam width measured from the wire target, the peak echo amplitude measured from the wire target, the mean estimated speckle intensity correlation cell area  $\hat{S}_c$ , the mean estimated lateral FWHM of the speckle intensity autocorrelation function  $\hat{C}_{gg}(x,0)$ , the mean estimated axial FWHM of the speckle intensity autocorrelation function  $\hat{C}_{gg}(0,z)$ , the mean speckle amplitude  $\mu_0$ , and the mean effective Rician  $k$  parameter. Each mean is for  $N=20$ .

$c_{as}/c$	Beam width	Peak	$\hat{S}_c$	$\hat{C}_{gg}(x,0)$	$\hat{C}_{gg}(0,z)$	$\mu_0$	$k$
0.92	2.344	0.335	1.418	1.321	0.979	0.526	0.636
0.96	2.031	0.500	1.277	1.186	0.981	0.723	0.692
0.98	1.484	0.726	1.091	1.041	0.993	0.891	0.816
0.99	1.062	0.917	1.051	1.002	0.999	0.965	0.950
1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.01	1.266	0.823	1.022	1.045	1.000	0.976	0.843
1.02	1.906	0.612	1.118	1.134	0.998	0.914	0.670
1.04	2.484	0.433	1.510	1.464	0.985	0.761	0.569
1.08	3.234	0.298	1.927	1.824	0.979	0.588	0.506

expediency.<sup>10</sup> The autocovariance functions of speckle in magnitude and intensity have been shown in theory and experiment to be very similar.<sup>10,11</sup> For these reasons we have calculated our second order speckle statistics for intensity. To estimate the speckle intensity correlation cell area  $S_c$ , we first estimated the two-dimensional autocovariance function  $C_{gg}$  of each frame of the envelope-detected, squared speckle data. For a two-dimensional real function  $g$  with mean value  $\mu_g$ , coordinate vector  $\mathbf{x}=(x,z)$ , and translation  $\boldsymbol{\delta}=(\delta_x, \delta_z)$ , the autocovariance function is defined:

$$C_{gg}(\boldsymbol{\delta}) = \langle (g(\mathbf{x}) - \mu_g)(g(\mathbf{x} - \boldsymbol{\delta}) - \mu_g) \rangle, \quad (1)$$

where  $\langle \rangle$  denotes expectation. Our estimate  $\hat{C}_{gg}$  of  $C_{gg}$  over a window of size  $D_x$  by  $D_z$  extended with zero padding follows from

$$\hat{C}_{gg}(\boldsymbol{\delta}) = \frac{1}{D_x - |\delta_x|} \frac{1}{D_z - |\delta_z|} \int_{\delta_x}^{D_x} \int_{\delta_z}^{D_z} (g(\mathbf{x}) - \mu_g) \times (g(\mathbf{x} - \boldsymbol{\delta}) - \mu_g) dx dz. \quad (2)$$

This is derived from Eq. (3) in Ref. 12. The normalization terms before the integral are necessary to counteract bias introduced by integration over zero-padded regions. We chose to apply Fourier techniques to this calculation. Given our discretized, zero-padded data matrix  $g_p$  sampled at spacing  $(\Delta x, \Delta z)$ , Eq. (2) becomes

$$\hat{C}_{gg}(m \Delta x, n \Delta z) = \frac{1}{D_x - |\delta_x|} \frac{1}{D_z - |\delta_z|} C_{gg \text{ biased}}(m \Delta x, n \Delta z), \quad (3)$$

where

$$C_{gg \text{ biased}} = \text{Real}[\mathbf{F}^{-1}(\mathbf{F}(g_p - \mu_g)\mathbf{F}^*(g_p - \mu_g))],$$

with  $\mathbf{F}$  and  $\mathbf{F}^{-1}$  denoting the forward and inverse discrete Fourier transforms, and  $*$  denoting the complex conjugate. Our average estimates of the lateral and axial  $\hat{C}_{gg}$  FWHMs are listed in Table I, normalized to the control.

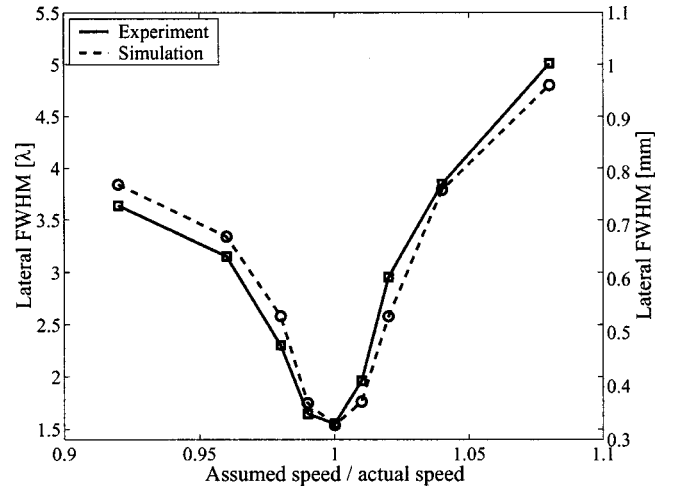


FIG. 1. The r.m.s. full-width, half-maximum (FWHM) beam width of the transducer used plotted as a function of sound speed error in units of mm (right axis) and normalized to the acoustic wavelength (left axis). We compare experimental findings (solid) with matched simulation results (dashed).

Following Ref. 13, the correlation cell size  $S_c$  and our discretized estimate  $\hat{S}_c$  are<sup>13</sup>

$$S_c = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{gg}(\boldsymbol{\delta}) / C_{gg}(0,0) d\delta_x d\delta_z, \quad (4)$$

$$\hat{S}_c = \sum_{m=-D_x}^{m=D_x} \sum_{n=-D_y}^{n=D_y} \hat{C}_{gg}(m,n) / \hat{C}_{gg}(0,0).$$

For the wire target, we calculated the r.m.s. beam plots from the RF system response data. The resulting FWHM beam widths are shown as a function of sound speed error in Fig. 1. The corresponding maximum amplitudes of the envelope-detected data are also shown in Fig. 2, exhibiting a marked decrease in point target echo amplitude with increasing error. We also calculated the axial r.m.s. FWHM of the pulse envelope, and found relatively little change ( $<4\%$  relative to the control) in the axial resolution as a function of sound speed error by this measure.

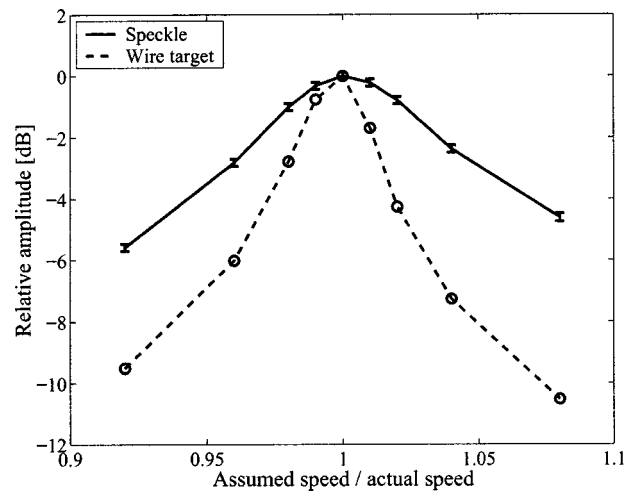


FIG. 2. The mean speckle amplitude from the speckle-generating phantom (solid) and the peak echo amplitude from the wire target (dashed) are plotted as a function of sound speed error. The error bars span  $\pm 2$  standard errors of the mean for 20 independent observations in the phantom.

We also determined the r.m.s. FWHM beam width and correlation cell area  $S_c$  using an acoustic field simulation program.<sup>14</sup> The simulation was matched to the transducer used (28.2-mm aperture, 60% relative bandwidth, 30-mm transmit focus). We set the simulated transducer pulse/echo impulse response to the echo experimentally measured from the wire target, and adjusted the transmit and receive focal delays to correspond to the desired sound speed errors. The simulated receive focus was also fixed in range, but scaled to reflect the effect of dynamic receive focusing for each sound speed error. These simulations sampled the acoustic field at 25- $\mu\text{m}$  spacing laterally and 1.3- $\mu\text{m}$  spacing axially.

To simulate the wire target, we placed a single point scatterer at the wire target range calculated from our experimental data (28.9 mm). The r.m.s. FWHM results of these simulations are compared to the experimental results in Fig. 1.

In simulation, we estimated  $S_c$  from the autocovariance functions of the simulated point spread function, which corresponds to that for speckle for a field of many randomly positioned subresolution scatterers.<sup>15</sup> Our simulations included the attenuation parameters of the phantom, while we assumed the attenuation for the wire target to be negligible. We assumed  $S_c$  to be separable into axial and lateral components:

$$S_c = S_{cx} S_{cz}. \quad (5)$$

Given the two-dimensional RF point spread function  $r$  and its envelope-detected counterpart  $d$ , we expect the speckle intensity lateral correlation cell length  $S_{cx}$  and our discretized estimate  $\hat{S}_{cx}$  to be

$$S_{cx} = \int_{-\infty}^{\infty} |C_{rr}(\delta_x, 0)/C_{rr}(0, 0)|^2 d\delta_x, \quad (6)$$

$$\hat{S}_{cx} = \sum_{m=-D_x}^{m=D_x} |\hat{C}_{rr}(m, 0)/\hat{C}_{rr}(0, 0)|^2,$$

and the axial correlation cell length  $S_{cz}$  and our discretized estimate  $\hat{S}_{cz}$  to be

$$S_{cz} = \int_{-\infty}^{\infty} |C_{dd}(0, \delta_z)/C_{dd}(0, 0)|^2 d\delta_z, \quad (7)$$

$$\hat{S}_{cz} = \sum_{n=-D_z}^{n=D_z} |\hat{C}_{dd}(0, n)/\hat{C}_{dd}(0, 0)|^2.$$

These solutions follow from (11) and (12b) in Ref. 10, and reflect that the expected speckle intensity autocovariance is dependent on the RF PSF laterally and the detected PSF envelope axially. Our average estimates of correlation cell area  $\hat{S}_c$  over 20 ROIs are listed in Table I, normalized to the control.

The r.m.s. FWHM results of these simulations are compared to the experimental results in Fig. 3. In both Fig. 1 and Fig. 3 we consider these simulation results to be in good agreement with our experimental findings. Our simulation method did not use speckle patterns and thus did not require multiple observations to estimate  $S_c$ , hence the absence of error bars for the simulation results. We expect some dis-

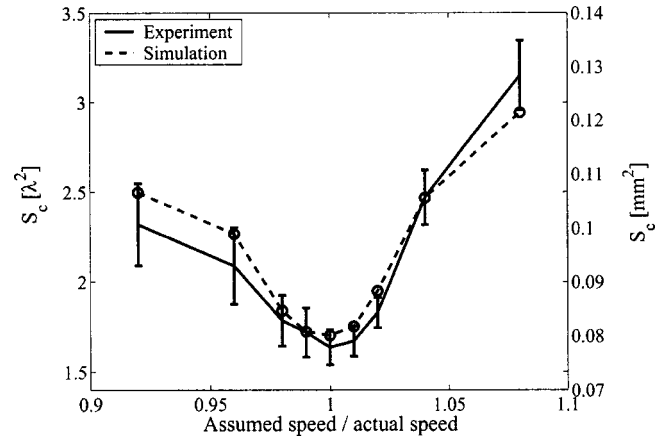


FIG. 3. The speckle intensity correlation cell area plotted as a function of sound speed error in units of  $\text{mm}^2$  (right axis) and normalized to the acoustic wavelength squared (left axis). We compare experimental findings (solid) with matched simulation results (dashed). The experimental measurement error bars span  $\pm 2$  standard errors of the mean for 20 independent observations in the phantom.

crepancies due to the use of fixed receive focus in the simulations versus dynamic receive focusing in the experimental work, electronic noise in the experimental results, variance in the estimates of speckle statistics using experimental data, and inevitable differences between a real ultrasound transducer and its idealized counterpart in simulation.

### III. LOW CONTRAST LESION DETECTABILITY

The detection of low contrast lesions using medical ultrasound is an important clinical task. The performance of both ideal and human observers in lesion detection has been characterized theoretically and experimentally for an idealized statistical description of this task.<sup>16–18</sup> The impact of several imaging parameters on such a performance has also been described.<sup>11,16,19</sup> We based our analysis on the work of Insana and Hall,<sup>18</sup> from which the development below is drawn, and which provides a thorough and accessible description and development of the quantities presented here with appropriate references to the literature. Our primary area of research is breast ultrasound, under which malignant lesions typically show marked hypoechoogenicity.<sup>20</sup> For this reason we limit our examples to negative contrast lesions.

Given a number of assumptions, the detectability of a circular low contrast lesion can be reduced to a function of the first and second order speckle statistics within the lesion and in the surrounding background. Specifically, the performance of an ideal observer in a 2-alternative forced choice (2AFC), signal known exactly (SKE) decision regime is quantified by

$$\text{SNR}_I^2 = 2 \left( \frac{\pi a^2}{4S_c} \right) \text{SNR}_0^2 C^2. \quad (8)$$

Here  $a$  is the lesion radius,  $S_c$  is the speckle intensity correlation cell size (in units of area),  $\text{SNR}_0$  is the speckle point-wise intensity signal-to-noise ratio (the mean over the standard deviation of squared speckle amplitude), and  $C^2$  is the lesion intensity contrast.

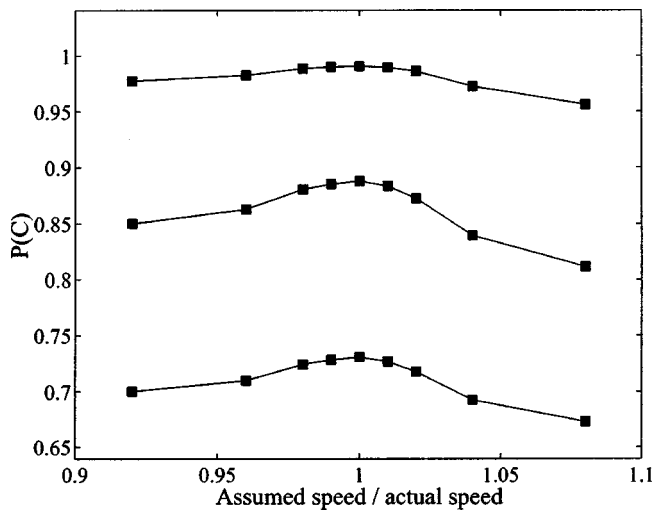


FIG. 4. The probability of a correct decision in detecting a circular, low contrast lesion in speckle noise plotted as a function of lesion contrast and sound speed error. The detection task is a 2 alternative, forced choice for a signal known exactly (2AFC,SKE). The lesion diameter is 3 mm, and the contrasts considered are  $-2$ ,  $-1$ , and  $-0.5$  dB (i.e., negative contrast) from top to bottom.

In a 2AFC, SKE decision, the observer is presented with two images in random order. One image contains the lesion combined with noise, one the noise alone. The observer's task is to determine which of these two images contains the lesion of interest. A third image is also available to the observer, showing a mask indicating the location, size, and shape of the lesion. The assumptions made in deriving and applying (8) are that the noise corresponds to fully-developed speckle, that the target is in the focal region of the transducer, and that the lesion is large relative to  $S_c$ , such that the number of independent speckles within the lesion ( $\pi a^2/4S_c$ ) is greater than ten.

The expected human performance in this detection regime is more intuitively presented as a probability of a correct decision  $P(C)$ , which in an actual human observer trial is estimated by calculating the fraction of correct responses.  $P(C)$  is related to  $\text{SNR}_I$  by

$$P(C) = \Phi\left(\frac{\text{SNR}_H}{2}\right), \quad \text{SNR}_H = \sqrt{\eta_{HI}} \text{SNR}_I, \quad (9)$$

where  $\Phi(x) \equiv 1/\sqrt{2\pi} \int_{-\infty}^x \exp(-t^2/2) dt$ , i.e., the integral of the standard normal distribution, and  $\eta_{HI}$  is the human observer efficiency relative to the ideal observer in this context.  $\eta_{HI}$  has been estimated using human observers, and we used the values reported in Ref. 18 ( $\eta_{HI}=0.59$  for negative-contrast,  $\eta_{HI}=0.58$  for positive-contrast). These results can be related directly to another familiar detection task by noting that  $P(C)$  for the 2AFC equals the area under the receiver operating characteristic (ROC) curve in a yes-no (YN) decision regime, in which the observer must decide whether the target is present in a single image. The ROC curve completely characterizes the observer's performance, and the area under the curve (AUC or  $A_z$ ) summarizes performance in a single value.<sup>21</sup>

Given the results for  $S_c$  shown in Fig. 3, we are thus able to predict the performance of a human observer in this

task as a function of lesion size, contrast, and sound speed error. As an example, the expected  $P(C)$  for 0.5-, 1-, and 2-dB negative-contrast lesions of 3-mm diameter are shown in Fig. 4. These contrast values are in amplitude, which is squared to form the intensity contrast  $C^2$  in (8). Given our calculations of speckle correlation cell size, this lesion contains  $\sim 96$  independent speckles in the control case and  $\sim 50$  independent speckles in the worst case with sound speed error. Equation (9) predicts that for relatively high contrast, ( $>2$  dB) the probability of a correct decision is high, approaching its upper limit of 1. For the very low contrast cases ( $\ll 0.5$  dB),  $P(C)$  approaches its minimum limit of 0.5, indicating that the observer in such cases performs no better than if the decision were based on chance alone (e.g., a coin toss). At either of these extremes, sound speed error has negligible impact on  $P(C)$ . The cases of interest lie at intermediate contrasts, for which subtle changes in sound speed, and by extension in  $S_c$ , cause the greatest changes in  $P(C)$ . For this size lesion, we see for the  $-1$ -dB case that a  $+4\%$  error in sound speed results in a 5% reduction in  $P(C)$ .

In comparing Figs. 1 and 3, we observe that sound speed errors have a larger impact on the lateral beam width than on the correlation cell size. Given that we found small ( $<4\%$ ) changes in axial resolution as a function of error, one would expect the correlation cell size to scale proportionately with the lateral beam width, which is not the case. We believe that this is the result of the curvature of the PSF caused by sound speed error. This curvature is illustrated in Fig. 5, which compares the PSFs measured from the wire target with and without a  $+4\%$  error. We believe this curvature leads to a shorter lateral speckle autocovariance length than would otherwise be expected from the lateral beam width. Note that to produce these figures we interpolated the measured PSFs by a factor of four in each dimension.

The broadening of the beam due to sound speed error illustrated in Figs. 1 and 5 will reduce the scanner's contrast resolution. This long range structure of the PSF will cause lateral smearing across the edges of a lesion, decreasing the contrast at its borders. This effect is illustrated in Fig. 6, in which the images of a 7-mm diameter  $+12$ -dB phantom contrast lesion are shown with and without a  $+4\%$  error, with all other scanner settings identical. Lateral smearing contributes to a marked subjective decrease in visibility in this example. The lesion detection task model applied in our study assumes that the area of the lesion is significantly greater than that of the correlation cell. Border smearing is thus implicitly considered negligible, which is a safe assumption for large lesions. We would expect this model to underestimate the impact of sound speed errors on the detectability of small lesions. [*N.B.*: We imaged the contrast lesion in Fig. 6 in a commercial phantom (Model 404GS, Radiation Measurements, Inc., Middleton, WI, 1540-m/s, 0.5-dB/cm/MHz attenuation) using a 35-mm transmit focus, with an experimental setup identical to that described for the phantom in Sec. I. We did not use this phantom for measurements because it did not feature adequate scatterer density in the background region to provide fully developed speckle under these conditions. At a 35-mm range the background scattering in this phantom exhibited a speckle amplitude SNR of only 1.4.]

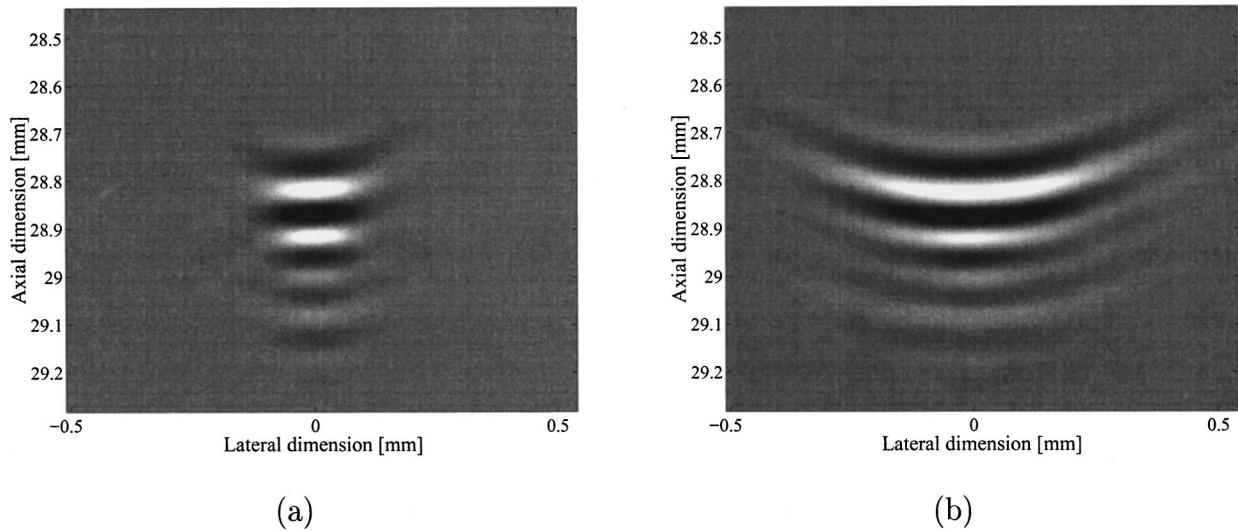


FIG. 5. Point spread functions (PSFs) experimentally measured with a wire target at the correct sound speed (a) and with a +4% sound speed error (overestimation) (b) showing the resulting increase in lateral beam width and PSF curvature. These results are shown with separate gray scales, and thus do not reveal the peak echo amplitude reduction that also results from this error. The transducer and experimental setup used to create these images are described in Sec. II.

#### IV. POINT TARGET DETECTABILITY

The authors have previously described a statistical approach for assessing the impact of imaging parameters on the visualization of a point-like target at the focus among diffuse scatterers in the context of the visualization of breast microcalcifications *in vivo*.<sup>22</sup> This approach considers the decision of an ideal observer based on applying a brightness threshold to the pointwise amplitude at the target location. We assume the diffuse scatterers to be of sufficient concentration to produce fully-developed speckle.<sup>23</sup> We do not consider the more complicated task of detecting such a target at some location, known or unknown, within a region of interest.

The pointwise amplitudes of envelope-detected speckle from a resolution cell containing diffuse scatterers with and without a strong scatterer exhibit Rician and Rayleigh probability density functions, which correspond in this context to the expected amplitude distributions for the target present ( $H_1$ ) and target absent ( $H_0$ ) cases:

$$p(a|H_1) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2 + s^2}{2\sigma^2}\right) I_0\left(\frac{as}{\sigma^2}\right), \quad (10)$$

$$p(a|H_0) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2}{2\sigma^2}\right).$$

Here  $a$  denotes detected speckle amplitude,  $s$  is the coherent echo amplitude,  $\sigma$  is the standard deviation of the diffuse component's (RF speckle) echo amplitude, and  $I_0$  is the modified Bessel function of the first kind, zero order.<sup>24</sup> The relative strength of the coherent component is parametrized by the variable  $k$ , defined  $k \equiv s/\sigma$ .

The optimum detector for a binary hypothesis task calculates a quantity known as the likelihood ratio, and makes a decision by applying a threshold to this quantity.<sup>25,26</sup> The likelihood ratio  $\lambda(a)$  for this detection task is defined:

$$\lambda(a) = \frac{p(a|H_1)}{p(a|H_0)}. \quad (11)$$

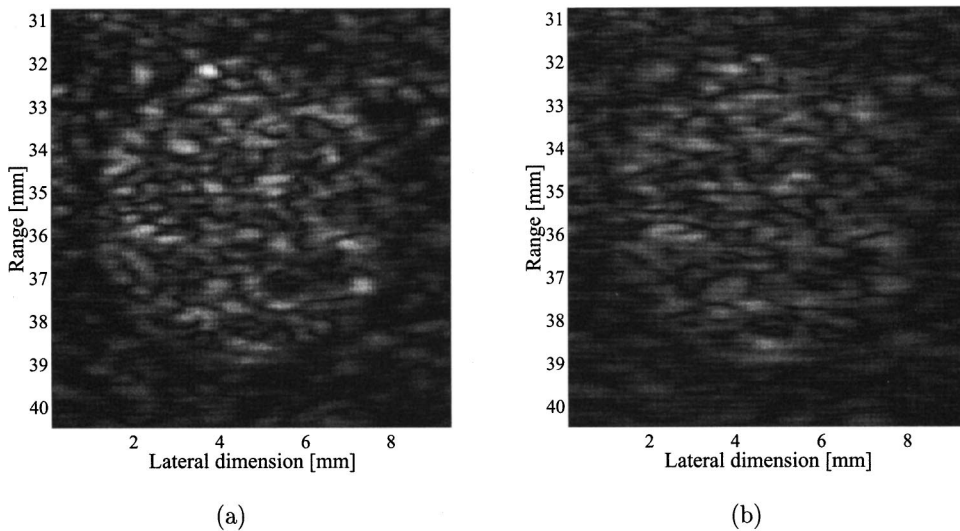


FIG. 6. A 12-dB positive contrast lesion in a tissue-mimicking phantom is shown (a) at the correct sound speed and (b) with a +4% sound speed error (overestimation) showing the resulting decrease in contrast resolution and mean speckle amplitude. These results are shown with the same gray scale and with otherwise identical system settings. The transducer used to create these images is described in Sec. I, while the phantom is described in Sec. IV.

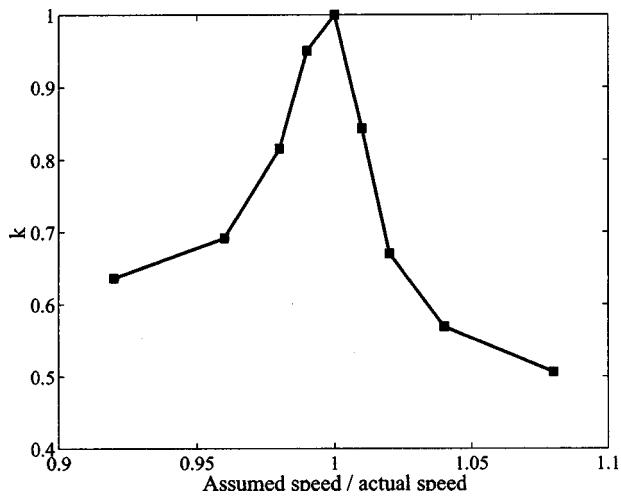


FIG. 7. The effective Rician  $k$  parameter plotted as a function of sound speed error, normalized to the control. The reduction in this quantity corresponds to the reduction in echo amplitude of a point target relative to the surrounding speckle background.

Thus the likelihood ratio detector for this task merely compares  $\lambda(a)$  to an amplitude (or “brightness”) threshold to make the decision. The practical implication of this is that for an ultrasound B-mode image of a subresolution coherent target surrounded by diffuse scatterers, for a decision made at a single image point on the basis of amplitude alone, the optimum detector is merely a brightness threshold.

The integrals of  $p(a|H_0)$  and  $p(a|H_1)$  above the brightness threshold applied to make the decision correspond to the probabilities of false alarm ( $P_F$ ) and of correct detection ( $P_D$ ), also known as the false positive and true positive fractions, FPF and TPF. The plot of  $P_D$  vs  $P_F$  parametrized by the amplitude threshold is the ROC curve for an ideal observer making the decision.

We derive the impact of sound speed errors on the  $k$  parameter from the experimental results shown in Fig. 2. The ratio of wire target peak echo amplitude to mean speckle

brightness, normalized to the control case, predicts the scaling of the  $k$  parameter as a function of sound speed error. Given that the wire peak echo amplitude is reduced more than the speckle brightness for all errors, this scaling represents a degradation in point target detectability in these cases. The expected scaling of  $k$  is shown in Fig. 7. For example, for a sound speed error of +4%,  $k$  is scaled by approximately 0.6. This suggests that this error reduces the brightness of this target relative to background by over 40%. Such errors essentially “push” point targets down into the background.

This effect is demonstrated in Fig. 8. In this example, we created simulated ultrasound images by convolving the interpolated PSFs shown in Fig. 5 with appropriate scattering functions and envelope-detecting the results. We represented the tissue’s diffuse scattering function with a field of normally-distributed random numbers with zero mean, and the simulated point scatterer with a single positive value within a field of zeros. For the control case and for a +4% sound speed error, we convolved the corresponding PSF with the same diffuse scattering and point scatterer fields and summed these results to create simulated RF data. We envelope-detected these results using the Hilbert transform to create the images shown in Fig. 8.<sup>8</sup> In this example we set the point scatterer amplitude to provide a  $k$  value of 5. We increased the brightness of the +4% error image to equalize the mean speckle brightness, allowing the reader to compare the contrast of the point scatterer in the two cases. This simulation method provided effective diffuse scatterer densities of over 330 per resolution cell in the control and over 860 per resolution cell in the +4% error case, far in excess of the scatterers required to provide fully-developed speckle.<sup>23</sup> The corresponding speckle signal-to-noise ratios for the realizations shown are 1.97 and 1.91, in good agreement with the expected value of 1.91.<sup>9</sup>

With knowledge of the scaling of  $k$ , we can integrate the PDFs in (10) with and without this scaling to predict the area

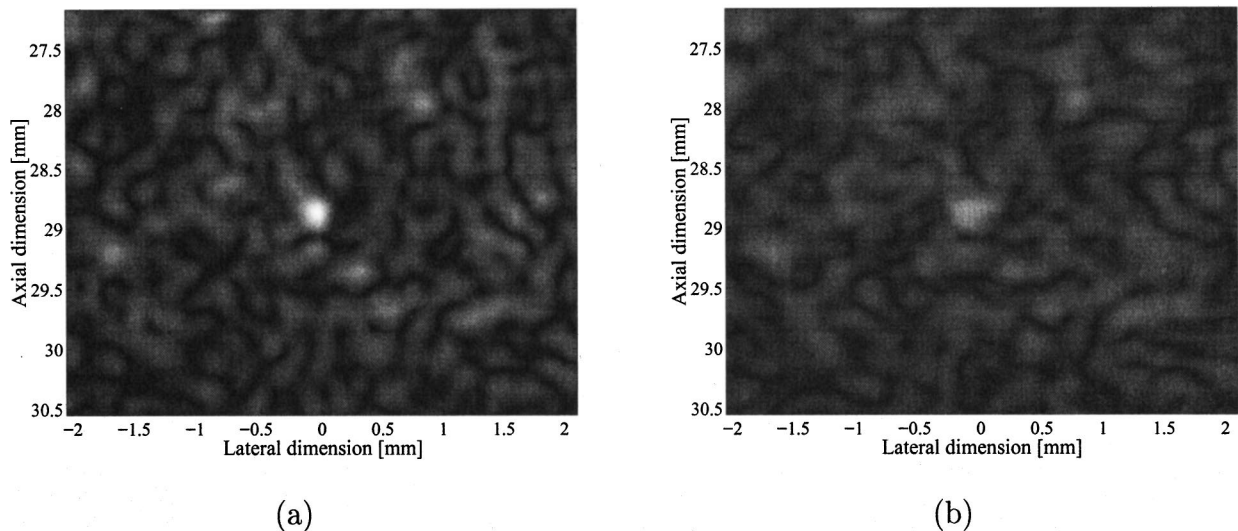


FIG. 8. A simulated speckle field with a bright point scatterer at the center having a Rician  $k$  parameter value of 5 is shown (a) at the correct sound speed and (b) with a +4% sound speed error (overestimation). The overall brightness in image (b) has been increased to equalize the mean speckle brightness in both images, while the grayscale gain is the same for both images. The sound speed error produces a marked decrease in the contrast of the point target, while the texture of the speckle does not appear significantly changed. The method used to create these images is described in Sec. IV.

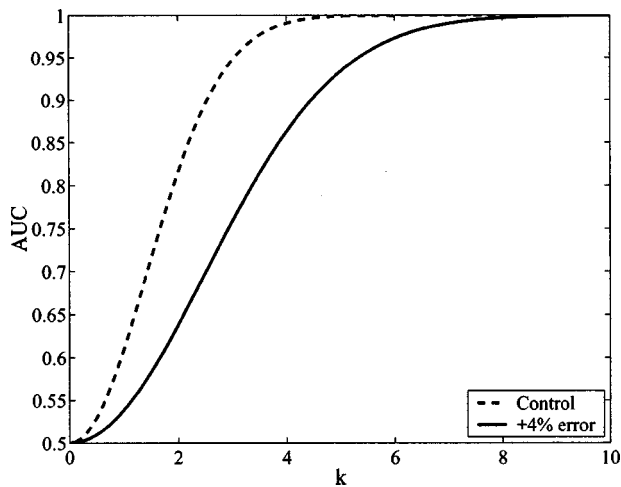


FIG. 9. The area under the ROC curve for an ideal observer deciding whether a point target is present at the imaging system focus, plotted as function of the Rician  $k$  parameter. The control case (dashed) is compared to a case with a +4% sound speed error (solid).

under the ROC curve (AUC) as a function of  $k$  with and without a sound speed error present. An example of this analysis for a +4% sound speed error is shown in Fig. 9. This result shows the expected reduction in ideal observer performance as a function of  $k$ . For a target having a  $k$  value of 3, the AUC in the control case (i.e., no sound speed error) of 0.95 is reduced to 0.76 by the sound speed error. The largest impact in this example is at  $k=2.6$ , where AUC is reduced by 0.20. This is a relative reduction of 22% in correct decisions in cases involving targets of this relative brightness. This is a considerable reduction for what we consider to be an error of relatively routine magnitude.

## V. SUMMARY DISCUSSION

Table I is a comprehensive list of the results shown in Figs. 1, 2, 3, and 7. For the transducer and targets used, we found that sound speed errors up to  $\pm 8\%$  can degrade the system lateral spatial resolution by up to a factor of three (Fig. 1). The impact on amplitude sensitivity was greater for the point target, which exhibited peak echo amplitude reductions up to  $-10.5$  dB (Fig. 2). The speckle intensity correlation cell area increased by up to 92% (Fig. 3).

For a typical sound speed error, Fig. 8 shows a marked decrease in point target contrast, while the speckle texture does not appear significantly changed. This has important implications for studies of sound speed error *in vivo*: it may be possible to alter the beamforming sound speed variable without observing remarkable changes in speckle-generating tissue texture, particularly when logarithmic brightness compression is being applied, which decreases the apparent magnitude of the amplitude sensitivity reduction associated with sound speed errors. We expect that rigorous subjective evaluation of sound speed error correction techniques in the clinic will require careful selection of regions and targets of interest to reliably demonstrate the impact of such errors and of mechanisms by which they may be corrected.

The beamforming distortions arising of sound speed errors are inversely proportional to the imaging system

$f$ -number and wavelength, thus affecting most prominently clinical applications in, e.g., the breast, small parts, and peripheral vasculature. These errors are compounded in systems that are dynamically focused on receive. Such systems consequently apply not only incorrect focal delays, but delays calculated based on erroneous focal ranges as well. A number of techniques for measuring the sound speed of an unknown medium with pulse-echo ultrasound have been proposed and recently reviewed.<sup>27</sup> The authors have recently demonstrated an additional method for such measurement.<sup>28</sup> A compensation technique for the sound speed errors that arise in water-coupled sonography has also been demonstrated.<sup>29</sup> As discussed in Ref. 28, sound speed errors are also a subset of imaging distortions arising from phase aberration, for which a variety of corrective techniques have been proposed and experimentally evaluated.<sup>6,30–36</sup> We expect techniques that compensate for phase aberrations to reduce sound speed errors as well.

The spreading of the pulse-echo system response that results from sound speed errors can be considered a spatial low-pass filter. In this light, it is easy to understand how such errors have a much greater impact on the visualization of image structures comprising higher spatial frequencies. This can be seen by comparing the results in Figs. 4 and 9, which show a much greater change in point target detectability than low contrast lesion detectability for the examples considered. While in this paper we consider the case of point target detectability, we expect this low-pass filtering to cause similar observer performance degradation in any image task relying on high spatial and/or contrast resolution, such as the visualization of spiculation or capsules surrounding lesions, the differentiation of lesions and cysts, ultrasound-guided biopsy, and the estimation and visualization of blood flow velocity in small vessels.

We consider the results shown in this paper to demonstrate the significant impact of routinely-encountered sound speed errors on important clinical imaging tasks. Given the good agreement with simulation demonstrated, it should also be possible to extend these analysts in simulation to other imaging systems, including hypothetical systems.

## ACKNOWLEDGMENTS

In-kind and technical support was provided by the Ultrasound Group of Siemens Medical Systems, Incorporated, and by the Intel Corporation. The authors thank Dr. Timothy Hall for his assistance with the phantom and Dr. Robert Wagner and Dr. Stephen Smith for helpful discussion. This work was supported by NIH grants No. R01-CA43334 and No. R01-CA76059.

<sup>1</sup>P. N. T. Wells, *Biomedical Ultrasonics* (Academic, London, 1977), Chap. 4, pp. 120–125.

<sup>2</sup>S. A. Goss, R. L. Johnston, and F. Dunn, "Comprehensive compilation of empirical ultrasonic properties of mammalian tissues," *J. Acoust. Soc. Am.* **64**, 423–457 (1978).

<sup>3</sup>S. A. Goss, R. L. Johnston, and F. Dunn, "Compilation of empirical properties of mammalian tissues. II," *J. Acoust. Soc. Am.* **68**, 93–108 (1980).

<sup>4</sup>N. Bilaniuk and G. S. K. Wong, "Speed of sound in pure water as a function of temperature," *J. Acoust. Soc. Am.* **93**, 1609–1612 (1993).

<sup>5</sup>C. M. Sehgal, B. R. Porter, and J. F. Greenleaf, "Ultrasonic nonlinear

- parameters and sound speed of alcohol-water mixtures," J. Acoust. Soc. Am. **79**, 566–570 (1986).
- <sup>6</sup> S. W. Flax and M. O'Donnell, "Phase aberration correction using signals from point reflectors and diffuse scatterers: basic principles," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **35**, 758–767 (1988).
- <sup>7</sup> J. Y. Lu, H. Zou, and J. F. Greenleaf, "Biomedical ultrasound beam forming," Ultrasound Med. Biol. **20**, 403–428 (1994).
- <sup>8</sup> R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill, New York, 1986), Chap. 12, pp. 267–271.
- <sup>9</sup> R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," IEEE Trans. Sonics Ultrason. **SU-30**, 156–163 (1983).
- <sup>10</sup> R. F. Wagner, M. F. Insana, and S. W. Smith, "Fundamental correlation lengths of coherent speckle in medical ultrasonic images," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **35**, 34–44 (1988).
- <sup>11</sup> S. W. Smith, G. E. Trahey, S. M. Hubbard, and R. F. Wagner, "Properties of acoustic speckle in the presence of phase aberration, part II: correlation lengths," Ultrason. Imaging **10**, 29–51 (1988).
- <sup>12</sup> C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Trans. Acoust., Speech, Signal Process. **24**, 320–327 (1976).
- <sup>13</sup> J. W. Goodman, "Statistical properties of laser speckle patterns," in *Laser Speckle and Related Phenomena*, edited by J. C. Dainty (Springer-Verlag, Berlin, 1984), pp. 36–48.
- <sup>14</sup> J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 262–267 (1992).
- <sup>15</sup> R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," IEEE Trans. Sonics Ultrason. **30**, 156–163 (1983).
- <sup>16</sup> S. W. Smith, R. F. Wagner, J. M. Sandrik, and H. Lopez, "Low contrast detectability and contrast/detail analysis in medical ultrasound," IEEE Trans. Sonics Ultrason. **30**, 164–173 (1983).
- <sup>17</sup> T. J. Hall, M. F. Insana, N. M. Soller, and L. A. Harrison, "Ultrasound contrast-detail analysis: a preliminary study in human observer performance," Med. Phys. **20**, 117–127 (1993).
- <sup>18</sup> M. F. Insana and T. J. Hall, "Visual detection efficiency in ultrasonic imaging: a framework for objective assessment of image quality," J. Acoust. Soc. Am. **95**, 2081–2090 (1994).
- <sup>19</sup> G. E. Trahey, J. W. Allison, S. W. Smith, and O. T. von Ramm, "A quantitative approach to speckle reduction via frequency compounding," Ultrason. Imaging **8**, 151–164 (1986).
- <sup>20</sup> A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: use of sonography to distinguish between benign and malignant lesions," Radiology **196**, 124–134 (1995).
- <sup>21</sup> C. E. Metz, "ROC methodology in radiologic imaging," Invest. Radiol. **21**, 720–733 (1986).
- <sup>22</sup> M. E. Anderson, M. S. Soo, R. C. Bentley, and G. E. Trahey, "The detection of breast microcalcifications with medical ultrasound," J. Acoust. Soc. Am. **101**, 29–39 (1997).
- <sup>23</sup> J. M. Thijssen and B. J. Oosterveld, "Texture in tissue echograms," J. Ultrasound Med. **9**, 215–229 (1990).
- <sup>24</sup> L. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I* (Wiley, New York, 1968), Chap. 2, pp. 23–37.
- <sup>25</sup> L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis* (Addison-Wesley, New York, 1968), Chap. 4, pp. 105–109.
- <sup>26</sup> J. W. Goodman, *Statistical Optics* (Wiley-Interscience, New York, 1985), Chap. 2, pp. 44–55.
- <sup>27</sup> D. E. Robinson, J. Ophir, L. S. Wilson, and C. F. Chen, "Pulse-echo ultrasound speed measurements: progress and prospects," Ultrasound Med. Biol. **17**, 633–646 (1991).
- <sup>28</sup> M. E. Anderson and G. E. Trahey, "The direct estimation of sound speed using pulse-echo ultrasound," J. Acoust. Soc. Am. **104**, 3099–3106 (1998).
- <sup>29</sup> J. Jellins and G. Kossoff, "Velocity compensation in water-coupled breast echography," Ultrasonics **11**, 223–226 (1973).
- <sup>30</sup> M. Hirama, O. Ikeda, and T. Sato, "Adaptive ultrasonic array imaging system through an inhomogeneous layer," J. Acoust. Soc. Am. **71**, 100–109 (1982).
- <sup>31</sup> L. Nock, G. E. Trahey, and S. W. Smith, "Phase aberration correction in medical ultrasound using speckle brightness as a quality factor," J. Acoust. Soc. Am. **85**, 1819–1833 (1989).
- <sup>32</sup> D. Rachlin, "Direct estimation of aberrating delays in pulse-echo imaging systems," J. Acoust. Soc. Am. **88**, 191–198 (1990).
- <sup>33</sup> R. Kanda, Y. Sumino, K. Takamizawa, and H. Sasaki, "An investigation of wavefront distortion correction: correction using averaged phase information and the effect of correction in one and two dimensions," Proc.-IEEE Ultrason. Symp. **2**, 1201–1206 (1991).
- <sup>34</sup> D. L. Liu and R. C. Waag, "Time-shift compensation of ultrasonic pulse focus degradation using least mean square error estimates of arrival time," J. Acoust. Soc. Am. **95**, 542–555 (1994).
- <sup>35</sup> G. C. Ng, P. D. Freiburger, W. F. Walker, and G. E. Trahey, "A speckle target adaptive imaging technique in the presence of distributed aberrations," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **44**, 140–151 (1997).
- <sup>36</sup> S. Krishnan, K. W. Rigby, and M. O'Donnell, "Adaptive aberration correction of abdominal images using PARCA," Ultrason. Imaging **19**, 169–179 (1997).

# 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.

## Displacement in a continuously twisted structurally chiral medium due to axial loading

Akhlesh Lakhtakia<sup>a)</sup> and Joseph A. Sherwin

CATMAS—Computational and Theoretical Materials Sciences Group, Department of Engineering Science and Mechanics, Pennsylvania State University, University Park, Pennsylvania 16802-6812

(Received 4 January 2000; revised 3 March 2000; accepted 5 March 2000)

Loading of a linear, continuously twisted, structurally chiral medium (CTSCM) parallel to its helicoidal axis is analyzed, after assuming that the medium is locally triclinic. The roles of certain composite compliances are established, with a view to eventual use of CTSCMs for ultrasonic, infrared and optical applications. © 2000 Acoustical Society of America.

[S0001-4966(00)02806-X]

PACS numbers: 43.20.Bi, 43.38.Fx [CBB]

### INTRODUCTION

A few years ago, a unidirectionally and periodically nonhomogeneous substance—the continuously twisted, structurally chiral medium (CTSCM)—was introduced in this journal.<sup>1</sup> A CTSCM has a helicoidal structure, and pile-of-plates realizations of the analogous cholesteric liquid crystals have been traced back to 1869.<sup>2</sup> More recently, physical vapor deposition techniques have been implemented to grow CTSCMs as thin films.<sup>3</sup> Viable optical applications have emerged since then,<sup>4</sup> and are now being demonstrated experimentally.<sup>5,6</sup> Although ultrasonic applications have been theoretically examined as well,<sup>7,8</sup> experimental confirmation has not yet become possible. This is because the fabrication of CTSCMs with large periods ( $>20 \mu\text{m}$ ) is still not reliable,<sup>3</sup> although research continues.<sup>9</sup> CTSCMs with large periods are particularly attractive because they offer the double advantage of being useful for ultrasonic as well as infrared applications.

Whether the applications of CTSCMs are ultrasonic, infrared or even optical, the fabricated materials must be mechanically robust for two reasons:

- (i) A CTSCM must be able to withstand indentation by scanning electron microscope (SEM) probes, etc., necessary for microstructural analyses as well as for quality control.
- (ii) CTSCM-based devices will have to withstand stresses created by their operating environments.

Although indentation techniques<sup>10</sup> may appear feasible to study the mechanical responses of CTSCMs, the only available results<sup>11</sup> display dramatic differences between experimental data and back-of-the-envelope formulas. Clearly, a basic theory must be developed from which CTSCM characterization researchers can take pointers to devise suitable experimental techniques. This communication provides a starting point.

### I. CONSTITUTIVE RELATIONS

Assuming linear elastic behavior, we write Hooke's law for a CTSCM in the Kelvin notation<sup>1,12</sup> as follows:

$$\begin{aligned} [\underline{\mathbf{S}}(\mathbf{r})] &= [\underline{\mathbf{s}}(z)] [\underline{\mathbf{T}}(\mathbf{r})] \\ &= [\underline{\mathbf{R}}(-z)]^t [\underline{\mathbf{s}}'] [\underline{\mathbf{R}}(-z)] [\underline{\mathbf{T}}(\mathbf{r})]. \end{aligned} \quad (1)$$

Here and hereafter,  $\mathbf{r}=[x; y; z]$  is the position vector; the superscript  $t$  denotes the transpose, while the column vectors

$$[\underline{\mathbf{S}}] = [S_{xx}; S_{yy}; S_{zz}; 2S_{yz}; 2S_{xz}; 2S_{xy}]^t, \quad (2)$$

$$[\underline{\mathbf{T}}] = [T_{xx}; T_{yy}; T_{zz}; T_{yz}; T_{xz}; T_{xy}]^t, \quad (3)$$

$$[\underline{\mathbf{u}}] = [u_x; u_y; u_z]^t, \quad (4)$$

respectively, contain components of the strain tensor, the stress tensor and the displacement vector. The CTSCM is helicoidally nonhomogeneous along the  $z$  axis, also called the *helicoidal axis*. This feature is captured by the  $6 \times 6$  Bond matrix

<sup>a)</sup>Electronic mail: AXL4@psu.edu



$$[\mathbf{R}(z)] = \begin{bmatrix} \cos^2 \pi z/\Omega & \sin^2 \pi z/\Omega & 0 & 0 & 0 & -\sin 2\pi z/\Omega \\ \sin^2 \pi z/\Omega & \cos^2 \pi z/\Omega & 0 & 0 & 0 & \sin 2\pi z/\Omega \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos \pi z/\Omega & \sin \pi z/\Omega & 0 \\ 0 & 0 & 0 & -\sin \pi z/\Omega & \cos \pi z/\Omega & 0 \\ 0.5 \sin 2\pi z/\Omega & -0.5 \sin 2\pi z/\Omega & 0 & 0 & 0 & \cos 2\pi z/\Omega \end{bmatrix}, \quad (5)$$

where  $2\Omega$  is the structural period. [Equation (5) holds for a structurally right-handed CTSCM. For a left-handed CTSCM, replace  $\Omega$  by  $-\Omega$  in that equation as well as the later ones.] The  $6 \times 6$  matrix  $[\mathbf{s}'] = [\mathbf{s}(0)] = \{s_{mn}\}$ , ( $1 \leq m \leq 6, 1 \leq n \leq 6$ ), is called the *reference compliance matrix*.<sup>3</sup> We assume that  $[\mathbf{s}']$  is symmetric, and that the CTSCM possesses a locally triclinic structure, which is very general. Our objective is to determine the relevant stress and displacement components when a static axial load is applied on a CTSCM, i.e., when  $\partial/\partial x = 0$  and  $\partial/\partial y = 0$ .

## II. AXIAL STATIC ANALYSIS

Following Ovaert<sup>10</sup> for thin solid films, we concentrate on axial elastostatic analysis here. Accordingly, the linearized equation of motion reads as

$$[\mathbf{A}] \frac{d}{dz} [\mathbf{T}(z)] = [\mathbf{0}], \quad (6)$$

where the  $3 \times 6$  matrix

$$[\mathbf{A}] = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \quad (7)$$

Clearly, the stress components

$$T_{xz}(z) = T_{xz0}, \quad T_{yz}(z) = T_{yz0}, \quad T_{zz}(z) = T_{zz0} \quad (8)$$

must therefore be constant, being determined by the applied axial load.

The determination of the remaining components of the stress tensor as well as the displacement components requires us to consider the matrix equation

$$[\mathbf{A}]^t \frac{d}{dz} [\mathbf{u}(z)] = [\mathbf{S}(z)]; \quad (9)$$

accordingly, the in-plane strains  $S_{xx}$ ,  $S_{yy}$  and  $S_{xy}$  are identically null-valued. Using (1) on the right side of (9), we obtain

$$[\mathbf{A}]^t \frac{d}{dz} [\mathbf{u}(z)] = [\mathbf{R}(-z)]^t [\mathbf{s}'] [\mathbf{R}(-z)] [\mathbf{T}(z)], \quad (10)$$

which matrix equation comprises three differential equations and three algebraic equations. The solution of these equations is greatly assisted by adopting the transformations

$$[\mathbf{T}(z)] = [\mathbf{R}(z)] [\mathbf{T}'(z)], \quad [\mathbf{u}(z)] = [\mathbf{A}(z)] [\mathbf{u}'(z)], \quad (11)$$

which utilize also the rotation matrix

$$[\mathbf{A}(z)] = \begin{bmatrix} \cos \pi z/\Omega & -\sin \pi z/\Omega & 0 \\ \sin \pi z/\Omega & \cos \pi z/\Omega & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (12)$$

that corresponds to the Bond matrix  $[\mathbf{R}(z)]$ . Then (10) simplifies to

$$[\mathbf{A}]^t \frac{d}{dz} [\mathbf{u}'(z)] = -\frac{\pi}{\Omega} [\mathbf{\Gamma}]^t [\mathbf{u}'(z)] + [\mathbf{s}'] [\mathbf{T}'(z)], \quad (13)$$

wherein the  $3 \times 6$  matrix

$$[\mathbf{\Gamma}] = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (14)$$

has been employed for notational compactness.

The solution of the three algebraic equations yields the remaining stress components as follows:

$$\begin{bmatrix} T_{xx}(z) \\ T_{yy}(z) \\ T_{xy}(z) \end{bmatrix} = - \begin{bmatrix} \cos^2 \pi z/\Omega & \sin^2 \pi z/\Omega & -\sin 2\pi z/\Omega \\ \sin^2 \pi z/\Omega & \cos^2 \pi z/\Omega & \sin 2\pi z/\Omega \\ 0.5 \sin 2\pi z/\Omega & -0.5 \sin 2\pi z/\Omega & \cos 2\pi z/\Omega \end{bmatrix} \times \begin{bmatrix} s_{11} & s_{12} & s_{16} \\ s_{12} & s_{22} & s_{26} \\ s_{16} & s_{26} & s_{66} \end{bmatrix}^{-1} \begin{bmatrix} s_{13} & s_{14} & s_{15} \\ s_{23} & s_{24} & s_{25} \\ s_{36} & s_{46} & s_{56} \end{bmatrix} \times \begin{bmatrix} 0 & 0 & 1 \\ -\sin \pi z/\Omega & \cos \pi z/\Omega & 0 \\ \cos \pi z/\Omega & \sin \pi z/\Omega & 0 \end{bmatrix} \begin{bmatrix} T_{xz0} \\ T_{yz0} \\ T_{zz0} \end{bmatrix}. \quad (15)$$

In general,  $T_{xx}$ ,  $T_{yy}$  and  $T_{xy}$  are not constant with respect to the coordinate  $z$ —unlike  $T_{xz}$ ,  $T_{yz}$  and  $T_{zz}$ .

The three differential equations in (13) may be compactly written as

$$\frac{d}{dz} \begin{bmatrix} u'_z(z) \\ u'_y(z) \\ u'_x(z) \end{bmatrix} = \frac{\pi}{\Omega} \begin{bmatrix} 0 \\ -u'_x(z) \\ u'_y(z) \end{bmatrix} + \begin{bmatrix} \tilde{s}_{33} & \tilde{s}_{34} & \tilde{s}_{35} \\ \tilde{s}_{34} & \tilde{s}_{44} & \tilde{s}_{45} \\ \tilde{s}_{35} & \tilde{s}_{45} & \tilde{s}_{55} \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 \\ -\sin \pi z/\Omega & \cos \pi z/\Omega & 0 \\ \cos \pi z/\Omega & \sin \pi z/\Omega & 0 \end{bmatrix} \begin{bmatrix} T_{xz0} \\ T_{yz0} \\ T_{zz0} \end{bmatrix}, \quad (16)$$

$$\begin{bmatrix} \tilde{s}_{33} & \tilde{s}_{34} & \tilde{s}_{35} \\ \tilde{s}_{34} & \tilde{s}_{44} & \tilde{s}_{45} \\ \tilde{s}_{35} & \tilde{s}_{45} & \tilde{s}_{55} \end{bmatrix} = \begin{bmatrix} s_{33} & s_{34} & s_{35} \\ s_{34} & s_{44} & s_{45} \\ s_{35} & s_{45} & s_{55} \end{bmatrix} - \begin{bmatrix} s_{13} & s_{23} & s_{36} \\ s_{14} & s_{24} & s_{46} \\ s_{15} & s_{25} & s_{56} \end{bmatrix} \cdot \begin{bmatrix} s_{11} & s_{12} & s_{16} \\ s_{12} & s_{22} & s_{26} \\ s_{16} & s_{26} & s_{66} \end{bmatrix}^{-1} \begin{bmatrix} s_{13} & s_{14} & s_{15} \\ s_{23} & s_{24} & s_{25} \\ s_{36} & s_{46} & s_{56} \end{bmatrix} \quad (17)$$

being a symmetric matrix.

Equation (16) can be solved easily using standard methods. Assuming that  $\pi/\Omega \neq 0$ , we obtained a straightforward expression for the displacement components in matrix notation, viz.,

with

$$\begin{bmatrix} u_x(z) \\ u_y(z) \\ u_z(z) \end{bmatrix} = \begin{bmatrix} u_{x0} \\ u_{y0} \\ u_{z0} \end{bmatrix} + \begin{bmatrix} \frac{\tilde{s}_{44} + \tilde{s}_{55}}{2} z + \phi_a(z) & \phi_b(z) & \phi_c(z) \\ \phi_b(z) & \frac{\tilde{s}_{44} + \tilde{s}_{55}}{2} z - \phi_a(z) & \phi_d(z) \\ \phi_c(z) & \phi_d(z) & \tilde{s}_{33} z \end{bmatrix} \begin{bmatrix} T_{xz0} \\ T_{yz0} \\ T_{zz0} \end{bmatrix}. \quad (18)$$

The first column vector on the right side of this equation contains the boundary values, i.e.,  $u_{x0} = u_x(0)$ , etc. The four functions used in the matrix on the right side of (18) are defined as follows:

$$\phi_a(z) = -\frac{2\tilde{s}_{45}(1 - \cos 2\pi z/\Omega) + (\tilde{s}_{44} - \tilde{s}_{55})\sin 2\pi z/\Omega}{4\pi/\Omega}, \quad (19)$$

$$\phi_b(z) = \frac{2\tilde{s}_{45}\sin 2\pi z/\Omega - (\tilde{s}_{44} - \tilde{s}_{55})(1 - \cos 2\pi z/\Omega)}{4\pi/\Omega}, \quad (20)$$

$$\phi_c(z) = -\frac{\tilde{s}_{34}(1 - \cos \pi z/\Omega) - \tilde{s}_{35}\sin \pi z/\Omega}{\pi/\Omega}, \quad (21)$$

$$\phi_d(z) = \frac{\tilde{s}_{34}\sin \pi z/\Omega + \tilde{s}_{35}(1 - \cos \pi z/\Omega)}{\pi/\Omega}. \quad (22)$$

According to (18), the mechanical response of a CTSCM to an axially applied load (whether virtually static or slowly varying in time) is determined by six quantities— $\tilde{s}_{33}$ ,  $\tilde{s}_{34}$ ,  $\tilde{s}_{35}$ ,  $\tilde{s}_{44}$ ,  $\tilde{s}_{45}$  and  $\tilde{s}_{55}$ —which may be called *composite compliances*. Of these,  $\tilde{s}_{33}$  may be the most important one for devices, and it is also the one that can be most easily measured. The remaining five composite compliances may be determined by loading CTSCM slabs of thicknesses that are integer multiples of  $\Omega/2$  and then measuring displacements across the slab faces. Obviously, the analysis presented here does not consider the finite lateral dimensions of a real sample, which would require numerically intensive research. However, the presented work has identified the

relative roles of several components of the compliance matrix. Based on optical experiments,<sup>4,6,9</sup> we conclude that the actual local symmetries of CTSCMs are likely to be much simpler than (and therefore special cases of) the triclinic symmetry that is covered in this communication.

<sup>1</sup>A. Lakhtakia, "Elastodynamic wave propagation in a continuously twisted structurally chiral medium along the axis of chirality," *J. Acoust. Soc. Am.* **95**, 597–600 (1994); Erratum, **95**, 3669 (1994).

<sup>2</sup>E. Reusch, "Untersuchung über Glimmercombinationen," *Ann. Phys. (Leipzig)* **138**, 628–638 (1869).

<sup>3</sup>A. Lakhtakia, K. Robbie, and M. J. Brett, "Spectral Green's function for wave excitation and propagation in a piezoelectric, continuously twisted, structural chiral medium," *J. Acoust. Soc. Am.* **101**, 2052–2058 (1997).

<sup>4</sup>R. Messier and A. Lakhtakia, "Sculptured thin films—II. Experiments and applications," *Mater. Res. Innovations* **2**, 217–222 (1999).

<sup>5</sup>I. Hodgkinson, Q. h. Wu, A. Lakhtakia, and R. Messier, "Linear and circular polarization filters using sculptured thin films," *OSA Opt. Photonics News* **10**, 30–31 (December 1999).

<sup>6</sup>Q. H. Wu, I. J. Hodgkinson, and A. Lakhtakia, "Circular polarization filters made of chiral sculptured thin films: Experimental and simulation results," *Opt. Eng.* (in press).

<sup>7</sup>A. Lakhtakia and M. W. Meredith, "Shear axial modes in a PCTSCM. Part IV: Bandstop and notch filters," *Sens. Actuators A: Phys.* **73**, 193–200 (1999).

<sup>8</sup>A. Lakhtakia, "Shear axial modes in a PCTSCM. Part V: Transmission spectral holes," *Sens. Actuators A: Phys.* (in press).

<sup>9</sup>*Engineered Nanostructural Thin Films and Materials*, edited by A. Lakhtakia and R. Messier, SPIE Proc. **3790** (1999).

<sup>10</sup>T. C. Ovaert, "On the indentation of a transversely isotropic half-space with application to thin solid films," *ASME J. Tribol.* **115**, 650–657 (1993).

<sup>11</sup>M. W. Seto, K. Robbie, D. Vick, M. J. Brett, and L. Kuhn, "Mechanical response of thin films with helical microstructures," *J. Vac. Sci. Technol. B* **17**, 2172–2177 (1999).

<sup>12</sup>B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger, Malabar, FL, 1990), Vol. I.

# An inexpensive passive acoustic system for recording and localizing wild animal sounds

Sean A. Hayes

*Biology Department, A316 Earth and Marine Science Building, University of California, Santa Cruz, California 95064*

David K. Mellinger

*Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, P.O. Box 628, Moss Landing, California 95039*

Donald A. Croll and Daniel P. Costa

*Biology Department and Institute of Marine Science, A316 Earth and Marine Science Building, University of California, Santa Cruz, California 95064*

J. Fabrizio Borsani

*Istituto Centrale per la Ricerca Applicata al Mare, via di Casalotti 300, 00166 Roma, Italy*

(Received 1 June 1999; accepted for publication 26 February 2000)

An inexpensive animal recording and tracking system was designed, which uses sound-recording buoys deployed at several locations simultaneously in a passive hydrophone array. Each buoy contains a global positioning system (GPS) Location logger, a stereo digital audio tape (DAT) recorder with a hydrophone connected to one channel, and a VHF radio signal for time synchronization connected to the second channel. In a calibration test, three buoys were deployed in triangle formation at 1.8-km spacing. Light bulb implosions were localized to an accuracy of 60 m at the array center. These buoys are far less expensive than most marine acoustic tracking systems. The instrument package can be used for drift, moored, or terrestrial applications. © 2000 Acoustical Society of America. [S0001-4966(00)02906-4]

PACS numbers: 43.30.Sf, 43.30.Yj, 43.58.Vb [WA]

## INTRODUCTION

Investigators of the acoustic behavior and movements of vocalizing wild animals are challenged by economic, logistical, and technical barriers. These barriers can be especially high in the open-ocean environment, where animals that spend the majority of their time underwater are often far from shore, fast-moving, highly dispersed, and interacting over distances greatly exceeding visual range (Costa, 1993). Addressing questions about social systems, foraging behavior, population densities, and management can be quite difficult.

A common need is to localize and track animals that are underwater, hidden beneath ice, in forests, or otherwise difficult to track visually. The marine environment is a poor conductor of light but an efficient propagation medium for sound. Many marine species, especially cetaceans, have evolved sophisticated sound production and reception mechanisms to aid in meeting their requirements for foraging and reproduction. These species' natural history can be studied through the acoustic signals produced during their activities. Such animals may be tracked acoustically by collecting sound from several locations simultaneously and using time-of-arrival differences to estimate locations (Spiesberger and Fristrup, 1990). Typical tools for acoustic research in the pelagic environment include towed hydrophone arrays, bottom-deployed arrays such as SOSUS arrays, large ships, multichannel signal conditioning/recording systems, and sonobuoy/receiver systems (Nishimura and Conlon, 1994). The price, signal processing skills, and military relationships associated with these systems make them inaccessible to many marine biologists.

To reduce the cost and improve the accessibility of acoustic tracking methods, an acoustic localization system was built consisting of commercial off-the-shelf components commonly available from hardware, marine supply, and audio electronics stores. It consists of several independently drifting, time-synchronized recording systems, similar to sonobuoys in concept except that these are recoverable buoys that record sound data instead of transmitting it by radio. In operation, these buoys record sound signals, time-synchronization signals, and GPS locations. Laboratory analysis of the recordings allows vocalizing animals to be located and tracked.

## METHODS

### Buoy design

Each buoy contains a stereo digital audio tape (DAT) recorder (Sony TCD-8; frequency response flat  $\pm 1$  dB from 9 Hz to 22 kHz) with one input channel connected to a hydrophone (Hi-Tech HTI-SSQ-41B; frequency response flat  $\pm 1$  dB from 10 Hz to 30 kHz) for collecting acoustic data. The second input channel of the DAT recorder receives the audio output of a marine VHF radio receiver; this signal is used during analysis to time-align the sounds recorded on the separate buoys. A nondifferential global positioning system (GPS) data logger (Garmin 45) documents the buoy's position as it drifts during recording sessions (Fig. 1).

Instruments and ballast are encased in waterproof spar-buoy PVC housings for deployment at sea (Fig. 1). The housing design utilizes a spar shape for the buoy. This shape has a small water-plane area, damping the impact of wave

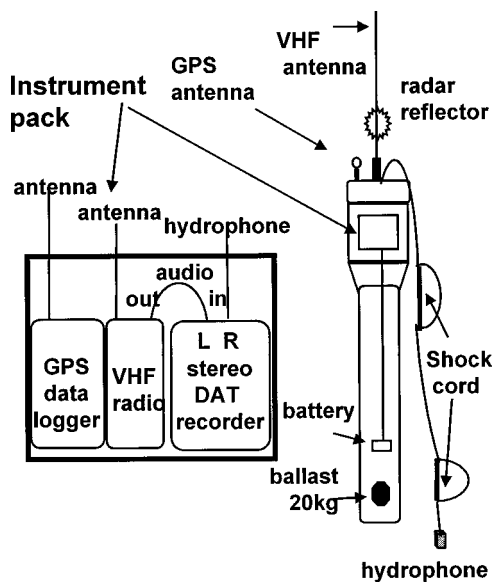


FIG. 1. Spar buoy instrument pack and housing. Each buoy contains an instrument pack with a stereo DAT recorder, a marine VHF radio receiver, and a GPS receiver/logger. The DAT has one input channel connected to an external hydrophone for collecting acoustic data, with the other channel connected to the VHF radio's audio output for time alignment. A 2.2-m waterproof spar-shaped PVC housing encases instruments for deployment at sea.

action on vertical buoy motion and reducing flow noise over the hydrophone. A 1–2-m length of shock cord is attached to the hydrophone cable near the buoy to further damp the effect of wave action and cable strum. Depending on the application and available resources, very high frequency (VHF) radio tags, strobes, and/or radar reflectors can be attached to the buoy to aid in tracking and recovery. Package price is approximately \$1900 per buoy.

### Deployment

These buoys are deployed in a drifting ring around target animals. Vocalizations are localized using time delays between buoys as explained below. Three or more buoys are needed for a localization system. Deployment time of the buoys in a 1-km triangle grid is approximately 1 h. Time is dependent on the spacing between buoys, sea state conditions, and speed of the deployment vessel. Maximum recording time, about 6 h, is achieved with a 90-m tape and the DAT recorder set to “long-play” mode (32-kHz sampling rate). A single recording session can therefore provide 5 h of time-synchronized data.

The hydrophone and cable trail upwind behind the buoy as it drifts. During recovery, it is best to approach drifting buoys from downwind to reduce the chance of propeller entanglement and a lost hydrophone. Moored buoys should be approached from upwind/upcurrent for similar reasons.

### Analysis

Localization analysis consists of three steps: time-alignment of the three (or more) hydrophone recordings, determination of differences in times that animal vocalizations occurred in the three recordings, and estimation of the animal's location.

### Time-alignment

Each stereo tape recorded in a buoy contains one channel with the hydrophone sound signal and the other channel with audio from the VHF radio receiver. This tape recording is transferred to a computer as a two-channel sound file. Time-alignment of the hydrophone recordings is accomplished using the VHF audio signal. The VHF receivers in all buoys are tuned to the same frequency, so that all the radio receivers relay the same audio signal synchronously. Hydrophone signals from all buoys are synchronized by time-aligning the corresponding VHF radio signals.

Alignment is performed by choosing one buoy as the reference. For the other two “aligning buoys,” the stored VHF audio signals are cross correlated with the reference VHF audio signal. The time offset of the cross-correlation function's peak is the amount by which the two signals are offset in time (van Trees, 1968). Each aligning buoy's hydrophone signal is time-shifted to bring it into alignment with the reference buoy's hydrophone signal. After all signals have been shifted and brought into time-alignment, the VHF signals are discarded, and the hydrophone signals are stored as a single sound file containing three time-synchronized channels.

### Time-delay estimation

To estimate the differences in arrival times of the animal vocalization at each buoy's hydrophone, either the waveform is measured directly or a cross correlation is calculated. In the direct measurement method, useful for loud, abrupt sounds, the waveform of each signal is examined. The onset time—the instant at which the sound first appears in the waveform—is measured. Time-of-arrival differences between hydrophone signals are calculated by subtraction of arrival times.

In the cross-correlation method, time differences are determined for each possible pair of hydrophones. The portion of two hydrophones' sound signals containing a vocalization are cross correlated. The time-offset peak in the cross-correlation function specifies the time difference between the arrivals of the vocalization at the hydrophones. Cross correlation can be limited to the frequency band of the vocalization, thus removing some noise (Clark *et al.*, 1996).

### Location estimation

Conceptually, a time-of-arrival difference between a pair of hydrophones determines a hyperbola on which the vocalizing animal must lie. In an ideal medium, the intersection of all the hyperbolas would be the animal's location. Since there is noise in each of the signals, the hyperbolas do not intersect at exactly the same point. A least-squared-error fit is used to determine the best location.

### Calibration

Two calibration sessions were conducted to determine the accuracy of this drifting buoy array over the submarine canyon in Monterey Bay, California, at 36° 47'N 122° 00' W, in water approximately 900-m deep. Three buoys were deployed in a rough equilateral triangle

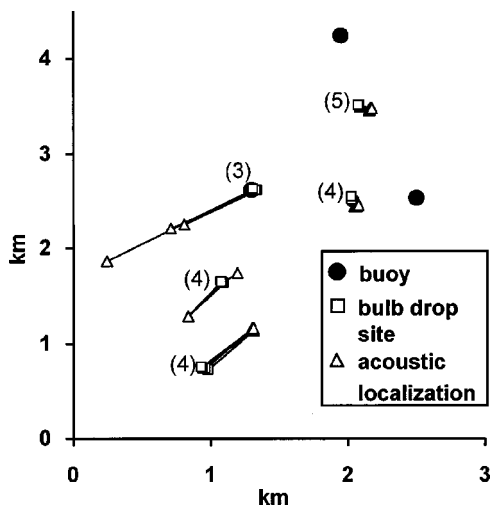


FIG. 2. Calibrating localizations with the spar buoy array. During the second of the two calibration sessions, 20 light bulbs were sunk and imploded around the array. One of the ensuing localizations was rejected (see text), resulting in the 19 localizations shown here. For bulbs dropped within the array, the mean distance between the recorded GPS position of the drop site and the acoustically determined position of each bulb was  $73 \pm 17$  m ( $n = 9$ ). For bulbs dropped outside the array, the mean distance was  $586 \pm 299$  m ( $n = 10$ ). The mean distance for all 19 points was  $306 \pm 344$  m ( $n = 19$ ).

approximately 1.8 km per side, with hydrophones at approximately 25 m depth. A total of 30 lead-weighted incandescent light bulbs were dropped and imploded at several positions in and around the array (Heard *et al.*, 1997). Implusions were recorded and localized by the waveform measurement method described above and the results were compared to positions measured by GPS on the research vessel deploying the light bulbs.

## Results

One source of error was drift in clock speed between the DAT recorders. Drift rates between machines were 0.5–2 ms/min. However, drift rate was consistent, and by sending several calibration signals per hour, the rates could be calculated. Based upon these rates, a correction factor was determined for each DAT and introduced into the time delay measurements.

The results of the calibration tests match theoretical predictions reasonably well (Fig. 2). The best localizations are predicted for sounds occurring within the array, while accuracy decreases with distance from the array, especially outside the corners. Of the 30 bulbs dropped during the two calibration sessions, two localizations outside the array were rejected because the localization analysis produced divergent, nonintersecting hyperbolas. Inside the array, the mean difference between the GPS positions of the drop sites and the acoustically determined positions was  $68 \pm 22$  m ( $n = 10$ ). Outside the array, mean error was  $567 \pm 642$  m ( $n = 18$ ).

Another potential source of error is multipath arrivals of a signal at the buoys. The light bulb implosion data were inspected for this. Several multipath arrivals with monotonically decreasing amplitude were observed when bulbs were imploded near buoys. The arrival times were assumed to

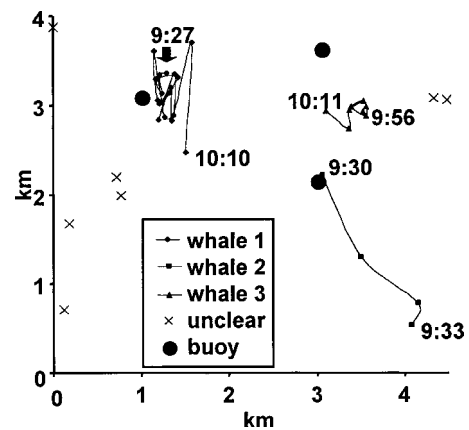


FIG. 3. Acoustic localizations of foraging blue whales recorded on July 19, 1996. Shown here are tracks of signals recorded from three or more animals; times are hours:minutes. Because the localizations for whale 2 are off one corner of the array, they have large range errors (see Fig. 2), and it is likely that the track shown here represents these errors rather than actual whale motion. These signals were attributed to one animal because of the consistent bearing from the array and the short time period between signals. It is also possible that whales 2 and 3 are the same individual, with whale 2 returning to the edge of the array and resuming vocalizations after a period of silence. “x” symbols denote localizations for which it was unclear whether the sounds were produced by whales 1–3 or other individuals.

occur in the following order: direct path, surface bounce, bottom bounce, surface-bottom bounce, and bottom-surface-bottom bounce. Based upon this assumption, we were able to calculate bottom depth, hydrophone depth, and bulb implosion depth. It was then possible to determine the exact time of implosion. Measured travel times to receivers matched most closely with direct path estimated travel times. The first arrival was at least 10 dB louder, and usually closer to 20 dB louder, than bottom-bounce arrivals, enabling us to ignore bottom bounces in subsequent use with vocalizing animals.

Field deployments of this system around blue whales in the southern California Channel Islands ( $34^{\circ} 05' N$   $120^{\circ} 00' W$ ) were successful, with buoys proving to be excellent platforms for recording whales' low-frequency sounds. Figure 3 shows the tracks of several vocalizing blue whales. Buoys survived repeated deployments in highly variable weather conditions.

## DISCUSSION

This system was designed as a low-cost, easy-to-operate recording system that could be used for tracking sound sources in the open ocean. The results of the light bulb experiments indicate accuracy to within 60 m inside the array. Outside the array, accuracy is greatly reduced. However, the bearing of signals relative to the array is maintained, providing information that is still useful. The error inside the array compares favorably to the GPS positioning error of approximately 40 m (August *et al.*, 1994) and could probably be greatly reduced through the use of differential GPS transmitters and receivers. Most calculated locations were consistently southeast of drop sites. Another of 10–15-m of error could be associated with the 40-m hydrophone trailing behind the buoys in the northwesterly direction at a 25–50° angle.

Figure 3 is included as practical demonstration of use of the array. During a session in which blue whales were tracked, the array was used to distinguish between several vocalizing whales, allowing at least two pods—probably two individuals—to be tracked in time and space. Tracks were determined by linking successive vocalizations that were nearby in time and space. In addition to showing how vocalizing blue whales move with respect to each other, this information can be compared with data on prey field structure, sighting info, and tracks of tagged individuals (Croll *et al.*, 1998).

The primary advantages to this acoustic tracking system are price and ease of use. These buoys are far less expensive than multichannel signal conditioning and recording systems connected to fixed bottom-mounted arrays or long towed arrays with large tow ships. Another advantage over towed arrays is that after deployment, the research vessel's movements are unconstrained and it is free to perform other functions such as behavioral observations and tagging. In addition, the left-right ambiguity associated with towed arrays does not exist with these buoys. The errors associated with this system in the pelagic environment are not large when compared with the movements of pelagic species. While sonobuoys are sometimes available to researchers without charge, sonobuoy systems still require receivers and multichannel recording systems. In addition, they transmit on regulated frequencies, potentially entailing extra permitting issues. Finally, not all researchers have access to sonobuoys, and those that do could be subject to political changes denying access.

Buoys are advantageous because they may be deployed from small vessels wherever animals are located. The impacts to both the environment and the animals whose behaviors are being studied are reduced through the use of small vessels and recoverable buoys. They have been used in sea state conditions through Beaufort 5. This system is durable, easy to use, and can be deployed and recovered with only two people.

The simplicity of this system makes it feasible for biologists without engineering backgrounds to collect and analyze sound data in new ways. The buoys are easily assembled from commercial off-the-shelf components. Most personal computers now have sufficient memory and hard disk space for the data analysis, and software for the time-alignment and localization steps is available from the authors.

This system may also be deployed as a moored array in the coastal environment, or as time-synchronized recording stations in terrestrial and polar-ice environments. Other studies with better positional accuracy than nondifferential GPS have achieved much higher animal-location accuracy (McGregor *et al.*, 1997; Janik *et al.*, in press). As the array does not move in these environments, increased positioning accuracy is possible, and onboard GPS units would not be required, further reducing cost.

This system has some disadvantages. Unlike sonobuoy recordings, data are not acquired until the buoys are recovered. However, no buoy has been lost in over 50 deployments. Due to DAT tape limitations, recording sessions last

only 5 h per deployment, not enough time for some applications. However, retrieving and redeploying the buoys with new tapes may extend this time limit, which takes only slightly longer than the initial deployment time. An alternative solution would be to use an A/D microcontroller computer system with programmable sampling rates and large storage medium.

It is hoped that this system will provide a much larger group of researchers with the acoustic tracking technology necessary to study the movements and other behaviors of marine and terrestrial organisms. By doing so, researchers may be able to gain new insights into questions about social interactions, foraging, population structure, and conservation.

## ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance of Capt. Rosalia Pavia of the STAR ROVER, Capt. Antonio Molinari of the BESTACCIA, Caitlyn Toropova, Steve Mitchell, John Speisberger, Vince Premus, Jason Gedamke, Michela Guisti, Monika Montefalcone, Vanesa Lopez, Anna Kaminska, Juan Carlos Salinas, Gabriela Valasquez, and Jen Jolly. We also appreciated all reviewers' comments. This work was supported by Office of Naval Research Grant No. N00014-95-10646 to D. Croll, Scripps Institution of Oceanography Grant No. ARPA MDA 972-93-1-003 to D. Costa, National Oceanic and Atmospheric Administration Contract No. 40ABNC602168 to D. Croll, Friends of Long Marine Lab, American Museum of Natural History, Lerner-Gray Fund, Earl and Ethyl Myers Trust Fund, a GAANN Fellowship to S. Hayes, and David and Lucile Packard Foundation funds to D. Mellinger.

- August, P., Michaud, J., Labash, C., and Smith, C. (1994). "GPS for environmental applications: Accuracy and precision of locational data," *Photogramm. Eng. Remote Sens.* **60**, 41–45.
- Clark, C. W., Charif, R., Mitchell, S., and Colby, J. (1996). "Distribution and behavior of the bowhead whale, *Balaena mysticetus*, based on analysis of acoustic data collected during the 1993 spring migration off Point Barrow, Alaska," *Sci. Report, Intl. Whaling Commn.* **46**, 541–554.
- Costa, D. P. (1993). "The secret life of marine mammals," *Oceanography* **6**, 120–127.
- Croll, D. A., Tershy, B. R., Hewitt, R. P., Demer, D. A., Fiedler, P. C., Smith, S. E., Armstrong, W., Popp, J. M., Keikhefer, T., Lopez, V. R., Urban, J., and Gendron, D. (1998). "An integrated approach to the foraging ecology of marine birds and mammals," *Deep-Sea Res., Part II* **45**, 1353–1371.
- Heard, G. J., McDonald, M., Chapman, N. R., and Jaschke, L. (1997). "Underwater light bulb implosions: A useful acoustic source," *Proc. Oceans '97* **2**, 755–762.
- Janik, V. M., Van Parijs, S. M., and Thompson, P. M. (2000). "A two-dimensional acoustic localization system for marine mammals," *Mar. Mammal Sci.* **16**, 437–447.
- McGregor, P. K., Dabelsteen, T., Clark, C. W., Bower, J. L., Tavares, J. P., and Holland, J. (1997). "Accuracy of a passive acoustic location system: Empirical studies in terrestrial habitats," *Ethol., Ecol. Evol.* **9**, 269–286.
- Nishimura, C. E., and Conlon, D. M. (1994). "IUSS dual use: Monitoring whales and earthquakes using SOSUS," *Mar. Technol. Soc. J.* **27**, 13–21.
- Spiesberger, J. L., and Frstrup, K. M. (1990). "Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography," *Am. Nat.* **135**, 107–153.
- van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory* (Wiley, New York), Vol. 1, 200 pp.

# Sidelobe reduction of limited diffraction beams with Chebyshev aperture apodization

Shiping He and Jian-yu Lu

Ultrasound Laboratory, Department of Bioengineering, The University of Toledo, Toledo, Ohio 43606

(Received 21 July 1998; accepted for publication 16 February 2000)

A limited diffraction beams (LDB) imaging system with Chebyshev weighting is presented. The objective of the paper is to reduce the sidelobes of the LDB without impacting on main-lobe performance and increase the contrast-resolution of the imaging system. The Chebyshev weighting is applied to the LDB and an analytic description and the simulation results are obtained. Theoretical analysis and simulation results show that the LDB with Chebyshev weighting can reduce sidelobes, and improve imaging system performances. © 2000 Acoustical Society of America.

[S0001-4966(00)04605-1]

PACS numbers: 43.35.Yb [HEB]

## INTRODUCTION

Limited diffraction beams (LDB) can propagate without changing their waveforms in both space and time provided that they are produced with an infinite aperture and energy.<sup>1,2</sup> Even if produced with a finite aperture, they have a large depth of field. Because of this advantage, limited diffraction beams have been used in medical imaging and many practical applications.<sup>3,4</sup> However, sidelobes of these beams are larger than conventional focused beams at their focuses. Sidelobes may lower the contrast in medical imaging. When the sidelobes of the LDB are reduced, the width of the main lobes are usually increased. A wide main lobe will reduce the lateral resolution of medical imaging systems.<sup>5,6</sup> In addition, sidelobes increase the effective sampling volume and thus average out spatially distinguished information in tissue identification. Sidelobes are also a source of multiple scattering that produces artifacts in nondestructive evaluation of materials.

In this paper, Chebyshev weighting function is used to reduce sidelobes of LDB without increasing the main-lobe width. To produce a limited diffraction beam with Chebyshev weighting in a circular two-dimensional (2D) array transducer, an aperture that produces LDB is multiplied with a Chebyshev weighting distribution array that is obtained from the Chebyshev function. A theoretical analysis for using the Chebyshev weighting method to reduce the sidelobes of LDB was developed. Simulations and analysis of the results with a finite aperture 2D annular array transducer using the Chebyshev weighting were performed. The analytical and simulation results show that LDB with Chebyshev weighting can reduce sidelobes, and thus improve the imaging system performances.

## I. LIMITED DIFFRACTION BEAMS WITH CHEBYSHEV WEIGHTING DISTRIBUTION

The geometrical configuration of the transducer with Chebyshev weighting distribution is shown in Fig. 1. The spectra of limited diffraction beams (*X* waves and Bessel beams) with Chebyshev weighting distribution are given by

$$\begin{aligned} \Phi_{X_n f_m} \left( \mathbf{r}, \frac{\omega}{c} \right) &= \frac{2\pi}{c} e^{in\phi} B \left( \frac{\omega}{c} \right) J_n \left( \frac{\omega}{c} r \sin \zeta \right) H \left( \frac{\omega}{c} \right) \\ &\times e^{- (\omega/c)(a_0 - iz \cos \zeta) f(I_m)} \\ &(m, n = 0, 1, 2, \dots) \end{aligned} \quad (1)$$

and

$$\begin{aligned} \Phi_{B_n f_m} \left( \mathbf{r}, \frac{\omega}{c} \right) &= J_n(\alpha r) e^{in\phi} e^{i(\beta z - \omega t)} f(I_m) \\ &(m, n = 0, 1, 2, \dots), \end{aligned} \quad (2)$$

respectively, where  $r$  is a radial distance,  $\omega$  is angular frequency,  $\alpha$  is a constant,  $m, n$  is an integer,  $B(k)$  is the transmitting or receiving transfer function of a transducer,  $J_n(x)$  is an  $n$ th-order Bessel beam function,  $H(\omega/c)$  is the Heaviside step function,  $a_0$  is the constant that determines the fall-off speed of the high frequency component of the *X* waves,  $\beta = \sqrt{(\omega/c)^2 - \alpha^2} > 0$  is real,  $c$  is the phase velocity, and  $f(I_m)$  is a Chebyshev weighting distribution.

Let us consider a 2D annular array of ten sources of uniform spacing  $d$  arranged as in Fig. 1. The individual sources have the amplitudes  $I_0, I_1, \dots, I_9$ , etc., as indicated, the amplitude distribution being symmetrical about the center of the array. The total field  $\Phi$  from the sources at a large distance in a direction  $\theta$  is then the sum of the fields of the symmetrical pairs of sources, or<sup>7</sup>

$$\Phi_{M_e}(\psi) = 2 \sum_{m=0}^{(M-1)/2} I_m \cos \left( \frac{2m+1}{2} \psi \right), \quad M \text{ even}, \quad (3)$$

$$\Phi_{M_o}(\psi) = 2 \sum_{m=0}^{M/2-1} I_m \cos \left[ (2m) \frac{\psi}{2} \right], \quad M \text{ odd}, \quad (4)$$

where  $\psi = (2\pi d/\lambda) \sin \theta$ ,  $\theta$  is the angle shown in Fig. 1,  $M_e$  is an even number,  $M_o$  is an odd number. Each term in Eqs. (3) and (4) represents the field due to a symmetrically disposed pair of the sources.

It is well-known that  $\cos m(\psi/2)$  can be expressed as a polynomial of degree  $m$ . Thus, Eqs. (3) and (4) are expressible as polynomials of degree  $2k+1$  and  $2k$ , respectively, since each is the sum of cosine polynomials of the form  $\cos m(\psi/2)$ . If we now set the array polynomial as given by

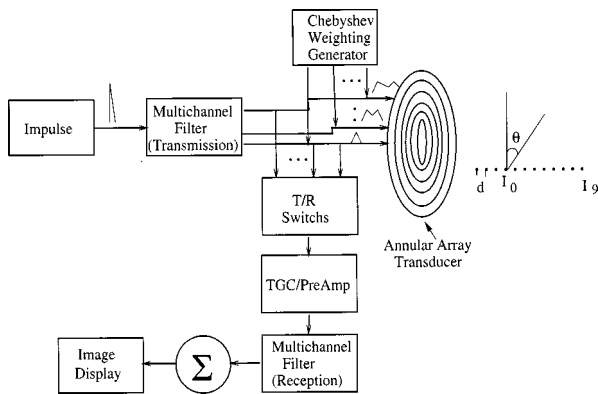


FIG. 1. 2D annular array imaging system with Chebyshev weighting distribution.

Eq. (3) or (4) equal to the Chebyshev polynomial of like degree ( $m=k-1$ ) and equate the array coefficients to the coefficients of the Chebyshev polynomial, then the amplitude distribution given by these coefficients is a Chebyshev distribution and the field pattern of the array corresponds to the Chebyshev polynomial of degree  $k-1$ .

Chebyshev weighting polynomial has a binomial series form as

$$T_m(x) = \cos^m \xi - \frac{m(m-1)}{2!} \cos^{m-2} \xi \sin^2 \xi + \frac{m(m-1)(m-2)(m-3)}{4!} \cos^{m-4} \xi \sin^4 \xi - \dots, \quad (5)$$

where  $T_m(x)$  is Chebyshev weighting polynomials and  $x = \cos \xi$ . The Chebyshev function in Eq. (5) has the following properties: First, the function of all orders passes the point (1,1). Second, for values of  $x$  in the range  $[-1,1]$ , the polynomials lie between  $-1$  and  $+1$ , and all roots of the polynomials are within  $[-1,1]$ .

Let the ratio of the main-lobe maximum to the sidelobe level be specified as  $R$ , i.e.,

$$R = \frac{\text{main-lobe maximum}}{\text{sidelobe level}}. \quad (6)$$

If  $x_0$  satisfies  $T_{m-1}(x_0) = R$ , the point  $(x_0, R)$  on the  $T_m(x)$  curve corresponds to the main-lobe maximum, while the sidelobes are confined to a maximum value of unity. The important property of the Chebyshev polynomial is that if the ratio  $R$  is specified, the beamwidth to the first null is minimized. The corollary also holds that if the beamwidth is specified, the ratio  $R$  is maximized (sidelobe level minimized).

The procedure will now be summarized. For an array of  $m$  sources, the first step is to select the Chebyshev polynomial of the same degree as the array polynomial, Eq. (3) or (4). This is given by  $T_{m-1}(x)$ . Next we choose  $R$  and solve  $T_{m-1}(x) = R$  for  $x_0$ . The beam pattern polynomial, Eq. (3) or (4), may now be expressed as a polynomial of  $\cos(\psi/2)$ . The final step is to equate the Chebyshev polynomial of  $T_{m-1}(x)$  and the array polynomial obtained by substituting  $x$

$= x_0 \cos(\psi/2)$  into Eq. (3) or (4), where the smallest value of  $\psi$  is 0 for  $x = x_0$  (center of the main lobe). Thus,

$$T_{m-1} = \Phi_{m-1}. \quad (7)$$

The coefficients of the array polynomial are then obtained from Eq. (7), yielding the Dolph-Chebyshev amplitude distribution which is an optimum for the sidelobe level specified.

## II. CHEBYSHEV WEIGHTING COEFFICIENT

An array of  $m=20$  in-phase isotropic sources, spaced apart, is to have a sidelobe level 23-dB below the main-lobe maximum. Find the amplitude distribution fulfilling this requirement that produces the minimum beamwidth between first nulls, and plot the field pattern. The Chebyshev polynomial of degree  $m-1$  is  $T_{19}(x)$  and the value of  $x_0$  for  $T_{19}(x) = R$ , where  $R = 15$  may be determined by trial and error from the expansion as given in Eq. (5) or may be calculated from

$$x_0 = \frac{1}{2} [(R + \sqrt{R^2 - 1})^{1/(m-1)} + (R - \sqrt{R^2 - 1})^{1/(m-1)}]. \quad (8)$$

For  $R = 15$  and  $m-1 = 19$  in Eq. (8), we have  $x_0 = 1.016$ .

Now substituting  $x = x_0 \cos(\psi/2)$  in Eq. (3), we have

$$\Phi_{19} = \frac{26\,214I_9}{x_0^{19}} x^{19} - \frac{1\,245\,184I_9 - 65\,536I_8}{x_0^{17}} x^{17} + \dots - \frac{19I_9 - \dots - I_0}{x_0} x. \quad (9)$$

The Chebyshev polynomial of like degree  $m-1$  is given by

$$T_{19} = 26\,214x^{19} - 1\,245\,184x^{17} + \dots - 19x. \quad (10)$$

Now equating Eqs. (9) and (10),

$$\Phi_{19} = T_{19}. \quad (11)$$

For Eq. (11) to be true requires that the coefficients of Eq. (9) are equal to those of the terms of like degree in Eq. (10). Therefore,

$$I_9 = x_0^{19} = 1.352. \quad (12)$$

In a similar way we obtain

$$I_8 = 1.514, \quad I_7 = 2.717, \quad I_6 = 4.231, \quad I_5 = 7.1, \quad I_4 = 12.7, \\ I_3 = 19.8, \quad I_2 = 103, \quad I_1 = 268, \quad I_0 = 758. \quad (13)$$

The relative amplitudes of the ten sources are then given by

$$f(I_{(0-9)}) = [1, 1.12, 2.01, 3.13, 5.24, 9.42, 14.65, 76, 198, 561]. \quad (14)$$

Weighting coefficient as a function of the elements of array is given in Eq. (14).  $f(I_m)$  is the Chebyshev weighting distribution that is normalized by its minimum. The values of  $f(I_m)$  corresponding to a given set of values of  $I_m$  as obtained from Eq. (14) are produced with a limited diffraction beam array. If the array is a 20-element linear array,  $f(I_{(0-9)}) = [1, 1.12, 2.01, 3.13, 5.24, 9.42, 14.65, 76, 198, 561, 561, 198, 76, 14.65, 9.42, 5.24, 3.13, 2.01, 1.12, 1]$  has to be produced by each element separately. If the array is a ten-element an-



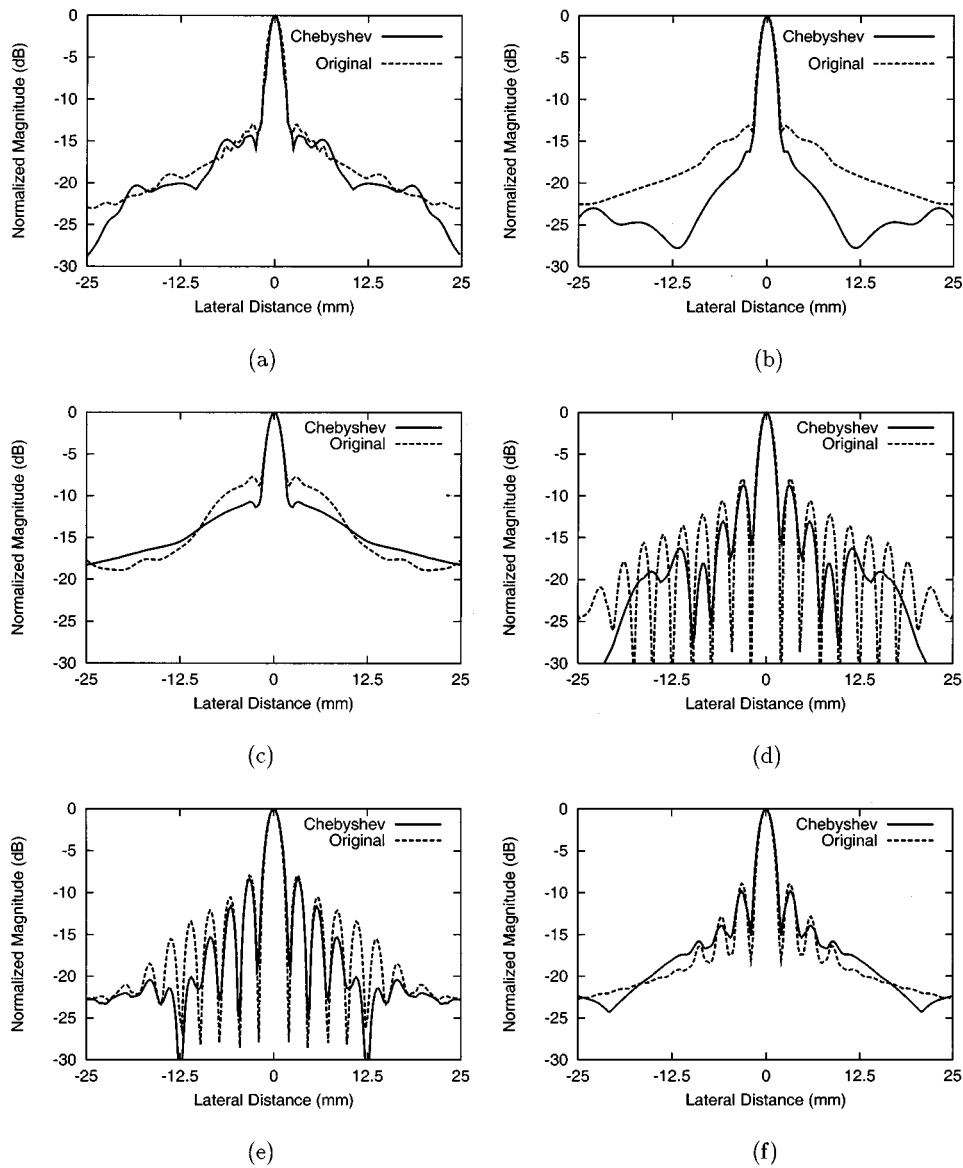


FIG. 2. Comparison between the peak lateral field responses of the original (dotted lines) and Chebyshev weighted (solid lines) limited diffraction beams. (a) X waves at  $z=50$  mm. (b) X waves at  $z=100$  mm. (c) X waves at  $z=200$  mm. (d) Bessel beams at  $z=50$  mm. (e) Bessel beams at  $z=100$  mm. (f) Bessel beams at  $z=200$  mm.

nular array,  $f(I_{(0-9)})=[1,1.12,2.01,3.13,5.24,9.42,14.65,76,198,561]$  is produced with each element separately because an annular array is symmetric about its center point.

The operating mode of a limited diffraction beam imaging system with Chebyshev weighting is illustrated in Fig. 1. A Chebyshev weighting generator, which is controlled by a computer, is used to produce the Chebyshev weighting distribution  $f(I_m)$ . The excitation signals of each transducer element are multiplied with the corresponding Chebyshev weighting distribution to produce a broadband modified limited diffraction beam. Echo signals are received with the same ultrasound annular array transducer that is used in transmission.<sup>6</sup> The received signal from each element is connected to a T/R (transmit/receive) switch and then preamplified and compensated for attenuation with a TGC (time gain control) circuit. After multichannel filtering, the received signals are coherently summed according to the amount of charges of each ring, and images are constructed and displayed.

### III. SIMULATION RESULTS

A simulation for transmission of Chebyshev weighted LDB with a ten-element ultrasonic annular array transducer was performed [use the Chebyshev weighting distribution theory given in Eqs. (1), (2), and (14), where  $n=0$ ]. The following parameters are chosen in our simulation program: beam types are X waves and Bessel beams, orders of Bessel beams or X waves are of zero order, rotation of weightings for higher-order Bessel and X wave is 0, scaling parameter for Bessel beams is  $1202.45 \text{ m}^{-1}$ , parameter for determining high frequency decay of X waves is 0.05 mm, Axicon angle of X waves is  $6.6^\circ$ , selection for lateral (maximum sidelobes) beam plot, axial distances from the transducer are 50, 100, and 200 mm, respectively, use the Fresnel approximation, field radius is from 0 to 25 mm, field step size is 0.25 mm, field rotation is  $0^\circ$ , central frequency is 2.5 MHz, aperture stop radius is 25 mm, and weighting function is one-way Blackman window function.<sup>2</sup>

Figure 2(a), (b), and (c) are beam plots of the simulation results for the original limited diffraction zeroth-order X

waves and the  $X$  waves with Chebyshev weighting, with axial distances of 50, 100, and 200 mm, respectively, from the transducer. In this simulation,  $B(\omega/c)$  was assumed to have the form of a Blackman window function with a central frequency of 2.5 MHz and a  $-6$ -dB one-way bandwidth about 2 MHz. From Fig. 2(a) and (b), we see that the sidelobes of the  $X$  waves with Chebyshev weighting distribution are lower than those of the original zeroth-order limited diffraction  $X$  waves. However, at  $z=200$  mm [Fig. 2(c)], the sidelobes start to increase (near the boundary of the depth of field).

Figure 2(d), (e), and (f) are beam plots of the simulation results for the original zeroth-order limited diffraction Bessel beams and the Bessel beams with Chebyshev weighting produced with a finite aperture, at axial distances from the transducer of 50, 100, and 200 mm, respectively. From Fig. 2(d) and (e), we see that the sidelobes of the Bessel beams with Chebyshev weighting are lower than those of the original zeroth-order Bessel beams. Again, the sidelobes start to increase near the depth of field [200 mm, Fig. 2(f)].

In summary, from Fig. 2(a)–(f) one can see that LDB with Chebyshev weighting can reduce the sidelobes while maintaining a small beamwidth over a large depth of field.

#### IV. CONCLUSIONS

This paper first briefly reviewed the theory of the LDB, and pointed out the importance of reducing sidelobes of these beams. Then, a limited diffraction beam imaging system with Chebyshev weighting is proposed to reduce the

sidelobes. Theoretical analysis and computer simulation were performed to study the relationship between sidelobes of LDB and Chebyshev weighting. Results show that Chebyshev weighting is effective to reduce sidelobes without increasing main-lobe width. This will increase the contrast of an imaging system while maintaining a small beamwidth (high resolution) over a large depth of field.

#### ACKNOWLEDGMENTS

This work was supported by Grant No. HL 60301 from the National Institutes of Health.

<sup>1</sup>Jian-yu Lu and J. F. Greenleaf, "Experimental verification of nondiffracting  $X$  waves," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**(3), 441–446 (1992).

<sup>2</sup>Jian-yu Lu and J. F. Greenleaf, "Nondiffracting  $X$  waves—exact solutions to free-space scalar wave equation and their finite aperture realizations," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**(1), 19–31 (1992).

<sup>3</sup>Jian-yu Lu, "Experimental study of high frame rate imaging with limited diffraction beams," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**(1), 84–97 (1998).

<sup>4</sup>Jian-yu Lu, Xiao-Liang Xu, Hehong Zou, and J. F. Greenleaf, "Application of Bessel beam for Doppler velocity estimation," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **42**(4), 649–662 (1995).

<sup>5</sup>Jian-yu Lu, "Bowtie limited diffraction beams for low-sidelobe and large depth of field imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**(6), 1050–1063 (1995).

<sup>6</sup>Jian-yu Lu and J. F. Greenleaf, "Sidelobe reduction for limited diffraction pulse-echo systems," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **40**(6), 735–746 (1993).

<sup>7</sup>Edward A. Wolff, *Antenna Design with Fiber Optics* (Artech House, 1996).

# Real-time and off-line comparisons of standard array configurations containing three and four microphones

M. W. Hoffman, C. Pinkelman, X. F. Lu, and Z. Li

*Department of Electrical Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska 68588-0511*

(Received 21 April 1998; revised 23 December 1998; accepted 22 January 2000)

This paper addresses the benefits and drawbacks of different configurations of three and four microphone adaptive arrays. In particular, broadside and endfire configurations of equally spaced and logarithmically spaced arrays are compared. The assessment uses arrays of simple electret microphones in actual acoustic environments. Fixed and robust adaptive approaches are compared given a single interference. Adaptive endfire configurations appear to provide 2–5 dB better average performance using the same resources (microphones, aperture, and processing complexity) than adaptive broadside implementations. This is true in all cases except for interfering signals close to ( $<30^\circ$ ) the desired signal look direction (with broadside providing 1–4 dB relative improvement).  
© 2000 Acoustical Society of America. [S0001-4966(00)02205-0]

PACS numbers: 43.60.Gk, 43.60.Lq, 43.66.Ts [JCB]

## INTRODUCTION

Recent interest (Flanagan *et al.*, 1991; Welker *et al.*, 1997; Elko and Pong, 1997) in microphone array processing for speech enhancement raises the question of the relative importance of system resources: microphones, filter taps, aperture, and processing complexity. In the present study, we are interested in evaluating actual microphone arrays in realistic acoustic environments with a single interfering signal. The class of arrays we are investigating are linear arrays with a small number (three or four) of microphones arranged in endfire and broadside configurations with both equal spacing and logarithmic spacing between the microphones in the array. A previously described (Hoffman and Buckley, 1995) robust adaptive processing approach is evaluated relative to block processed optimum results and fixed array results in a variety of acoustic interference scenarios. The primary goals of the present study are to provide some direction for choosing microphone configurations and to generate some sense of what types of gains are possible with adaptive processing. Simple arrays of omnidirectional electret microphones are used to provide a clear estimate of the gains that the *robust adaptive processing* provides—actual systems may well employ more directional sensors.

## I. METHODS

### A. Speech materials

All input speech and interference signals are sampled at 10 kHz and converted through 12 bit analog-to-digital converters. The microphone outputs are amplified and anti-alias filtered with an adjustable Precision Filters R-6000 signal conditioning system (sixth order filters). Eighteen isolated spondees are used as the desired speech input into the array. Multi-talker babble (edited to remove periods of excessive activity or unusual quiet) is used as the interfering sound source. Both the spondees and the multi-talker babble have been supplied on Digital Audio Tape by *Auditec* of St. Louis ([www.auditec.com](http://www.auditec.com)). The speech-to-babble power ratio at the input to the array is set to  $-6$  dB.

### B. Acoustic environment

The room geometry and array setup are illustrated in Fig. 1. The angles marked are:  $-10$ ,  $0$ , and  $10$  degrees for desired signal locations and  $30$ ,  $60$ ,  $90$ , and  $150$  degrees for interfering source locations. Each test angle has three corresponding distances as can be seen in Fig. 1. These three distances are:  $0.5$  m,  $0.75$  m, and  $1.0$  m. These distances roughly correspond to direct-to-reverberant energy ratios ( $D/R$ ) of  $11$ ,  $8$ , and  $5$  dB, respectively.

### C. Array geometry

The standard array configurations tested consist of broadside and endfire arrays with three or four (Knowles EK-3132 omnidirectional) microphones that are either equally or logarithmically spaced. In the equally spaced cases, the distance between the microphones is  $d = 3.5$  cm. For the logarithmically spaced cases, the spacings are  $d$ ,  $2d$ , and  $4d$ , with  $4d$  being used only in the four microphone cases and  $d = 3.5$  cm in all cases.

### D. Robust algorithms

Robust adaptive algorithms (Hoffman and Buckley, 1995) are used for all array geometries. For comparison purposes the equivalent geometry fixed quiescent processors and the data block optimum processors (as described in Hoffman *et al.*, 1994) are also evaluated. The robust processor is designed by measuring the desired source to microphone transfer functions for a nominal location ( $0$  degrees and  $0.75$  m) and extrapolating from these measurements with modeled data to generate robust constraints (as per Hoffman and Buckley, 1995) for the desired signal region of interest: a  $\pm 10$  degree sector from  $0.5$  m to  $1.0$  m range. The same constraints are used for real-time processing and off-line sampled data processing. The number of taps per microphone was  $16$  for the three element arrays and  $12$  for the four element arrays so that for all systems the array weight vectors had a total length of  $48$ . This amount of processing was feasible for the real-time system. The real-time implementa-

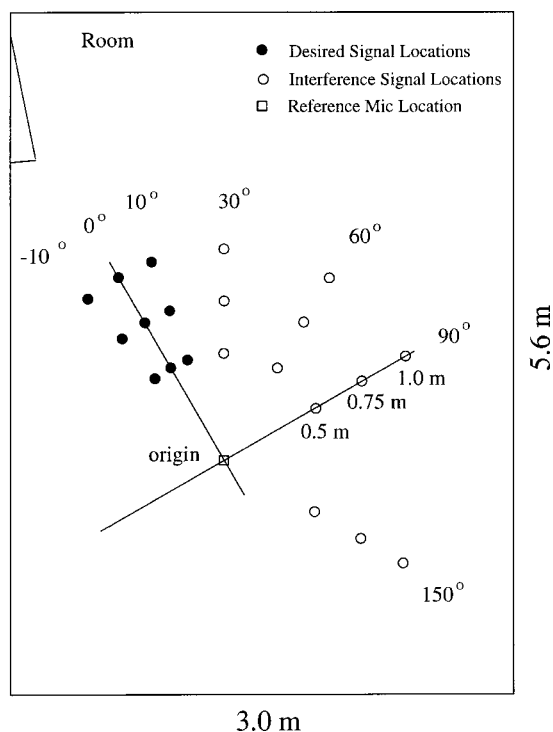


FIG. 1. Placement grid in testing room.

tion required a periodic block update of the adaptive algorithm weights every  $K$  samples (versus every sample). The value of  $K$  ranged from 20 to 30 and caused only a minor reduction in adaptive processing effectiveness.

### E. Real-time processor

The Texas Instruments TMS320C30 microprocessor (33 MHz) was used as the engine for the real-time system. A Spectrum Signal Processing C30 processing board and sister A/D board allowed extensive testing and intermediate output functions to be performed (such as freezing the adaptive weights and concurrent generation of fixed weight outputs). Sampled speech and babble signals from this real-time system were also saved to disk and used to allow off-line processing comparisons.

### F. Performance measure

The performance measure used is intelligibility averaged gain,  $G_{ai}$ . The calculation of the intelligibility averaged gain has been given in Hoffman *et al.* (1994).

Two notes should be made about the results presented in this study. First, all signals have been pre-emphasized [as suggested in Link and Buckley (1993)] prior to processing. This results in larger processing gains for the adaptive spatial filter than is possible with normal speech signals, whose total energy is dominated by low frequency components. Second, all performance measures are relative to a single omnidirectional microphone.

## II. RESULTS

In this section we examine three different aspects of the microphone array performance: optimum off-line versus

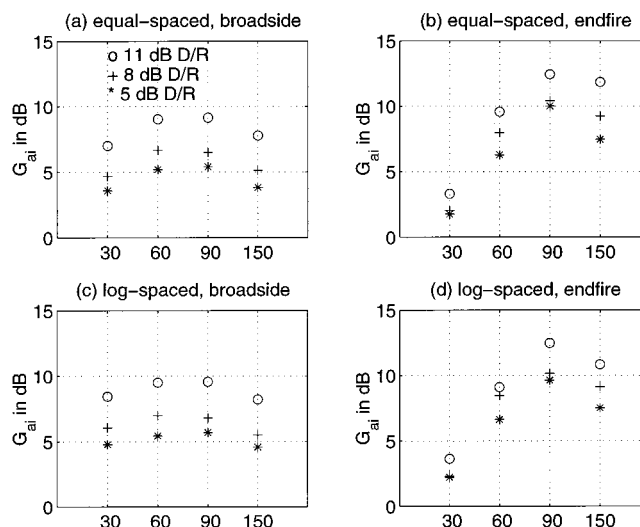


FIG. 2. Adaptive real-time processing results for various three microphone configurations with different levels of reverberation plotted versus interference angles of 30°, 60°, 90°, and 150°.

adaptive real-time processing, the impact of varying levels of reverberation on the processing results, and adaptive versus fixed array performance.

### A. Optimum off-line versus adaptive on-line processors

Experiments have been conducted to explicitly compare optimum off-line processing results with those obtained from adaptive real-time processing. The adaptive real-time system consistently provided intelligibility averaged gains,  $G_{ai}$ , within about 2 dB of the optimum off-line system. In many cases, performance was within 1 dB or even equal for the adaptive real-time and optimum off-line systems. The broadside systems demonstrated expected similarity in the 30° and 150° interference performance gains. The endfire systems showed a wider range of performance gains, generally providing enhanced performance relative to the broadside systems for all but the near-in 30° interference cases.

### B. Impact of reverberation level

In this section we report results parametrically as the  $D/R$ 's are varied from 5 to 11 dB. The change in  $D/R$  was accomplished by moving the sources to 0.5 m, 0.75 m, and 1.0 m. In other respects, the room acoustics were identical. In Fig. 2  $G_{ai}$ 's are illustrated for the various three microphone configurations as a function of interference angle of arrival. An "○" indicates the least reverberant 11-dB  $D/R$  condition, while the "+" and "\*" symbols represent results obtained in 8-dB and 5-dB  $D/R$  conditions, respectively. The  $D/R$ 's are approximate since they vary for each microphone in the array—they have been reported for the reference microphone, i.e., the one nearest the center of the array aperture. Note that the endfire arrays provide better performance for all but the 30° interfering signal case. The improved directional selectivity at array broadside accounts for the better performance of the broadside arrays for the 30° interference cases. Also, the endfire configurations provide better and

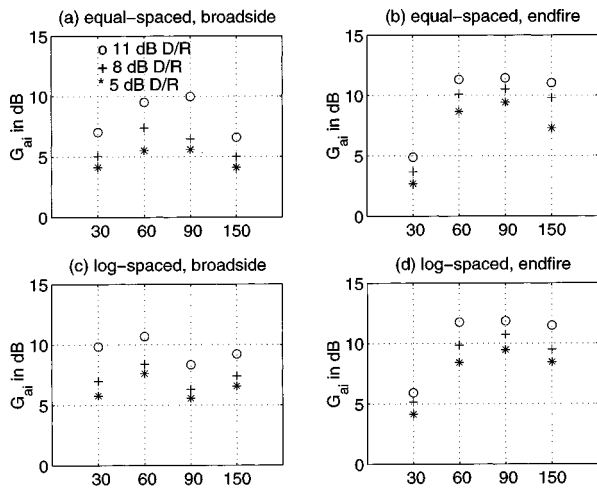


FIG. 3. Adaptive real-time processing results for various four microphone configurations with different levels of reverberation plotted versus interference angles of 30°, 60°, 90°, and 150°.

more consistent performance in more reverberant environments than do the broadside arrays. The increases in reverberation reduce the processing gain  $G_{ai}$  provided by the array, with decreases of 3–4 dB being common for a 6-dB increase in reverberant energy.

In Fig. 3 a comparable set of results are presented for the four microphone array configurations. Note the general similarity of the results to those shown for the three microphone arrays. For the broadside arrays, the performance for each of the configurations is remarkably similar. The largest difference occurs for the four microphone, log-spaced broadside array. This array has the largest aperture considered and it outperforms the other configurations for the near-in 30° interference case. For the endfire configurations the four microphone processing approaches appear to outperform the three microphone configurations for the 30° and 60° interference cases. Less of a performance difference is apparent at the higher angle interferences.

### C. Adaptive versus fixed processors

The adaptive implementation used in this study exploits the Generalized Sidelobe Canceller (GSC) (Griffiths and Jim, 1982) structure to allow a relatively simple implementation of the constrained minimum variance optimization approach. The GSC consists of a fixed upper path (fixed array processor) and an adaptive lower path that are combined to form our adaptive processor. Since the fixed path of the GSC provides fixed arrays with identical geometries to the adaptive processors we have tested, it is interesting to compare the performance of these fixed and adaptive processors. In Fig. 4  $G_{ai}$ 's are compared for fixed (“\*”) and adaptive (“+”) real-time processors in an 11-dB  $D/R$  condition. The fixed arrays provide from 0 to 6 dB of gain in various conditions. The arrays of three microphones presented have the smallest aperture considered (7 cm), while the arrays of four microphones use the largest aperture considered (24.5 cm). The difference in performance between the fixed responses and the robust adaptive responses were obtained using a GSC with the equivalent fixed processor and an identical

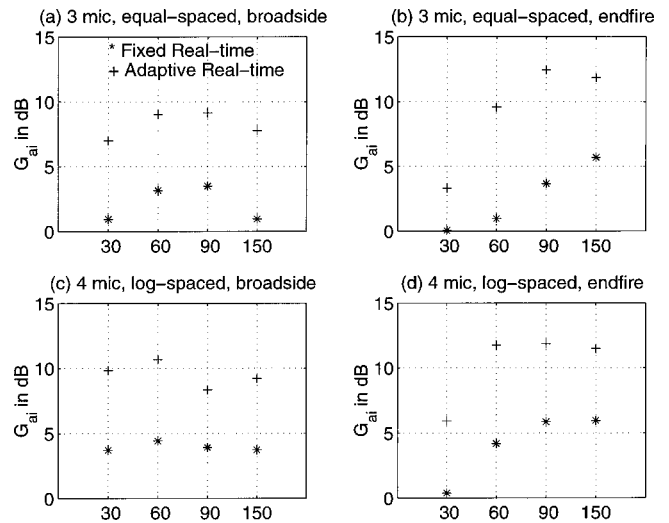


FIG. 4. Fixed versus adaptive real-time processing results for various three and four microphone configurations with  $D/R = 11$  dB plotted versus interference angles of 30°, 60°, 90°, and 150°.

array geometry. While better designs for fixed processors exist, these fixed processors could also be easily incorporated in the GSC structure by adding a single linear constraint to the robust constraint set (Griffiths and Buckley, 1987). The results presented here simply illustrate the performance improvement adaptivity provides over the fixed array. With an enhanced fixed beamformer in the GSC the adaptive processor would still provide improved performance over the same enhanced fixed processor in a similar scenario with a higher interference power. This is true because the adaptive processor attempts to minimize output power and is robust in the sense that it does not cancel the signal of interest. The effects of increasing reverberation on the fixed processors are small. Since the array gains provided by the fixed arrays are relatively low to start with, they degrade less with increasing reverberation. However, the performance remains substantially below that of the robust adaptive processors given an identical array geometry.

### III. DISCUSSION

An investigation into various standard three and four microphone array geometries has been conducted. Sampled signals from actual microphone arrays in real acoustic environments have been used to assess the interference cancellation of the different configurations. Generally, the robust adaptive beamforming outperforms fixed beamforming in identical situations—with the array aperture being the key factor for improving the fixed array performance. Among the adaptive beamforming geometries, the endfire configurations tended to outperform the broadside configurations—with the exception being in the case of “near-in” interference (i.e., 30° azimuth). The arrays with logarithmically spaced sensors performed better in situations with near-in interference. For multiple microphone applications, the use of larger apertures appears to be advantageous when the application permits. In addition, the performance of the endfire configuration appears to be more consistent with changes in reverberation level and angle of interference. However, the endfire con-

figuration would not be recommended for situations in which the expected interference is close (within 30° for our apertures) to the desired signal direction.

A couple of implications for future microphone array investigations can be drawn based on our systematic study of standard linear microphone arrays. First, the use of some combination of endfire and broadside placements in arrays of more microphones appears necessary for both good “near-in” interference suppression and robust performance in more reverberant environments. Second, the addition of a more directive fixed array in the robust adaptive GSC processor will provide an interesting set of comparisons to fixed processing. While the current study focussed on a single interfering source, more comparisons of fixed and robust adaptive processing in situations including both directional interfering sources and diffuse noise sources would be useful.

### ACKNOWLEDGMENTS

This work has been supported by the Whitaker Foundation, the National Science Foundation—IRI 9405286, and

the University of Nebraska-Lincoln’s Center for Communication and Information Sciences.

- Elko, G., and Pong, A. N. (1997). “A steerable and variable first order differential microphone array,” in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing, ICASSP*, pp. 223–226, Munich, Germany.
- Flanagan, J., Berkley, D., Elko, G., West, J., and Sondhi, M. (1991). “Autodirective microphone systems.” *Acustica* **73**, 58–71.
- Griffiths, L., and Buckley, K. (1987). “Quiescent pattern control in linearly constrained adaptive arrays,” *IEEE Trans. Acoust., Speech, Signal Process.* **35**, 917–926.
- Griffiths, L. J., and Jim, C. W. (1982). “An alternative approach to linearly constrained adaptive beamforming,” *IEEE Trans. Antennas Propag.* **AP-30**, 27–34.
- Hoffman, M., and Buckley, K. (1995). “Robust time-domain processing of broadband acoustic data,” *IEEE Trans. Speech Audio Process.* **3**, 193–203.
- Hoffman, M., Trine, T., Buckley, K., and VanTasell, D. (1994). “Robust microphone array processing for hearing aids: Realistic speech enhancement predictions,” *J. Acoust. Soc. Am.* **96**, 759–771.
- Link, M., and Buckley, K. (1993). “Pre-whitening for intelligibility gain in hearing aid arrays,” *J. Acoust. Soc. Am.* **93**, 2139–2145.
- Welker, D., Greenberg, J., Desloge, J., and Zurek, P. (1997). “Microphone array hearing aids with binaural output—Part II: A two microphone adaptive system,” *IEEE Trans. Speech Audio Process.* **5**, 543–551.

# Results of an independent analysis of the inverse beamformer for use in the advanced processor build process

G. Clifford Carter<sup>a)</sup>

Naval Undersea Warfare Center Division, Submarine Sonar Department, Newport, Rhode Island 02841

(Received 13 October 1999; revised 22 February 2000; accepted 8 March 2000)

The inverse beamformer (IBF) is a signal-processing algorithm developed for underwater applications. The advanced processor build (APB) is a recently developed process for sonar test and evaluation. This correspondence describes both and summarizes the results of an independent “data-driven” test and evaluation of the IBF by using the APB. Discussion includes the underlying reasons for gains observed over conventional systems and applicability to uncluttered acoustical environments. [S0001-4966(00)04106-0]

PACS numbers: 43.60.Gk, 43.60.Pt, 43.30.Wi [JCB]

## INTRODUCTION

The purpose of this manuscript is twofold. First, it provides members of the underwater acoustical signal processing community with a brief summary of a new advanced processor build, or APB, process for sonar test and evaluation. This process is being used for acoustical signal-processing algorithms being considered for use in sonar systems. Second, it provides a summary of the inverse beamformer (IBF) and a summary of the results of applying the APB process to the IBF system.

The remainder of the manuscript is organized as follows: Background on the APB Process, the IBF Algorithm, Results of Test and Evaluation of IBF, Observations and Discussions, Conclusions, and Acknowledgments.

## I. BACKGROUND ON THE ADVANCED PROCESSOR BUILD (APB) PROCESS

The U.S. Navy has begun a process to upgrade the hardware and software in its underwater, acoustic, signal-processing systems. Shrinking budgets and evidence that sonar systems could perform better led the U.S. Navy to develop and test new sonar signal-processing software and to increase the use of commercial-off-the-shelf (COTS) hardware.<sup>1</sup> Briefly, the new process develops and subjects signal-processing algorithms to a four-step test and evaluation process that is referred to as the APB process. This process is done nominally on an annual basis and, for example, APB-98 refers to the version of the APB being tested in 1998.

The first step in the APB process is a test by the algorithm developer of the type that might lend itself to patent performance claims by the inventor and publication in the literature. This first step is apt to “showcase” the advantages of the new algorithm with limited attention to its shortcomings. The second APB step is to subject the algorithm to independent verification and validation or IV and V testing. In the second step, an independent test group stresses the algorithm to determine its limits of performance and range of applicability. In this step, “closed” data sets are used, that

is, data sets that the developer has not had the opportunity to use. The third step is to integrate the proposed algorithm with other algorithms and test them in a real-time laboratory or “in-plant” setting using both synthetic and recorded data. This step, by its nature, is constrained by the available hardware. These tests are performed by an independent group and conducted at the organization with systems integration responsibility. The fourth and final step is to take a real-time, integrated sonar processing system to sea.

Following a successful fourth step, algorithms that have completed this process are adjusted, integrated, and become part of the Acoustic Rapid COTS Insertion (ARCI) system delivered to the U.S. Navy. (Guertin and Miller<sup>2</sup> provide an excellent overview of the ARCI program.) RADM Fages described the benefits of ARCI as contrasted with the earlier AN/BSY-2 legacy sonar system. He stated,

“... the most successful of this strategy is the next submarine sonar system, AN/BQQ-10, also known as Acoustic Rapid COTS Insertion, or ARCI. ARCI development costs were one-tenth (1/10) the cost of BSY-2 and shipset cost was less than one-thirtieth (1/30)... ARCI's flexible design allows the incorporation of periodic improvements through Advanced Processor Builds or APBs.”<sup>3</sup>

It is this new APB process that has offered the U.S. Navy the opportunity to investigate a number of innovative sonar-signal-processing algorithms. The IBF algorithm, which has been the subject of promising reports in the *Journal of the Acoustical Society of America*, was evaluated. The IBF summary-evaluation results are reported here as an illustration of the APB process and to provide a summary of test results to those who have read early reports on IBF.

## II. THE INVERSE BEAMFORMER (IBF) ALGORITHM

Nuttall, Carter, and Montavon<sup>4</sup> developed a method that optimally estimated power as a function of bearing. This work formed the theoretical underpinnings for an algorithm that would provide evidence that legacy sonar systems could be improved.<sup>5</sup> As with many theoretical developments, the initial work was conducted with simplifying assumptions necessary for mathematical tractability. These assumptions

<sup>a)</sup>Electronic mail: C.Carter@IEEE.org

included that the signals and receivers were in the farfield (that is, planar wave fronts) and in the same plane, in particular, that the acoustic environment was direct path without multipath. The output of this system was the noise (or signal) directionality spectrum, that is, estimated power versus bearing (or azimuth). This initial work was also conducted at a time when it was not practical to perform the extensive calculations in real time. Its use, therefore, languished primarily for computational reasons, but also owing to the apparent limited applicability to direct path propagation, that is, the theoretical development was limited to the case of no multipath. Two advances overcame this obstacle. First, the improvements in computer speed (Moore's law) coupled with an algorithmic advance by A. Nuttall allowed rapid computation for an equispaced linear array. Second, a successful sea-test was organized by J. Donald of the Naval Undersea Warfare Center (NUWC) and J. Wilson<sup>6</sup> of Neptune Sciences, Inc. This at-sea test was successful despite the fact that theory had only been shown to hold in the direct path acoustic environment. Indeed, the thought-provoking results suggested performance gains over older legacy sonar systems were possible, but the reported results seemed too good to be true. Also, at the time the cost of introducing engineering changes to legacy sonar systems was large compared to the U.S. Navy's soon-to-be implemented APB and ARCI process.

In 1993, Donald, Nuttall, and Wilson patented their processing system and called it an inverse beamformer or IBF system.<sup>7</sup> In 1995, two published articles described the promising results of early testing. This testing is now referred to as step-one APB testing. Referencing the earlier work of others, Wilson<sup>5</sup> stated that "... IBF consists of three separate algorithms ... the Fourier integral method (FIM) ... the eight nearest-neighbor peak picker (ENNPP) and the ' $M$  of  $N$ ' tracker ... it is shown that IBF has a 3-dB array noise gain ... (that) ... leads to a minimum detectable level (MDL) advantage of over 10 dB for IBF ..."

In the same 1995 issue of *J. Acoust. Soc. Am.*, Fabre and Wilson<sup>9</sup> stated that: "Inverse beamforming (IBF) has been very successful in detecting signals of very low levels in measured ocean data and during real-time at-sea experiments." ... They go on to state: "The large performance gains are attributed to ... FIM, ENNPP, and the  $M$  of  $N$  tracker." However, no comment is made about which of the three IBF components provides most of the gain. Fabre and Wilson further state that: "It is shown that the IBF MDLs are over 10 dB lower than those for CBF." Later, Carter (who is the author of this manuscript), working with others,<sup>10</sup> interpreted the term to mean conventional beamforming and energy detection *with the frequency band and integration settings of traditional legacy sonar systems*. Indeed, it was found and is reported here, that it was by changing three fundamental parameters: the integration time, the bandwidth and the center frequency, that IBF was able to perform better than legacy sonar systems. We therefore explicitly note here that the comparison of IBF to a legacy sonar system performed by Fabre and Wilson includes the integration processing times and the frequency bands of the older legacy sonar systems with longer IBF integration times and also

with IBF bandwidths better matched to the signals of interest. Fabre and Wilson then stated that, "The important point to emphasize is that these IBF tracker principles have performed extremely well for real submarine signals on measured data and in real time at sea as well as on the MDL evaluation in this paper."

### III. RESULTS OF TEST AND EVALUATION OF IBF

In July 1995, Carter<sup>11</sup> reported, based upon preliminary reports and preprints, that IBF is a submarine sonar signal-processing COTS system that has been taken to sea and has been the subject of some encouraging results.<sup>8,9</sup> The IBF system detects acoustic signals received using frequencies different than conventional systems. The IBF system senses the acoustic pressure field at a number of sensors and makes limited assumptions on signal type, but assumes uncorrelated noise and multiple uncorrelated plane wave sources, that is, no correlated multipath. This method inverts the received pressure field to detect and estimate the bearing to the incoming source or sources of acoustic energy. The IBF system then is effectively estimating the (signal and) noise directionality field. IBF uses a clever method devised by J. Wilson (called the Fourier Series Method, or FSM) and revised and improved for computational efficiency by A. Nuttall (Nuttall's method is called the Fourier Integral Method, or FIM, and is not to be confused with the more common Fisher Information Matrix). The FIM, an improved FSM, initially considers the acoustic array of sensors to be infinite in length, thereby allowing for solution of otherwise intractable integral equations for the noise directionality. The method then reintroduces the finite array length constraint so that the end result no longer utilizes an infinite array assumption.

The FIM portion of IBF has been evaluated and extended recently by Solomon, Knight, and Greening.<sup>12</sup> They looked at an adaptive beamformer, namely the minimum variance distortionless response (MVDR), and a conventional beamformer (what they call the Bartlett beamformer) and concluded, in part, that: "The results indicate that MVDR performs better with gain/phase corrections, and is better than FIM and the Bartlett beamformer." A more complete appraisal can be found in their excellent paper. Their analysis led them to develop an innovative extension of FIM. In particular, their extension was "FIM with non-equal weighting of covariance lags which are (they) called weighted FIM (WFIM)." Based upon their initial analysis with a limited data set, they also stated that: "The results obtained clearly indicate the superiority of WFIM ... The MVDR with array calibration performed almost as well as WFIM in directions where it was accurately calibrated ..." Solomon *et al.* have proposed additional research, including at-sea trials and a comparison with another inversion method. Through initial independent testing reported here and by Solomon *et al.*, we see that, in some environments, improvements to IBF are possible. We note that other more recent signal processing methods, such as the NOLOSS<sup>13</sup> method, matched field processing,<sup>14</sup> and WFIM, may also offer performance gains. These methods await independent step-two test and evaluation.



Any performance gains of IBF over legacy sonar systems, which have been reported in this journal, may come from several places. Gains may come from any of the three main IBF components, or combination of components, or from design decisions, including frequency range and integration time. The IBF system, like other algorithms in the development cycle, evolves and has more than one version. In addition, any particular version can be run with different parameter settings, making comparisons challenging. In the early investigation of IBF, it was determined that the bandwidth, frequency range, and effective integration time of IBF and an architecture that provided only a power output precluded testing of individual sub-system components. This was despite the fact that subsystem testing would have been preferred, in order to allow integration of successful IBF sub-systems components into APB-98. As a result of several discussions with the IBF developers, we determined that performance gains cited in this journal in 1995 came from two areas. Reported IBF gains over legacy sonar systems came from a better match to the frequency band of coherent energy and from an effective increase in integration time as compared to conventional, that is, legacy sonar systems. However, it was felt that a system's level test would provide IBF the best opportunity to demonstrate its capability, that is, what it "brought to the party." Therefore, the developers of the IBF processor met at NUWC Division, Newport, Rhode Island, during February 1998 for a step-two APB test. During this time, the IBF system, in its configuration as of February 1998, was tested with underwater acoustic data sets from a variety of environments.

The system detection performance was examined to address two questions. Does the IBF system have features that should be incorporated in a future APB, and under what circumstances does the IBF processor show performance that surpasses current APB technology? Given these goals, the detection performance was compared to APB-98 performance for the same data cuts. The testing was limited to comparisons of spectrograms (that is, estimated power versus frequency and time).

The data were broken into two categories: cluttered (with say, typically more than a dozen acoustic contacts, some of which are closely spaced in bearing and may move and cross through different beams) and uncluttered (with less than, say, a half dozen acoustic contacts and more typically two or three contacts). For the uncluttered data sets, IBF performance was essentially the same as APB-98 when APB-98 used the same filter bandwidth and integration times.<sup>15</sup> Detection metrics were essentially the same for both systems. For the cluttered cases, IBF generally held contacts less time than APB-98, and in at least one case, no IBF detection was obtained. Overall, the IBF system performance was generally inferior to that of the APB-98 system for cluttered environments. That is, environments where there were large numbers of acoustic contacts.<sup>16</sup>

Thus, in answer to one key question: no instances of superior performance by IBF under normal operational constraints were witnessed during our February 1998 tests. Given these results, it was decided that the IBF system as it

existed during the February 1998 tests not be installed in future APB implementations.<sup>17</sup>

Based on our evaluation, it appeared that the signal phenomena, around which the *M-of-N* tracker was designed, may have exploitable features. The nature of the tests performed and the IBF implementation tested did not allow a full evaluation of these signal features, or the mechanisms applied within the *M-of-N* tracker. The *M-of-N* tracker may have not only longer integration times than traditional sonar systems, but an ability to integrate along dynamic source trajectories. Additional step-two-like testing of FIM by the Australian researchers led to WFIM, discussed earlier.

#### IV. OBSERVATIONS AND DISCUSSIONS

As a result of extensive APB testing, IBF was not selected for integration into a U.S. Naval sonar processing chain. However, IBF is one of the early successful COTS sonar systems and did perform well in an uncluttered environment as a stand-alone system. Therefore the initial results of IBF, that performance of legacy sonar systems could be improved with increased integration time and different frequency coverage, were independently corroborated. Also, because of its initial apparent success, WFIM is being considered for use by the Australians as an extension to the FIM portion of IBF.<sup>18</sup>

In the opinion of the author, there are numerous reasons why IBF was not selected for incorporation in the APB series. One factor is the new way of acquiring sonar signal-processing systems that is now being utilized by the U.S. Navy. This process relies heavily on COTS and cooperative peer review and independent four-step testing using representative sea data. As such, the process is called "data-driven" in contrast to "model-driven" development. There are advantages and disadvantages to each approach. A further factor for evolutionary advances here, as contrasted with revolutionary advances, includes working closely with the targeted market, that is, the system in which one's algorithm is to be integrated. This can be done through various peer-review groups, including the Signal Processing Working Group (SPWG).

As illustrative of system integration challenges, the word "beamformer," in the name inverse beamformer, raises the false expectation that the IBF system output is voltage proportional to acoustic pressure in the water. But, the IBF system produces (squared and averaged) power estimates and not voltage proportional to acoustic pressure. This is important if one desires to insert an algorithm into a mature, developed systems architecture that anticipates access to voltages for reprocessing acoustic signals. Hence it is important to be mindful of how the completed research is to be included by the system's integrator if the target market is evolutionary.

#### V. CONCLUSIONS

A new APB evaluation process is described together with the IBF and its evaluation. A new way of acquiring sonar signal-processing systems has been put in place with lower cost and higher performance. This APB approach uses

COTS equipment to achieve transition to ARCI production sonar systems. It involves annual incremental processing improvements and a four-step test and evaluation process of, respectively, individual testing, independent stress testing, independent integration testing, and, finally, independent at-sea testing. As a result of independent testing of IBF, we concluded that there were no instances of superior performance by IBF that were witnessed during the February 1998 tests. Given these results, it was recommended that the IBF system as it existed in February 1998 not be installed in future sonar systems. There is, however, reason to explore the  $M$ -of- $N$  tracker. Also, more recently, Australian researchers have tentatively identified improvements to the FIM portion of IBF, which deserve further investigation. In addition, other underwater acoustic signal-processing algorithms will be developed and tested under the APB process for use in production sonar systems as advances occur.

### ACKNOWLEDGMENTS

The author gratefully acknowledges the U.S. Navy sponsorship of this work, including most recently, Mr. James Thompson, Capt. John Polcari, the late Mr. Art Bisson, and particularly, Mr. Robert Zarnich, who has been technically engaged in the IBF evaluation. The author also acknowledges the efforts of Dr. James Wilson, Neptune Sciences, Inc., Dr. Albert Nuttall, NUWC Division Newport, and Mr. James Donald, NUWC Division Newport, the developers of the IBF algorithm, for their explanation and aggressive defense of their work. These inventors provided early demonstrations of improved sonar signal-processing gains in an uncluttered environment using COTS equipment with new displays, matched frequency bands, and longer integration times. Indeed, the gains that they demonstrated were, in some sense, part of the benchmarks that were met or exceeded by a competing, integrated set of processing algorithms, called APB-98. The author further acknowledges the contributions of Dr. Richard Gramann and Frank Rule, Chief Sonar Technician (submarines) who participated in the technical evaluation of the IBF algorithm. They prepared a technical report that was briefed to U.S. Navy panels for their consideration. One such panel was co-chaired by this author and Dr. Clark Penrod. Portions of the work reported here are excerpted from the executive summary report by Gramann and Rule<sup>17</sup> and the Chief of Naval Operations letter report<sup>19</sup> signed out following reviews by Navy panels. The author also acknowledges the technical contributions of Dr. Harry Cox, Dr. Ira Dyer, and Dr. Art Baggeroer, who, together with the author, and other members of the SPWG and Technical Advisory Group, reviewed IBF and other proposed APB processing algorithms.

- <sup>1</sup>R. E. Zarnich, "A fresh look at broadband passive sonar processing," *Proceedings of the Adaptive Sensor Array (ASAP) Workshop*, 10–11 March 1999, 99–103 (1999).
- <sup>2</sup>N. H. Guertin and R. W. Miller, "A-RCI-The right way to submarine superiority," *Nav. Eng. J.* **110**(2) (1998).
- <sup>3</sup>M. I. Fages, "Submarine programs: A resource sponsor's perspective," *Sub. Rev.* 53–59 (1998).
- <sup>4</sup>A. H. Nuttall, G. C. Carter, and E. M. Montavon, "Estimation of the Two-Dimensional Spectrum of the Space-Time Noise Field for a Sparse Line Array," *J. Acoust. Soc. Am.* **55**, 1034–1041 (1974).
- <sup>5</sup>There were actually several demonstrations of the fact that legacy sonar systems could be improved by use of various acoustic signal-processing algorithms.
- <sup>6</sup>Dr. Wilson, a Naval Academy graduate and submarine-qualified officer, was also a Captain in the U.S. Naval Reserve and adjunct faculty member at the Naval Postgraduate School in Monterey, CA.
- <sup>7</sup>J. Donald, A. H. Nuttall, and J. H. Wilson, "Inverse beamforming sonar system and method," United States Patent 5,216,640, 1 June 1993.
- <sup>8</sup>J. H. Wilson, "Applications of inverse beamforming theory," *J. Acoust. Soc. Am.* **98**, 3250–3261 (1995).
- <sup>9</sup>J. P. Fabre and J. H. Wilson, "Minimum detectable level evaluation of inverse beamforming using outpost SUNRISE data," *J. Acoust. Soc. Am.* **98**, 3262–3278 (1995).
- <sup>10</sup>Other evaluators included R. Gramann, F. Rule, H. Cox, C. Penrod, and R. Zarnich.
- <sup>11</sup>G. C. Carter, internal—private communication, providing an initial technical evaluation of IBF to U.S. Navy decision makers (1995).
- <sup>12</sup>I. S. D. Solomon, A. J. Knight, and M. V. Greening, "Sonar array signal processing for sparse linear arrays," *Fifth International Symposium on Signal Processing and its Applications (ISSPA)*, Brisbane, Australia, 1999, Vol. 2, pp. 527–530.
- <sup>13</sup>I. M. G. Lourtie and G. C. Carter, "Signal detectors for random ocean media," *J. Acoust. Soc. Am.* **92**, 1420–1427 (1992).
- <sup>14</sup>A. 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).
- <sup>15</sup>For purposes of comparison with IBF, we can think of APB-98 as traditional signal processing (energy detection and cross correlation) with: (a) improved filter parameter settings for bandwidth, center frequency, and integration time, and (b) adaptive beamforming.
- <sup>16</sup>These test results are consistent with the recent Australian testing that showed MVDR adaptive beamforming is better than FIM under certain conditions. In particular, in the important case where the gain and phases of the sensor array elements were well known and well corrected for (that is, properly calibrated), the Australian researchers found that MVDR adaptive beamforming performed better than IBF with FIM. Also, they introduced an extension to the FIM component of IBF and call it WFIM, which their initial results find to be superior to FIM. The results of Solomon *et al.* are consistent with our test results that show adaptive beamforming outperforms IBF with FIM.
- <sup>17</sup>R. Gramann and F. Rule, Executive Summary, private communication (1998).
- <sup>18</sup>In comparing WFIM to other methods, one must use caution to do the comparison on an equal footing, in particular, since weighting in the covariance domain is a convolution in the frequency domain and effectively changes the shape of the frequency bandwidth (resolution) of the sonar system and vice versa. That is, averaging or convolution in the covariance domain is equivalent to frequency weighting.
- <sup>19</sup>Chief of Naval Operations, Washington, DC, private communication (14 December 1998).

# Localization of amplitude-modulated high-frequency noise

Geoff Eberle

*School of Psychology, Deakin University, Geelong, Australia*

Ken I. McAnally<sup>a)</sup> and Russell L. Martin

*Air Operations Division, Defence Science and Technology Organisation, Melbourne, Australia*

Patrick Flanagan

*School of Psychology, Deakin University, Geelong, Australia*

(Received 21 September 1999; revised 5 January 2000; accepted 7 March 2000)

This study investigated whether listeners can use interaural time differences (ITDs) in the amplitude envelope to localize high-frequency sounds in a free field. Localization accuracy was measured for high-frequency (7 to 14 kHz) noise with and without an imposed amplitude modulation (AM) at 20, 80 or 320 Hz. Only AM at 320 Hz led to more accurate localization relative to the nonmodulated condition. The results of a control experiment suggest that the improvement in localization accuracy was due to an increase in stimulus bandwidth, rather than the temporal cues provided by the modulation. © 2000 Acoustical Society of America. [S0001-4966(00)03306-3]

PACS numbers: 43.66.Qp [DWG]

## INTRODUCTION

Early accounts of human sound localization proposed that horizontal plane localization is mediated by two sets of cues. Low-frequency sounds are localized on the basis of interaural time differences (ITDs), whereas high-frequency sounds (above about 1.5 kHz) are localized on the basis of interaural intensity differences (IIDs; e.g., Rayleigh, 1907). This simple account of sound localization has long been regarded as incomplete. Mills (1972), for example, noted the availability at all frequencies of interaural time of onset and offset differences. More recent studies have explored the utility of ongoing ITD cues in the envelopes of complex, low- and high-frequency sounds (e.g., Amenta *et al.*, 1987; Bernstein and Trahiotis, 1994; Henning, 1974a, b, 1980; McFadden and Pasanen, 1976; Nuetzel and Hafter, 1976; Rule, 1994). These latter studies have shown that listeners can use interaural envelope delay to lateralize headphone-presented complex sounds, with the extent of laterality increasing with increasing ITD (Hebrank and Wright, 1974; Henning, 1974a; McFadden and Pasanen, 1976; Nuetzel and Hafter, 1976). Although these delays can be used to lateralize both low- and high-frequency sounds, ITDs in the fine structure dominate at low frequencies (e.g., Bernstein and Trahiotis, 1985a,b; Trahiotis and Bernstein, 1986; Trahiotis and Stern, 1989).

The above findings suggest that ITDs in the sound's envelope can provide listeners with an indication of horizontal location. However, it is not clear whether the additional information provided by these ITDs allows more accurate localization in the free field than can be achieved on the basis of other cues alone. This is because IIDs and monaural spectral cues covary with ITDs in the free field. We examined the utility of envelope ITD cues for free-field localization by comparing localization accuracy for amplitude-modulated and nonmodulated high-frequency noise. Amplitude modulation (AM) rates from 20 to 320 Hz were studied. Sensitivity to AM is known to be highest for modulation rates up to

approximately 20 Hz (Viemeister, 1979), whereas high-frequency carriers are best lateralized for AM rates up to several hundred Hertz (Henning, 1980; Bernstein and Trahiotis, 1985b).

Amplitude modulation is known to result in the generation of side-bands which increase the bandwidth of the stimulus, and previous studies (e.g., King and Oldfield, 1997) have shown that increases in stimulus bandwidth lead to an improvement in localization. Therefore, we conducted a control experiment where the localization of an amplitude-modulated stimulus was compared with that of a nonmodulated stimulus which was matched for bandwidth.

## I. METHODS

### A. Participants

Five males and one female aged between 26 and 40 years participated in the study. All had normal hearing (thresholds within one standard deviation of age-related norms) as assessed using a two-alternative forced-choice procedure for pure-tones of 1, 2, 4, 8, 12, and 16 kHz. Two of the participants were co-authors of this paper. In addition, all participants were able to localize broadband (0.05–20 kHz) noise with an average error of less than 16 degrees.

### B. Apparatus and stimuli

The stimuli were band-pass high-frequency (7 to 14 kHz) noises with a duration of 328 ms (including 20-ms cosine-squared rise and fall times). Carriers were calculated in the frequency domain and had a rectangular spectral profile. In addition to a nonmodulated condition, each noise was 100% sinusoidally amplitude modulated at 20, 80 and 320 Hz. A new stimulus was generated for each trial, converted to an analog signal at 50 kHz by a 16-bit analog to digital converter (Tucker-Davis Technologies System II), amplified (Hafler Pro 1200) and presented through a tweeter (Bose, Freespace) at 60 dB (A-weighted). Sound pressure level was measured using a measuring amplifier (Brüel & Kjaer 2209)

<sup>a)</sup>Electronic mail: ken.mcanally@dsto.defence.gov.au

and microphone (Brüel & Kjær 4145). Acoustic distortion products were at least 50 dB below the primaries.

As AM results in an increase in the bandwidth of the signal, a control experiment was performed in which the localization of high-frequency noise, amplitude modulated at 320 Hz, was compared with a nonmodulated noise, the bandwidth and spectral profile of which were equivalent to those of the modulated noise (6.68 to 14.32 kHz).

The experiment was conducted in an anechoic chamber within which background noise was less than 10 dB sound pressure level in all 1/3-octave bands with center frequencies from 0.5 to 16 kHz. Participants were seated on a stool with their heads at the center of rotation of a remotely controlled, 1-m-radius hoop which allowed the speaker to be located at any azimuth and from  $-50$  to  $+80$  degrees elevation. The participant's view of the hoop and speaker was obscured by an acoustically transparent cloth sphere. Enough light was provided to allow visual orientation within the cloth sphere.

Each participant wore a head-mounted laser pointer, which was used to indicate the perceived direction of the sound source on the inside surface of the cloth tent. The orientation and location of the laser pointer were measured with a magnetic tracker (Polhemus, 3Space Fastrak). The participant indicated when they judged the laser pointer to be directed toward the perceived source location by pressing a button on a hand-held control box.

### C. Procedure

Each experimental session consisted of a block of 42 trials. A pseudo-random procedure was followed in allocating speaker positions where the part-sphere from  $-47.6$  to  $+80.0$  degrees elevation and 360 degrees of azimuth was divided into 42 segments of equal surface area. For each trial, a segment was selected randomly without replacement. The speaker was then randomly positioned within the segment. This resulted in a reasonably even spread of source locations within each session.

At the beginning of each trial, participants oriented towards a light emitting diode (LED) located at zero degrees of azimuth and elevation. The speaker was moved first to a randomly chosen location before it was moved to the target location in order to prevent participants estimating the final speaker position from the duration of speaker movement. Participants were instructed to hold their heads still during the presentation of a single stimulus. Following stimulus presentation, participants indicated the perceived source location. Feedback was then provided by illuminating a speaker-mounted LED which was visible through the cloth.

A repeated-measures design was used in which each of the six participants was involved in the four conditions of AM. For each condition, participants completed two blocks of 42 trials. The order of conditions was random for each participant. For the control experiment, there were two conditions (high-frequency noise modulated at 320 Hz and non-modulated high-frequency noise with matching bandwidth). Again, participants completed two blocks of 42 trials for each condition. The order of conditions was counterbalanced across participants.

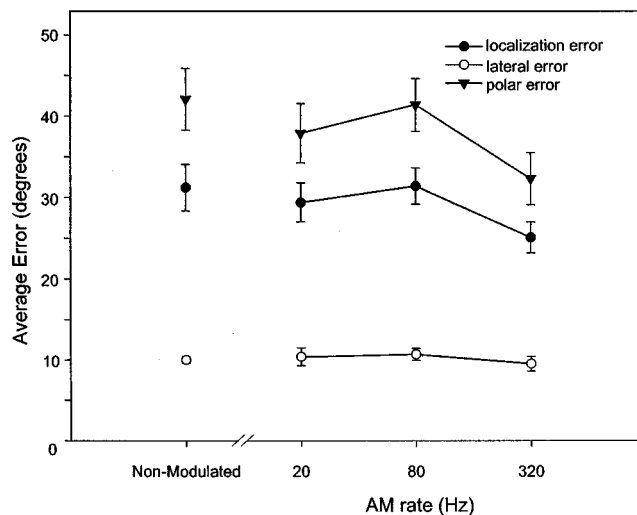


FIG. 1. Average localization error, lateral error and polar error for high-frequency (7–14 kHz) noise across conditions of AM. Error bars indicate one standard error of the mean.

### D. Data collection and analysis

Localization error was defined as the angle between the two vectors from the center of the sphere of rotation of the hoop to the actual and perceived sound source locations. Using a spherical model of the head, an ITD or IID indicates a surface of possible source locations referred to as the cone of confusion. Such cones are described by a constant angle from the median plane, defined as the lateral angle. Normal to the lateral angle is the polar angle which specifies the position around the cone of confusion (Middlebrooks, 1999). These two angles describe a single-pole coordinate system, as do azimuth and elevation, but with the pole oriented in alignment with the interaural axis. Localization error was partitioned into lateral and polar error components. Lateral error was defined as the difference between the lateral angles of the actual and perceived source locations and polar error as the difference between the polar angles (subtended from the frontal horizon) of the actual and perceived source locations (Middlebrooks, 1999).

Localization errors, lateral errors and polar errors were each analyzed using a repeated-measures analysis of variance (ANOVA) for the main and control experiments. Greenhouse–Geisser correction was applied to all tests of significance where appropriate (i.e.,  $df > 1$ ; Keppel, 1991). An alpha level of 0.05 was used for all tests of significance.

## II. RESULTS

Average localization, lateral and polar errors are presented for each condition in Fig. 1. Average localization error varied significantly across conditions [ $F(1.4,6.7) = 5.45$ ,  $p = 0.047$ ]. Planned comparisons revealed that the only AM condition for which performance was significantly different from that for the nonmodulated condition was 320 Hz [ $F(1,5) = 8.441$ ,  $p = 0.034$ ]. The average localization error for this condition was lower than that for any other condition. Average lateral errors did not differ significantly across conditions and planned comparisons between each of the AM conditions and the nonmodulated condition were not

TABLE I. Average localization, lateral and polar errors across conditions of the control experiment. Standard errors are shown in parentheses.

Stimulus	<i>N</i>	Localization error (degrees)	Lateral error (degrees)	Polar error (degrees)
Nonmodulated	6	27.6 (2.6)	10.2 (1.1)	36.6 (5.2)
320 Hz AM	6	26.5 (1.5)	9.8 (1.3)	33.4 (2.1)

significant. Average polar errors also did not differ significantly across conditions. However, planned comparisons between each of the AM conditions and the nonmodulated condition revealed a significantly lower average polar error for the 320-Hz AM condition [ $F(1,5) = 9.454, p = 0.028$ ].

The results of the control experiment are shown in Table I. The difference between the 320-Hz modulated and nonmodulated with matched bandwidth conditions was not significant for the average localization error, the average lateral-angle error or the average polar error.

### III. DISCUSSION

Lateralization studies have demonstrated the ability of listeners to discriminate ITDs in the amplitude envelopes of ongoing stimuli for both amplitude-modulated tones (Bernstein and Trahiotis, 1985a, b; Henning, 1974a, 1980; Nuetzel and Hafter, 1976; Rule, 1994) and band-limited noise (Henning, 1974b; McFadden and Pasanen, 1976; Trahiotis and Bernstein, 1986; Trahiotis and Stern, 1989). The present study investigated the utility of envelope ITD cues for free-field localization to determine whether they allow more accurate localization than can be achieved on the basis of other cues alone.

There are two possible mechanisms by which envelope ITDs could improve localization of high-frequency sound beyond that possible using other cues. It is possible that envelope ITDs specify the appropriate cone of confusion more accurately than IIDs. If this is the case, the ability to determine lateral angle for high-frequency noise would be expected to improve with AM. However, the results of the present study indicate that lateral error did not decrease with AM. Alternatively, envelope ITDs may help resolve ambiguity inherent in IIDs provided by the fine structure of high-frequency stimuli. Recordings made from real heads show that iso-ITD and -IID contours are not identical (Wightman and Kistler, 1999). A given ITD and IID may together specify a unique point at the intersection of the respective iso-ITD and -IID contours. If this is the case, polar error would be expected to decrease with AM. This study has shown that AM at 320 Hz significantly decreased polar errors for high-frequency noise. Although this supports this second hypothesis, the control experiment showed that an identical increase in localization accuracy could be achieved by increasing the bandwidth of the nonmodulated stimulus to equal that of the amplitude-modulated stimulus. This result suggests that the increase in localization accuracy for the amplitude-modulated stimulus in the main experiment was due to the increase in bandwidth associated with AM rather than the temporal cues provided by the modulation.

It could be argued that AM did not lead to an improvement in localization accuracy because there were sufficient envelope ITD cues present in the unmodulated noise. The energy passing through auditory filters tuned to high frequencies would fluctuate, even when stimulated by unmodulated noise. However, envelope fluctuations for unmodulated noise would have been rapid (around 770 Hz for the filter tuned to 12 kHz) and relatively shallow (Middlebrooks and Green, 1990). Therefore, useable envelope ITDs would have been far more prevalent for the modulated conditions than for the unmodulated condition. The limited utility of ITDs in nonmodulated high-frequency noise is supported by the dominance of IIDs in the localization of high-pass noises when ITDs and IIDs conflict (Wightman and Kistler, 1992).

In summary, the results of the present study show that ITDs provided by AM of the envelope of high-frequency stimuli do not allow increased accuracy of localization in the free field above that possible using other cues.

- Amenta, C. A., Trahiotis, C., Bernstein, L. R., and Nuetzel, J. M. (1987). "Some physical and psychological effects produced by selective delays of the envelope of narrow bands of noise," *Hear. Res.* **29**, 147–161.
- Bernstein, L. R., and Trahiotis, C. (1985a). "Lateralization of low-frequency, complex waveforms: The use of envelope-based temporal disparities," *J. Acoust. Soc. Am.* **77**, 1868–1880.
- Bernstein, L. R., and Trahiotis, C. (1985b). "Lateralization of sinusoidally amplitude-modulated tones: Effects of spectral locus and temporal variation," *J. Acoust. Soc. Am.* **78**, 514–523.
- Bernstein, L. R., and Trahiotis, C. (1994). "Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise," *J. Acoust. Soc. Am.* **95**, 3561–3567.
- Hebrank, J., and Wright, D. (1974). "Are two ears necessary for localization of sound sources on the median plane?" *J. Acoust. Soc. Am.* **56**, 935–938.
- Henning, G. B. (1974a). "Detectability of interaural delay in high-frequency complex waveforms," *J. Acoust. Soc. Am.* **55**, 84–90.
- Henning, G. B. (1974b). "Lateralization and the binaural masking-level difference," *J. Acoust. Soc. Am.* **55**, 1259–1262.
- Henning, G. B. (1980). "Some observations on the lateralization of complex waveforms," *J. Acoust. Soc. Am.* **68**, 446–454.
- Keppel, G. (1991). *Design and Analysis. A Researcher's Handbook* (Prentice-Hall, Englewood Cliffs, NJ).
- King, R. B., and Oldfield, S. R. (1997). "The impact of signal bandwidth on auditory localization: Implications for the design of three-dimensional audio displays," *Hum. Factors* **39**, 287–295.
- McFadden, D., and Pasanen, E. G. (1976). "Lateralization at high frequencies based on interaural time differences," *J. Acoust. Soc. Am.* **59**, 634–639.
- Middlebrooks, J. C. (1999). "Virtual localization improved by scaling non-individualized external-ear transfer functions in frequency," *J. Acoust. Soc. Am.* **106**, 1493–1510.
- Middlebrooks, J. C., and Green, D. M. (1990). "Directional dependence of interaural envelope delays," *J. Acoust. Soc. Am.* **87**, 2149–2162.
- Mills, A. W. (1972). "Auditory localization," in *Foundations of Modern Auditory Theory*, edited by J. W. Tobias (Academic, New York), pp. 303–348.
- Nuetzel, J. M., and Hafter, E. R. (1976). "Lateralization of complex waveforms: Effects of fine structure, amplitude, and duration," *J. Acoust. Soc. Am.* **60**, 1339–1346.
- Rayleigh, Third Baron (1907). "On our perception of sound direction," *Philos. Mag.* **13**, 214–232.
- Rule, S. J. (1994). "Lateralization of high frequency, amplitude modulated sounds as a function of carrier frequency," *Can. J. Exp. Psychol.* **48**, 435–443.
- Trahiotis, C., and Bernstein, L. R. (1986). "Lateralization of bands of noise and sinusoidally amplitude-modulated tones: Effects of spectral locus and bandwidth," *J. Acoust. Soc. Am.* **79**, 1950–1957.
- Trahiotis, C., and Stern, R. M. (1989). "Lateralization of bands of noise:

- Effects of bandwidth and differences of interaural time and phase," J. Acoust. Soc. Am. **86**, 1285–1293.
- Viemeister, N. F. (1979). "Temporal modulation transfer functions based on modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.
- Wightman, F. L., and Kistler, D. J. (1992). "The dominant role of low-frequency interaural time differences in sound localization," J. Acoust. Soc. Am. **91**, 1648–1661.
- Wightman, F. L., and Kistler, D. J. (1999). "Resolution of front-back ambiguity in spatial hearing by listener and source movement," J. Acoust. Soc. Am. **105**, 2841–2853.

# Individual differences in speech adaptation to an artificial palate

Shari R. Baum<sup>a)</sup>

*School of Communication Sciences & Disorders, McGill University, 1266 Pine Avenue West, Montréal, Québec H3G 1A8, Canada*

David H. McFarland

*École d'orthophonie et d'audiologie, Université de Montréal, CP 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7, Canada*

(Received 15 November 1999; accepted for publication 23 February 2000)

This preliminary investigation examined the ability of individual speakers to adapt to a structural perturbation to the oral environment in the production of [s]. In particular, the experiment explored whether previous evidence of relatively quick adaptation subsequent to intensive practice would be replicated, whether vowel environment would influence the degree of adaptation, whether adaptive strategies would carry over to normal productions and/or similar sounds (i.e., cause negative aftereffects), and whether adaptive strategies developed during the practice phase could be recalled 1 h later. Results of acoustic and perceptual analyses generally revealed improvement after practice, few consistent effects of vowel context, few negative aftereffects, and an absence of quick recall of adaptive strategies. Moreover, extensive individual differences were found in both the degree of initial perturbation and the extent of adaptation. Implications of the results for issues in speech adaptation are briefly discussed. © 2000 Acoustical Society of America. [S0001-4966(00)01406-5]

PACS numbers: 43.70.Aj [AL]

## INTRODUCTION

In a recent series of investigations, we have been studying the adaptive responses of the speech production system to functional and structural modifications of the oral cavity (Baum and McFarland, 1997; Baum, McFarland, and Diab, 1996; McFarland and Baum, 1995; McFarland *et al.*, 1996). Of particular relevance, we examined whether adaptation to an artificial palate with a thicker than normal alveolar region in [s] production could occur in a relatively brief period, given intensive, focused practice (Baum and McFarland, 1997). The data revealed a gradual improvement in [s] production with the palate in place, as reflected in fricative spectra as well as quality ratings by phonetically trained listeners. Somewhat surprisingly, changes were also found between [s] productions in normal conditions pre- and post-practice, suggesting potential negative “aftereffects” (Anstis, 1995).

One of the hallmarks of previous investigations of adaptation to perturbations has been the apparent individual variability in compensatory abilities (see also Flege *et al.*, 1988; Savariaux *et al.*, 1995, 1999). For example, in the investigation just described, three of the seven subjects appeared to have adapted much more completely than the other speakers after the 1-h practice period (Baum and McFarland, 1997). Other authors have suggested that individual speakers may differ in their overall articulatory “skill” and their ability to modify articulatory dimensions in response to alterations in vocal tract configuration (Savariaux *et al.*, 1995, 1999).

The present investigation, with three main objectives, was designed to provide a preliminary within-subject analysis of the effects of intensive practice on the development of speech adaptation to the presence of an artificial palate. One

goal was to examine whether the ability to adapt to a perturbation in fricative production is influenced by vowel environment. One might predict that certain vowel contexts may impede adaptive processes because of large or contrasting articulatory demands of the vowel and consonant. In contrast, if one adopts the hypothesis that, despite coarticulatory influences, invariant acoustic properties characterize phonetic segments, one might predict that vowel context should not affect adaptation because speakers are targeting the same main acoustic goal regardless of vowel environment. A second goal was to explore whether intensive practice will also affect other speech sounds (produced under normal conditions), particularly those with similar production characteristics, such as [ʃ], due to the development of a distributed and adaptive mode of articulatory programming (Baum and McFarland, 1997). The third primary objective was to assess speakers’ ability to accommodate without further practice to a previously adapted perturbation and the degree to which adaptive articulatory strategies are automatized, much like normal articulatory programs (see Hamlet and Stone, 1978; Hamlet *et al.*, 1978).

## I. METHODS

### A. Acoustic analyses

1. *Subjects.* Four adult female native speakers of (Quebec) French served as subjects, none of whom reported a history of speech, language, or hearing deficits. Three of the four (Speakers 2–4) had a history of orthodonture, with two (Speakers 3 and 4) wearing removable appliances for 1 year or less.

2. *Stimuli and procedures.* A specially designed artificial palate was constructed for each speaker with a 6-mm build-up of acrylic at the alveolar ridge to perturb [s] pro-

<sup>a)</sup>Electronic mail: c3cr@musica.mcgill.ca

duction. Details of the palate design are described elsewhere (Baum and McFarland, 1997). Speech production was sampled at three time intervals: immediately upon insertion of the palate (time 0), after 60 min of [s]-intensive practice (see below) with the palate in place (time 1), and 60 min later (time 2) after normal speech activity without the palate. At time 0, ten repetitions of the syllables [si sa su šī šā šū] were elicited prior to insertion of the artificial palate, comprising a normal baseline condition. Once the palate was inserted, an additional ten repetitions of [si sa su] were elicited for the time 0 perturbed condition. At time 1, another set of ten repetitions of [si sa su] was recorded with the palate in place to determine whether the 60 min of practice had influenced [s] production. A second unperturbed set of stimuli (both [s] and [š]) was also elicited to examine any potential aftereffects or carryover of altered articulatory patterns. Finally, at time 2, the same set of recordings was made—i.e., ten repetitions of [si sa su] with the palate in place, and ten repetitions of [si sa su šī šā šū] under normal conditions. (The normal condition was recorded first at this time interval.)

During the 60-min practice interval, subjects read [s]-laden passages aloud to facilitate adaptation to the palatal perturbation. Target stimuli were recorded on DAT tape using a Sony DTC-57ES recorder and head-mounted directional microphone (AKG-HD421U) which ensured a constant mic-to-mouth distance.

3. *Analyses.* All target utterances were digitized at a rate of 20 K samples/s with a 9-kHz low-pass filter and 12-bit quantization using the BLISS speech analysis system (Mertus, 1989). Centroid frequencies were computed at the midpoint of the frication noise in each stimulus as an indication of degree of compensation to the palatal perturbation (see also Baum and McFarland, 1997; McFarland and Baum, 1995; McFarland *et al.*, 1996).<sup>1</sup>

## B. Perceptual analyses

1. *Subjects.* Ten native French-speaking adults with training in articulatory assessment served as raters in the perception experiment. None of the listeners had served as a speaker in the experiment.

2. *Stimuli and procedures.* For each individual speaker, a perceptual test was constructed consisting of all of the /sV/ tokens in both perturbed and unperturbed conditions, for a total of 180 stimuli per speaker. Ten practice training trials preceded each test: 5 “high-quality” exemplars and 5 “poor” exemplars chosen from the normal and perturbed conditions at time 0, respectively. Those stimuli with centroid frequencies furthest from normal served as examples of poor-quality productions and listeners were informed about the defined quality of the example stimuli. The perception tests were presented in four separate sessions with order of presentation counterbalanced across subjects.

Stimuli were presented to listeners over closed headphones (with a high-frequency limit of 22 kHz) with a 5-s ISI, during which listeners were required to rate the quality of the fricative on a 10-cm visual analog scale based on the examples and their articulatory assessment experience. The

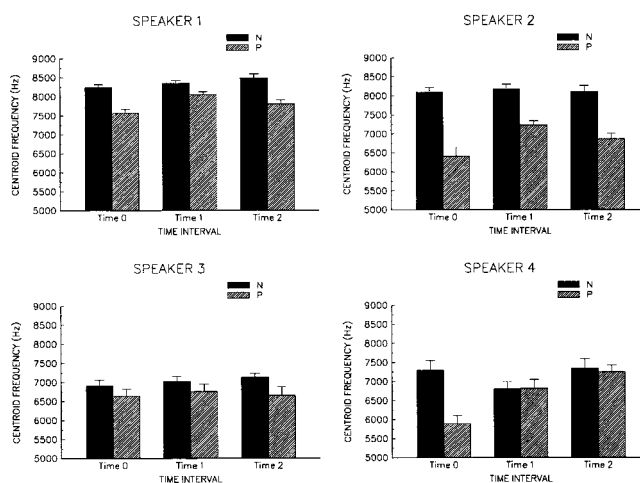


FIG. 1. Mean centroid frequencies (and standard errors) for [s] in palate and normal conditions across time intervals for each speaker.

endpoints of the scale were labeled “unintelligible” and “perfect” and listeners were asked to make a mark at an appropriate point along the scale.

## II. RESULTS

Mean centroid frequencies for [s] computed at each time interval in both perturbed and unperturbed conditions are displayed in Fig. 1(a)–(d) for each speaker.<sup>2</sup> Separate time interval (time 0, time 1, time 2) × palatal condition (normal, perturbed) × vowel ([i a u]) analyses of variance (ANOVA) were carried out for each speaker individually. The ANOVA for Speaker 1 revealed main effects for time interval [ $F(2,18) = 5.268, p < 0.05$ ], palatal condition [ $F(1,9) = 85.166, p < 0.001$ ], and vowel [ $F(2,18) = 6.262, p < 0.01$ ], as well as interactions of time × condition [ $F(2,18) = 3.473, p = 0.05$ ] and time × vowel [ $F(4,36) = 2.819, p < 0.05$ ]. *Post hoc* analyses of the time × condition interaction using the Newman–Keuls procedure ( $p < 0.05$ ) revealed significantly lower centroids in the perturbed conditions at time 0 and time 2. At time 1, the centroids in the normal and perturbed conditions did not differ significantly, suggesting some adaptation had occurred after the practice period.<sup>3</sup> Within the perturbed condition, only centroids at time 0 and time 1 differed significantly, confirming the compensation just noted. Because no three-way interaction emerged, the vowel-related effects were not explored further.

The ANOVA for Speaker 2 yielded significant main effects and interactions of all variables. Of particular interest was the time × condition × vowel interaction [ $F(4,36) = 2.639, p = 0.05$ ], which was subjected to *post hoc* analysis as described above. These tests revealed significant differences in centroid values computed in the normal and perturbed conditions for all vowels at all time intervals, with the exception of [su] stimuli at time 1 (for which a trend emerged). For this speaker, despite apparent improvements in adaptation over the practice interval, significant perturbation remained.

The ANOVA computed on Speaker 3’s data revealed main effects of palatal condition [ $F(1,9) = 21.955, p < 0.001$ ] and vowel [ $F(2,18) = 352.382, p < 0.001$ ] and interactions of condition × vowel [ $F(2,18) = 24.747, p < 0.001$ ]



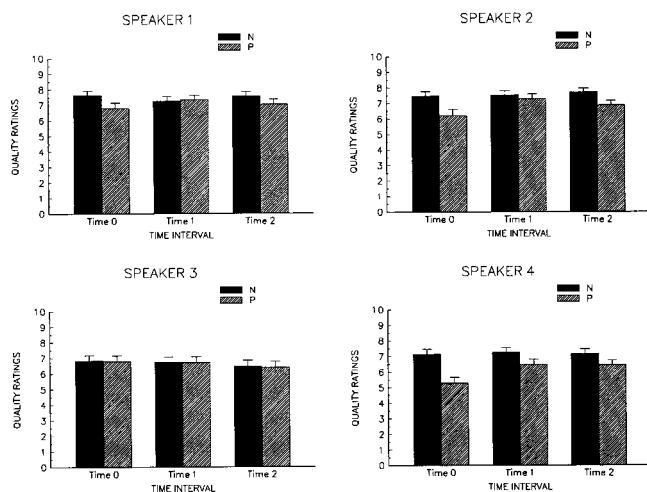


FIG. 2. Mean quality ratings (and standard errors) for [s] in palate and normal conditions across time intervals for each speaker.

and time $\times$ condition $\times$ vowel [ $F(4,26) = 3.23, p < 0.05$ ]. *Post hoc* analysis of the three-way interaction yielded only a few significant comparisons: at each time interval, centroids for [su] stimuli only were significantly lower in the perturbed relative to the normal condition. Thus, even immediately after insertion of the artificial palate, this speaker's [s] articulation was not greatly affected.

Finally, like Speaker 2, the ANOVA for Speaker 4 yielded three significant main effects and numerous interactions. *Post hoc* testing of the time $\times$ condition $\times$ vowel interaction [ $F(4,36) = 4.540, p = 0.005$ ] revealed significant effects of the perturbation at time 0 for all vowels, no significant differences at time 1 (except for [si] stimuli, which displayed a difference in centroid frequencies in the unexpected direction), and differences between perturbed and normal conditions at time 2 for all but [sa] stimuli. For this speaker, within the normal condition, centroids for [si] and [sa] stimuli were lower at time 1 as compared to time 0 and time 2, suggesting an influence of the palate on normal [s] production. Within the perturbed condition, centroids were higher at time 1 relative to time 0 across vowels and remained high at time 2.

Mean centroid frequencies were also calculated at each time interval for normal (unperturbed) [ʃ] stimuli to determine whether the altered articulatory patterns developed in compensation for the palate would affect nonalveolar sounds. Individual time interval $\times$ vowel ANOVAs conducted on [ʃ] centroid frequencies revealed few significant differences of interest for any of the speakers, despite apparent variability in mean centroid frequencies of up to several hundred Hz across the time intervals. (Details are available upon request.)

Individual ratings along the visual analog scale on the perception tests for each speaker were computed in cm and averaged across the ten repetitions of each target stimulus. These values (displayed in Fig. 2) were submitted to separate time interval $\times$ palatal condition $\times$ vowel ANOVAs for each speaker. The ANOVA for Speaker 1 revealed no significant main effects or interactions. For Speaker 2, all main effects and interactions were significant. Of particular interest was the three-way interaction of time $\times$ condition $\times$ vowel

[ $F(4,40) = 9.043, p < 0.001$ ], which was further analyzed using the Newman–Keuls procedure. Initial pairwise comparisons focused on the quality ratings in the normal versus the palatal conditions at each time interval. At time 0, ratings were significantly lower in the palate condition relative to the normal condition for all but [sa] stimuli. At time 1, only the [si] stimuli in the palate condition were rated significantly more poorly than in the normal condition. Finally, at time 2, ratings for the stimuli produced under normal and palate conditions differed for all but the [su] stimuli. Comparisons were also made across the time intervals within conditions. In the normal condition, the ratings for [si] stimuli at time 0 were significantly lower than at times 1 and 2. No other significant differences emerged. In the palate condition, ratings for stimuli produced at time 0 were significantly lower than at both times 1 and 2 for almost all comparisons, with the exception of the [sa] stimuli at time 0 (where no effect of the palatal perturbation was seen). Surprisingly, for [sa] stimuli at time 2, quality ratings were worse than at time 0.

The ANOVA for speaker 3 revealed a main effect of time [ $F(2,16) = 8.324, p < 0.003$ ] and interactions of time $\times$ vowel [ $F(4,32) = 2.826, p < 0.05$ ] and condition $\times$ vowel [ $F(2,16) = 20.567, p < 0.001$ ]. *Post hoc* analyses of the latter interaction revealed significantly lower quality ratings for [su] stimuli in the palate compared to the normal condition, collapsed over time intervals. In contrast, for [si] stimuli, ratings were unexpectedly higher in the palate condition relative to the normal condition, while for [sa] stimuli, no significant differences across condition emerged. For speaker 4, all main effects and interactions were significant with the exception of the condition $\times$ vowel interaction ( $F < 1$ ). *Post hoc* analysis of the three-way interaction [ $F(4,40) = 4.262, p < 0.01$ ] revealed a significant effect of the palatal manipulation on quality ratings at time 0 for all vowels; at time 1, only [si] stimuli were still rated lower in the palate compared to the normal condition. At time 2, ratings for both [si] and [sa] productions were again lower in the perturbed condition. Within the normal condition, no differences emerged across time intervals; however, within the palate condition, for both [sa] and [su], ratings at time 0 were significantly lower than at times 1 and 2.

### III. DISCUSSION

This investigation was designed to provide preliminary data representing detailed within-subject comparisons of the development of speech adaptation to a structural modification of the oral cavity. Emerging from these analyses were sometimes striking differences in individual compensatory responses. For example, evidence of speech adaptation for [s] production after 1 h of intensive practice was observed in the acoustic and perceptual data for Subject 4, in the acoustic data only for Subject 1, and in the perceptual data only for Subject 2 (with a trend towards adaptation observed in the acoustic results). Further, Subject 3 showed essentially no effect of the artificial palate on speech articulation at the first and subsequent measurement intervals. At present, we can only speculate as to some of the many factors that may contribute to these individual differences. We and others (Baum *et al.*, 1996; McFarland *et al.*, 1996; Savariaux *et al.*, 1995,

1999) have suggested that individual subjects may differ in their ability to integrate altered sensory feedback in the modification of articulatory gestures appropriate to the change in oral form. An additional, and related, possibility is that individual differences in oral form or articulatory placement may have rendered the artificial palate more or less perturbing to [s] production. However, consistent with our previous findings (McFarland *et al.*, 1996), no consistent relationship was observed between measures of palatal dimensions and individual responses to articulatory perturbation.<sup>4</sup> This does not rule out the possibility that articulatory positioning for [s] production may have influenced the speech perturbation of the artificial palate. For example, certain subjects may normally adopt a more posterior positioning of the tongue for [s] outside of the perturbed “zone” (Savariaux *et al.*, 1999).<sup>5</sup>

Individual differences were also observed in the extent to which speech adaptation to perturbation subsequent to an hour of intensive practice influenced nonperturbed articulatory planning. In fact, changes in [s] production in the normal conditions pre- and post-practice, suggesting negative “aftereffects,” were observed consistently in the acoustic data of only one experimental subject. In our previous investigation we found that five of the seven subjects exhibited negative aftereffects using a criterion of a 1000-Hz decrease in normal centroid frequencies after the practice period (see Table II, Baum and McFarland, 1997). Based on the perceptual data, three of the seven subjects exhibited such aftereffects. These data suggest that there are individual differences in the distributed influence of the development of adapted articulation to global articulatory programming. Similarly, intensive practice intended to facilitate [s] production under conditions of perturbation had little effect on other speech sounds with similar production characteristics. No consistent effects of the practice interval were observed on [ʃ] production, suggesting that the altered articulatory processes associated with speech adaptation were specific to the highly practiced sounds. Our data also suggest that speakers targeted a specific consonantal acoustic goal under conditions of perturbation regardless of vowel context, as no systematic effect of vowel environment on the adaptive response to the oral articulatory perturbation was found.

Finally, the present results suggest that adapted speech articulation resulting from intensive target-specific practice results in compensatory responses that are fragile and not easily recalled 1 h after the practice interval. This is in contrast to an earlier investigation of speech adaptation using an artificial palate very similar to that used in the current study (Hamlet *et al.*, 1978) with the exception that there was a 4-mm build-up of acrylic as contrasted to 6 mm in the present work. In that investigation, however, subjects were provided with a much longer period of adaptation during which they wore the artificial palate for 2 weeks except while sleeping. After the 2-week adaptation period, subjects were tested again and phonetic/perceptual judgments revealed

more normal consonant production after only the third repetition of the test utterance, suggesting a rapid retrieval of the modified articulatory strategies. Although the period of intensive practice used in our previous and current investigation appears to accelerate the adaptation process, longer exposure to the presence of the change in oral form may be necessary to stabilize the modified articulatory programs.

## ACKNOWLEDGMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

<sup>1</sup>Although there are obviously good reasons to compute both temporal measures and additional spectral moments (e.g., skewness and kurtosis) in providing an accurate characterization of fricatives, we limited the reported measures to centroid frequencies for several reasons. First, our previous data did not reveal differences across conditions in either temporal or other spectral measures. Second, our primary goal was a comparison of perturbed and unperturbed conditions, not specifically the most comprehensive spectral characterization of fricatives.

<sup>2</sup>It should be noted that some of the speakers’ productions yielded high centroid values, approaching the cutoff frequency of the filter. It is, therefore, possible that the spectra are not fully representative of the [s] productions. However, because all comparisons are within-subject, across perturbation conditions, this was unlikely to have affected the specific questions under investigation.

<sup>3</sup>One must, of course, always be cautious about interpreting a null result. However, given the significant differences which were found at other time intervals, the question of sufficient statistical power does not appear to be an issue.

<sup>4</sup>Interestingly, speakers’ history of orthodonture also showed no consistent influence on adaptation.

<sup>5</sup>It is interesting to note that the acoustic and perceptual findings, although largely consistent with one another, were not exactly parallel. Such a pattern is in keeping with previous investigations (e.g., Savariaux *et al.*, 1999) and emphasizes the need for multiple, complementary measures in assessing articulatory compensation.

Anstis, S. (1995). “Aftereffects from jogging,” *Exp. Brain Res.* **103**, 476–478.

Baum, S., and McFarland, D. (1997). “The development of speech adaptation to an artificial palate,” *J. Acoust. Soc. Am.* **102**, 2353–2359.

Baum, S., McFarland, D., and Diab, M. (1996). “Compensation to articulatory perturbation: Perceptual data,” *J. Acoust. Soc. Am.* **99**, 3791–3794.

Flege, J., Fletcher, S., and Homiedan, A. (1988). “Compensating for a bite block in /s/ and /t/ production: Palatographic, acoustic, and perceptual data,” *J. Acoust. Soc. Am.* **83**, 212–228.

Hamlet, S., and Stone, M. (1978). “Compensatory alveolar production induced by wearing a dental prosthesis,” *J. Phonetics* **6**, 227–248.

Hamlet, S., Stone, M., and McCarty, T. (1978). “Conditioning prostheses viewed from the standpoint of speech adaptation,” *J. Prosthet. Dent.* **40**, 60–66.

McFarland, D., and Baum, S. (1995). “Incomplete compensation to articulatory perturbation,” *J. Acoust. Soc. Am.* **97**, 1865–1873.

McFarland, D., Baum, S., and Chabot, C. (1996). “Speech compensation to structural modifications of the oral cavity,” *J. Acoust. Soc. Am.* **100**, 1093–1104.

Mertus, J. (1989). *BLISS User’s Manual* (Brown University, Providence).

Savariaux, C., Perrier, P., and Orliaguet, J. (1995). “Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production,” *J. Acoust. Soc. Am.* **98**, 2428–2442.

Savariaux, C., Perrier, P., Orliaguet, J.-P., and Schwartz, J.-L. (1999). “Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis,” *J. Acoust. Soc. Am.* **106**, 381–393.

# An analytical error invalidates the “depolarization” of the perceptual magnet effect

Frank H. Guenther<sup>a)</sup>

Department of Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, Massachusetts 02215 and Speech Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 1 July 1999; revised 17 January 2000; accepted 28 January 2000)

In a recent article, Lotto *et al.* [J. Acoust. Soc. Am. **103**, 3648–3655 (1998)] presented experiments investigating the role played by perceived phonemic identity in demonstrations of decreased discriminability for prototypical vowel sounds. The authors interpreted their results as evidence against shrinkage of perceptual space near vowel category prototypes. In this letter, it is shown that this interpretation is based on a flawed data analysis in which a key confounding term has been neglected. © 2000 Acoustical Society of America. [S0001-4966(00)00905-X]

PACS numbers: 43.71.An, 43.71.Es [JMH]

The perceptual magnet effect (PME) reported by Kuhl (1991) has been a hotly debated topic in the speech perception literature in recent years. According to the Kuhl *et al.* account, the PME is a case of perceptual space being “shrunk” near prototypical examples of vowel sounds, as compared to near nonprototypical examples, such that the same spectral difference seems smaller near prototypical vowels than near nonprototypical vowels. In a recent JASA article entitled “Depolarizing the perceptual magnet effect,” Lotto, Kluender, and Holt (1998) attempt to show that the PME is nothing more than a “further demonstration that general discriminability is greater for cross-category stimulus pairs than for within-category pairs” (p. 3648). In other words, Lotto *et al.* posit that PME experimental results can be completely explained by a simple assumption, rooted in classical treatments of categorical perception (CP), without reference to shrunken representations of spectral space near prototypical vowels. The assumption is that differences in the ability to discriminate prototypical versus nonprototypical sounds can be attributed to differences in the probability that two sounds in a discrimination pair are given different phonemic labels by the listener. That is, two nonprototypical vowel sounds are more likely to be given different phonemic labels by the listener, and are thus more likely to be discriminated by the listener, than two prototypical vowel sounds.

In order to assess this claim, Lotto *et al.* perform some interesting PME experiments that correct for contextual effects on phonemic identification. After analyzing their experimental results, Lotto *et al.* claim that “Following application of time-worn models of Categorical Perception and classical findings of phonetic perception in context, PME theory makes the wrong predictions” (p. 3653). However, an analytical error in the authors’ analysis invalidates this claim. Before investigating this issue, it is useful to define some terminology. In the following, P will be used to refer to the prototype stimulus (i.e., a “good” example of /i/) and NP to the nonprototype (a “bad” example of /i/). The probability of generalizing two sounds (i.e., responding that they are the same when they are actually different) will be denoted by

$p(\text{gen})$ . Finally, following Lotto *et al.* (1998), the term “PME theory” will be used to refer to the assertion by Kuhl and colleagues that the PME represents a shrinking of perceptual space in the neighborhood of prototypical sounds and that it is not simply the result of “phonemic identity” processes as described above.

Lotto *et al.* compare the difference between the generalization of P and NP sounds as measured in their experiment with a predicted difference in generalization of P and NP sounds obtained under the assumption that consideration of phonemic identity alone accounts for the difference. Quoting Lotto *et al.* (p. 3653):

“... if PME is due to a shrinking of the perceptual space around the prototype, then the differences between predicted generalization scores [predicted under the assumption that phonetic identity alone is responsible] for P and NP conditions should be substantially less than the scores obtained [in their experiment]. For these comparisons, it is the differences in the generalization scores and not the actual scores themselves that are important. Predictions of discriminability using this method understandably underestimate the discriminability of speech sounds (e.g., Miyawaki *et al.*, 1975). If one assumes that other bases for discrimination other than perceived identity (e.g., guessing, spectral difference) are equally potent for the P and NP conditions, then it is appropriate to compare the predicted differences and observed differences.”

The assumption underlying this comparison can be written mathematically as follows:

$$p(\text{gen}) = p(\text{gen by PI}) + p(\text{gen by spect}), \quad (1)$$

where  $p(\text{gen by PI})$  is the probability that the two sounds in a discrimination trial are identified as the same phoneme by the listener (i.e., the probability of generalizing based on phonemic identity), and  $p(\text{gen by spect})$  is the probability that the two sounds cannot be discriminated based on spectral distance or other means. Lotto *et al.* use their identification scores from experiment 1 to predict the difference in

<sup>a)</sup>Electronic mail: guenther@cns.bu.edu

generalization scores for P and NP based on phonemic identity alone:

$$\text{Diff\_PI} = p(\text{gen by PI, P}) - p(\text{gen by PI, NP}), \quad (2)$$

where  $p(\text{gen by PI, P})$  is the probability of generalization based on phonemic identity in the prototype case, etc. The difference predicted by phonemic identity alone is then compared to the measured difference in generalization scores, which, from Eq. (1), is assumed to correspond to the following:

$$\begin{aligned} \text{Diff\_measured} &= p(\text{gen, P}) - p(\text{gen, NP}) \\ &= p(\text{gen by PI, P}) + p(\text{gen by spect, P}) \\ &\quad - [p(\text{gen by PI, NP}) \\ &\quad + p(\text{gen by spect, NP})] \\ &= \text{Diff\_PI} + p(\text{gen by spect, P}) \\ &\quad - p(\text{gen by spect, NP}). \end{aligned} \quad (3)$$

Because the difference as predicted from phonemic identity alone (Diff\_PI, calculated to be 8.57 by Lotto *et al.*) is greater than the difference as determined from the measured generalization scores (Diff\_measured, which was 5.61 in the Lotto *et al.* experiment), the authors conclude that  $p(\text{gen by spect, P})$  must be less than  $p(\text{gen by spect, NP})$ , which is the opposite of what is predicted by PME theory: according to this analysis, it is easier to discriminate, based on spectral difference, sounds near the prototype than sounds near the nonprototype. This prompts the authors to claim (p. 3653) that ‘‘If category goodness plays any role in predicting discriminations it must be in a manner opposite to what has previously been suggested!’’

The following hypothetical example can be used to highlight the problem with the Lotto *et al.* analysis. Assume that, in the NP case, all test pairs fall into different phonemic categories. That is,  $p(\text{gen by PI, NP}) = 0\%$ . Assume further that all test pairs in the P case fall into the same phonemic category, so  $p(\text{gen by PI, P}) = 100\%$ . For this case, Diff\_PI is 100%. Next, assume that PME theory is correct, such that  $p(\text{gen by spect}) = 30\%$  for the NP case and 60% for the P case. That is, perceptual space is warped around the prototype such that the same spectral difference is half as likely to be discriminated near P than near NP. Since all NP stimuli can be discriminated by phonemic identity alone,  $p(\text{gen, NP}) = 0\%$  regardless of  $p(\text{gen by spect, NP})$ , and since no P stimuli can be distinguished by phonemic identity,  $p(\text{gen, P}) = p(\text{gen by spect, P}) = 60\%$ . The value of Diff\_measured as calculated from generalization data following the Lotto *et al.* method thus would be 60%. Since Diff\_PI is 100% and Diff\_measured is only 60%, using the Lotto *et al.* logic one would conclude that there is a very strong antimagnet effect, as Lotto *et al.* conclude from their experimental results. Clearly this is incorrect in our hypothetical example since  $p(\text{gen by spect})$  is twice as large for the P case than the NP case; that is, there is a strong magnet effect, not a strong antimagnet effect.

Although this example represents an extreme case for the sake of illustration, the same problem exists for less ex-

treme cases. The problem lies with the assumption summarized in Eq. (1). Generalization will take place only if a stimulus pair generalizes based on phonemic *and* based on spectral difference. Equation (1) is thus incorrect, and the following equation should have been used instead:

$$\begin{aligned} p(\text{gen}) &= p(\text{gen by PI and gen by spect}) \\ &= p(\text{gen by PI}) + p(\text{gen by spect}) \\ &\quad - p(\text{gen by PI or gen by spect}). \end{aligned} \quad (4)$$

This form of the equation follows from a fundamental axiom of probability theory (e.g., Ross, 1976):

$$p(A \text{ and } B) = p(A) + p(B) - p(A \text{ or } B). \quad (5)$$

The exclusion of the  $p(\text{gen by PI or gen by spect})$  term completely confounds the Lotto *et al.* analysis and conclusions. The Lotto *et al.* comparison of predicted and measured differences in generalization would be valid only if  $p(\text{gen by PI or gen by spect})$  were the same in the P and NP conditions, since in this case they would cancel each other in the equation for Diff\_measured. Given that the Lotto *et al.* experimental results indicate that  $p(\text{gen by PI})$  is 8.57% less for NP than for P, it is not reasonable to assume that this is the case. It can be shown that, for PME theory to be supported by the Lotto *et al.* results,  $p(\text{gen by PI or gen by spect})$  need only be 3% less in the NP case than in the P case. Since there is no way to estimate the values of  $p(\text{gen by PI or gen by spect})$  for the P and NP conditions based on the Lotto *et al.* experimental results (nor the results of any other experiment of which I am aware), the Lotto *et al.* results do not provide evidence against PME theory despite the authors’ claims to the contrary.

Of course, a lack of evidence against PME theory should not be interpreted as evidence supporting PME theory. It appears clear from Kuhl (1991) and followup studies including Lotto *et al.* (1998) that our auditory perceptual space for vowels is warped in the following sense: discriminability as measured by  $d'$  is worse for some parts of acoustic space than other parts, and this appears to have something to do with the locations of vowel categories in the listener’s language. However, the relative roles of labeling processes, category prototypes, and other influences in this warping remains to be clarified.

## ACKNOWLEDGMENT

This work was supported by Grant No. R29 DC02852 from the National Institute on Deafness and other Communication Disorders.

Kuhl, P. K. (1991). ‘‘Human adults and human infants show a ‘perceptual magnet effect’ for the prototypes of speech categories, monkeys do not,’’ *Percept. Psychophys.* **50**, 93–107.

Lotto, A. J., Kluender, K. R., and Holt, L. L. (1998). ‘‘Depolarizing the perceptual magnet effect,’’ *J. Acoust. Soc. Am.* **103**(6), 3648–3655.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimora, 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.

Ross, S. (1976). *A First Course in Probability* (Macmillan, New York).

# Reply to “An analytical error invalidates the ‘depolarization’ of the perceptual magnet effect” [J. Acoust. Soc. Am. 107, 3576–3577 (2000)]

Andrew J. Lotto<sup>a)</sup>

Department of Psychology & Parmlly Hearing Institute, Loyola University Chicago,  
6525 North Sheridan Road, Chicago, Illinois 60626

(Received 24 November 1999; revised 15 December 1999; accepted 28 January 2000)

In a previous paper, Lotto *et al.* [J. Acoust. Soc. Am. **103**, 3648–3655 (1998)] claimed that experiments investigating a perceptual magnet effect (PME) failed to demonstrate a novel effect of goodness judgments on speech sound discrimination due to a confound in the procedure. New analyses of these data and a review of the recent literature fail once again to support the existence of PME. © 2000 Acoustical Society of America. [S0001-4966(00)01105-X]

PACS numbers: 43.71.An, 43.71.Es [JMH]

The perceptual magnet effect (PME) has been offered as a novel perceptual effect in work on the discrimination of speech sounds (e.g., Kuhl, 1991; Iverson and Kuhl, 1995). The data from several experiments have been interpreted as an indication that category “goodness” influences discrimination above and beyond the effects of spectral distance and phonemic identification. For example, Iverson and Kuhl (1995) demonstrated that vowel tokens that were labeled as “good” members of the /i/ category [the **Prototype (P)** condition] were harder to discriminate from neighboring exemplars than were members that were considered poor examples of /i/ [the **Nonprototype (NP)** condition]. Importantly, these discriminations were all considered to be *intracategory* comparisons. Lotto *et al.* (1998) suggested that this experiment contained a confound that made the interpretation in terms of PME problematic. The problem is that identification functions were collected for stimuli presented in isolation and discrimination occurred for stimuli presented in pairs. Lotto *et al.* (1998) demonstrated that the perceived phonemic identities of these stimuli were greatly affected by context in the pair condition and, therefore, that the identification functions obtained in isolation were not good indicators of identity in the discrimination task.

Of greatest theoretical significance was the fact that a shift in vowel identity was more likely to occur for the poor vowel exemplars (i.e., the **NP** condition) than for the good exemplars (i.e., the **P** condition). Thus, discriminations in the **NP** case were more likely to be *intracategorical* and the **P** discriminations were more likely to be *intracategorical*. If this is the case, then the PME is not a novel demonstration. It has long been observed that discriminations between sounds from different phonemic categories are easier than discrimination of sounds from within one speech category (e.g., Liberman *et al.*, 1957; Pisoni, 1971; Wood, 1976).

To see if there was any evidence for an effect of category goodness on discrimination above and beyond the established effects of spectral distance and phonemic identity, Lotto *et al.* examined the *generalization scores* (misses).

This measure had been used by Kuhl in previous demonstrations of PME (Grieser and Kuhl, 1989; Kuhl, 1991; Kuhl *et al.*, 1992). The obtained difference in generalization scores for the **P** and **NP** conditions was compared to the predicted difference in generalization scores based solely on phonemic identification. In making this comparison, it was presumed that effects of spectral difference were equivalent (these were experimentally equated in the two conditions by using the mel scale) and that any joint probabilities were equivalent between conditions. Because there is no way in this experimental setup to estimate the joint probabilities (e.g., the probability of discriminating on the basis of phonemic identification *and* perceptual space [goodness] differences), this was considered the most conservative route. The results demonstrated that the obtained scores did not differ from the predicted scores in a way that would serve as evidence for PME. There was no evidence that the differences between the **P** and **NP** conditions were due to category goodness. The conclusion reached was that there was a lack of positive evidence for PME.

Guenther (2000) points out that if the joint probabilities differ in the **P** and **NP** conditions, then the outcomes of generalization difference scores could tell a different story. He presents a clear and powerful example in which a warped perceptual space could lead to generalization difference scores that incorrectly appear to demonstrate no magnet effect. It should be noted that Lotto *et al.* (1998) were not attempting to disprove the existence of PME with the difference equations, but that they were looking for positive evidence for PME using Kuhl’s generalization score. However, Guenther’s (2000) point is significant for the extensive PME literature. Basically, Guenther’s analysis leads one to conclude that the use of generalization scores in all of the PME work is of dubious merit.

The following example elucidates the problem. Presume that the data demonstrate that the probability of discrimination based on phonemic identity is equal in both the **P** and **NP** condition of a PME experiment, let’s say, 0.50 in either case. Now, presume that the probability of discriminating based on perceptual space distance (warped in relation to goodness) is 0.60 for the **P** condition and 0.40 for the **NP**

<sup>a)</sup>Electronic mail: alotto@luc.edu

condition. That is, the better members of a category are more easily discriminated from neighbors than are poorer members. This is the opposite of PME: it is an anchoring effect. What will the generalization scores show? This depends on the joint probability of discriminating based on phonemic identification *and* perceptual space distance. Let's say that for the **P** condition the joint probability is 0.40 and that for the **NP** condition it is 0.10. Plug these values into Guenther's equation (??):

$$[1 - (0.50 + 0.60) - 0.40] - [1 - (0.50 + 0.40) - 0.10] = 0.10.$$

The difference in generalization scores for this examples looks like PME: generalization scores are higher for the **P** condition than for the **NP** condition. However, the values presented here are for an underlying perceptual space that is warped in the opposite direction. It appears that if one allows the joint probabilities to differ, then generalization scores can be quite deceiving.

Guenther (2000) suggests that this analysis of joint probabilities invalidates the conclusions of Lotto *et al.* (1998) that there is no evidence for PME. However, Guenther's analysis, in fact, supports this conclusion by pointing out that previous empirical work using generalization scores do not provide compelling evidence for PME. Note that this is true for any experiment that allows cross-category pairs in discrimination. If phonemic identity is equivalent for all stimuli, then the joint probabilities are zero. Lotto *et al.* (1998) demonstrated that the discrimination pairs in Iverson and Kuhl (1995) often contained cross-category comparisons. Furthermore, the strong context effect on perceived identity that Lotto *et al.* demonstrated for stimuli in pairs indicates that many of the demonstrations of PME probably have contained cross-category discriminations. Thus, the analysis of Guenther (2000) does not invalidate the conclusions of Lotto *et al.* (1998), but rather the two works taken together provide a challenge to the entire PME enterprise at least in relation to the use of generalization scores.

## I. NEW DATA ANALYSIS

If the generalization score is a dubious measure for testing PME, is there another analysis that will provide information about the effect of goodness ratings on discrimination above and beyond the effects of spectral distance and phonemic identity? One possibility is to model discrimination using multiple regression to see if goodness ratings account for a significant amount of the variance.

To further test for PME, the discrimination data from experiment 2 of Lotto *et al.* (1998) were modeled using stepwise regression with variables representing spectral distance, discrimination predicted from perceived phonemic identity, goodness rating of the standard (P or NP), and average goodness rating of both members of each discrimination pair. In stepwise regression, factors are entered into the model depending on the amount of variance (*R*-squared) accounted for in the dependent variable. A significance test determines which factors get entered into the model with the factors accounting for the most variance being added first. For this analysis, the factors were the following.

- (1) Spectral distance: Euclidean distance between the standard and the comparison stimuli in the mel-scaled first formant×second formant space (30, 60 or 90 mel).
- (2) Probability that both stimuli in pair will have the same perceived phonemic identity (PI): This is the predicted probability of a miss dependent solely on the identification functions for stimuli presented in pairs from Lotto *et al.* (1998) experiment 1.
- (3) Goodness rating for standard: Mean goodness rating for P or NP for the particular pair being discriminated as determined from experiment 1 of Lotto *et al.*
- (4) Average goodness rating for pair: The average goodness rating for the two stimuli in the pair as obtained in experiment 1 of Lotto *et al.*

Evidence for PME would correspond to either of these last two factors accounting for a significant amount of the variance in discrimination scores beyond the variance accounted for by the first two factors.

Two models were calculated: one for the dependent variable of hits (i.e., "different" responses to stimuli that are in fact different) and one for the bias-free measure of sensitivity, *d'* (Green and Swets, 1966). For the model of hits, spectral distance was entered, first having accounted for the most variance [*R*-square=0.79, *F*(1,10)=36.67, *p*<0.0001]. The factor PI was then entered, resulting in a model that accounted for a healthy 91% of the variance in hits [*F*(2,9)=46.41, *p*<0.0001]. Both of the goodness factors failed to contribute significantly to the model (*p*'s >0.67; criterion of selection=0.05).

The model of *d'* data followed the same pattern. Spectral distance was entered into the model first [*R*-square =0.86, *F*(1,0)=59.97, *p*<0.0001] followed by PI [full model, *R*-square=0.94, *F*(2,9)=72.24, *p*<0.0001]. The two goodness-rating factors failed to account for any significant added variance (*p*'s>0.19).

These results offer a compelling lack of evidence for PME. In fact, they demonstrate that spectral distance and phonemic identity do a remarkably good job of predicting discrimination for these vowel stimuli (94% of the variance accounted for in *d'* scores!).

## II. CONCLUSIONS

Taken together, these new regression models, the empirical work of Lotto *et al.* (1998), and the analysis by Guenther (2000) offer a sobering picture of PME. As Lotto *et al.* point out, PME has been one of the more fertile areas of recent speech perception research, spawning numerous empirical, theoretical, and modeling papers. However, the evidence underlying this enterprise is tenuous at best.

Guenther (2000) ends with the statement, "It appears clear from Kuhl (1991) and follow-up studies including Lotto *et al.* (1998) that our perceptual spaces for vowels is warped." But the data as they stand do *not* allow for such a conclusion. The regression models indicate that one can account for discrimination by looking at spectral distance and binary phonemic identity decisions without any additional *intra* category information. In addition, there is a growing literature of empirical work that fails to demonstrate percep-

tual warping for vowel categories when proper controls are taken. For example, Lively and Pisoni (1995) found no difference in P and NP discrimination when individual prototypes were determined for each subject. Renda *et al.* (1995) failed to find PME for the vowel /·/. Kewley-Port and Neel (1998) and Sharma and Dorman (1998) have also demonstrated failures of PME hypotheses. Thus, there does not appear to be sufficient evidence in the literature to back up Guenther's (2000) conclusion<sup>1</sup> and there appears to be little support for the concept of PME.

<sup>1</sup>There *are* models of perceptual space warping associated with categorical perception (e.g., Livingston *et al.*, 1998). This may be a fruitful avenue for the study of categorization effects on perceptual space.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Grieser, D., and Kuhl, P. K. (1989). "Categorization of speech by infants: Support for speech-sound prototypes," *Dev. Psychol.* **25**, 577–588.

Guenther, F. H. (2000). "An analytical error invalidates the 'depolarization' of the perceptual magnet effect," *J. Acoust. Soc. Am.* **107**, 3576–3577.

Harnad, S., Hanson, S. J., and Lubin, J. (1995). "Categorical perception and the evolution of supervised learning in neural nets," in *Proceedings of the AAAI Spring Symposium on Machine Learning of Natural Language and Ontology*, edited by D. W. Powers and L. Uhr (DFKI, Kaiserslautern, Germany), pp. 65–74.

Iverson, P., and Kuhl, P. K. (1995). "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling," *J. Acoust. Soc. Am.* **97**, 553–562.

Kewley-Port, D., and Neel, A. T. (1998). "Relation between discrimination and identification of English vowels," *J. Acoust. Soc. Am.* **103**, 2983.

Kuhl, P. K. (1991). "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not," *Percept. Psychophys.* **50**, 93–107.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). "Linguistic experience alters phonetic perception in infants by 6 months of age," *Science* **255**, 606–608.

Lieberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (1957). "The discrimination of speech sounds within and across phoneme boundaries," *J. Exp. Psychol.* **54**, 358–368.

Lively, S. E., and Pisoni, D. B. (1995). "On prototypes and phonetic categories: A critical assessment of the perceptual magnet effect in speech perception," *J. Exp. Psych.: Hum. Percept. Perform.* **23**, 1665–1679.

Livingston, K. R., Andrews, J. K., and Harnad, S. (1998). "Categorical perception effects induced by categorical learning," *J. Exp. Psych.: Learning Memory and Cognition* **24**, 732–753.

Lotto, A. J., Kluender, K. R., and Holt, L. L. (1998). "Depolarizing the perceptual magnet effect," *J. Acoust. Soc. Am.* **103**, 3648–3655.

Pisoni, D. B. (1971). "On the nature of categorical perception of speech sounds," Doctoral Thesis, University of Michigan.

Renda, S. C., Hawks, J. W., and Klich, R. (1995). "An investigation of the perceptual magnet effect in adults," *J. Acoust. Soc. Am.* **97**, 3420.

Sharma, A., and Dorman, M. F. (1998). "Exploration of the perceptual magnet effect using the mismatch negativity auditory evoked potential," *J. Acoust. Soc. Am.* **104**, 511–517.

Wood, C. C. (1976). "Discriminability, response bias, and phoneme categories in discrimination of voice onset time," *J. Acoust. Soc. Am.* **60**, 1381–1389.

# Frequency beamforming of dolphin whistles using a sparse three-element towed array

Aaron Thode<sup>a)</sup>

Marine Physical Laboratory, Scripps Institution of Oceanography, San Diego, California 92093-0205

Thomas Norris<sup>b)</sup> and Jay Barlow

Southwest Fisheries Science Center, National Oceanographic and Atmospheric Administration, P.O. Box 271, San Diego, California 92038

(Received 9 February 1999; accepted for publication 21 January 2000)

Acoustic bearings are obtained from dolphin whistles using frequency-domain (FD) beamforming techniques on signals recorded on a three-element 9-m aperture towed array. Due to the wide element separation, the high-frequency (kHz range) signals generate numerous grating lobes, but these lobes shift bearing with beamformed frequency, allowing identification of the true bearing whenever the whistles have over 1 kHz bandwidth. This method was validated by matching a sighting of a compact group of dolphins with acoustic bearing estimates. The system was subsequently used to detect and determine bearings from animals at least 3 km away and in Beaufort 5+ conditions. Frequency-domain beamforming has advantages over temporal cross correlation when the signals are faint and/or overlapping. © 2000 Acoustical Society of America.

[S0001-4966(00)00405-7]

PACS numbers: 43.80.Lb [WA]

## INTRODUCTION

Acoustic bearings from kilohertz-range marine mammal vocalizations are typically obtained via two methods: either by time-delay (TD) measurements of signals recorded on widely spaced hydrophones,<sup>1,2</sup> or more recently<sup>3</sup> by beamforming in the frequency domain (FD), using signals collected from many closely spaced hydrophones. Although mathematically equivalent to delaying and summing signals in time, frequency-domain methods demonstrate two advantages over time-domain beamforming. First, FD methods enable greater coherent gain when used with multielement arrays<sup>4</sup> and thus should be able to obtain bearings from signals too weak to be localized by TD methods. Second, delphinid vocalizations often overlap in time and frequency. Under these circumstances time-domain methods may fail due to interference between the signals. Frequency-domain methods can isolate overlapping signals from each other, because they allow precise control over which frequency components in a data sample are processed.

A potential disadvantage of FD beamforming occurs whenever the spacing between adjacent hydrophones exceeds half an acoustic wavelength at a given frequency, and the beamformer cannot distinguish between the true signal bearing (the mainlobe) and multiple false bearings (i.e., grating lobes, or sidelobes). For example, FD beamforming on a 10-kHz pure tone (15-cm wavelength) with hydrophones spaced 5 m apart generates over 60 grating lobes, making mainlobe identification impossible.

Most dolphin whistles are frequency modulated,<sup>5</sup> and thus have wide frequency bandwidth. Because grating lobe

maxima shift position with beamformed frequency, this bandwidth can be exploited to identify the mainlobe. In this letter these sidelobe properties were exploited to conduct FD beamforming on kilohertz-range marine mammal calls recorded on three hydrophones separated by 4 and 5 m, a situation that routinely generated 30–60 grating lobes at each frequency.

## I. BACKGROUND

Conventional frequency-domain algorithms<sup>4</sup> decompose a received signal into its frequency components, and then generally assume that each component represents a plane wave arriving at a bearing  $\theta$  from the array. In the following discussion a bearing of  $0^\circ$  refers to an angle arriving perpendicular (broadside) to the line array. The algorithm compares the data with that expected from a plane wave arriving from a test bearing  $\theta'$ , and produces the following output:

$$B(f, \theta) = \left| \sum_n e^{i(2\pi d/\lambda)(\sin \theta - \sin \theta')} \right|^2. \quad (1)$$

Here,  $d$  is the spacing between array hydrophones, and  $\lambda$  is the wavelength of the signal at frequency  $f$ . Note how the beamformer output will always attain a maximum (mainlobe) when  $\theta = \theta'$ . However, if the ratio  $d/\lambda$  is greater than  $1/2$ , then other bearings  $\theta'_m$  will produce beamformer maxima, or grating lobes. Their bearings are given by the following formula:

$$\sin \theta_m = \sin \theta \pm \frac{cm}{fd}, \quad m = 0, \pm 1, \pm 2, \pm 3, \dots \quad (2)$$

where  $c$  is the waterborne sound speed,  $f$  is the signal frequency, and  $m$  is an integer.

From Eq. (2) it follows that grating lobe bearings (which

<sup>a)</sup>Current address: Dept. of Ocean Eng., Room 5-212, MIT, 77 Massachusetts Ave., Cambridge, MA 02139.

<sup>b)</sup>Current address: SAIC Maritime Division, 3990 Old Town Ave., San Diego, CA 92110.



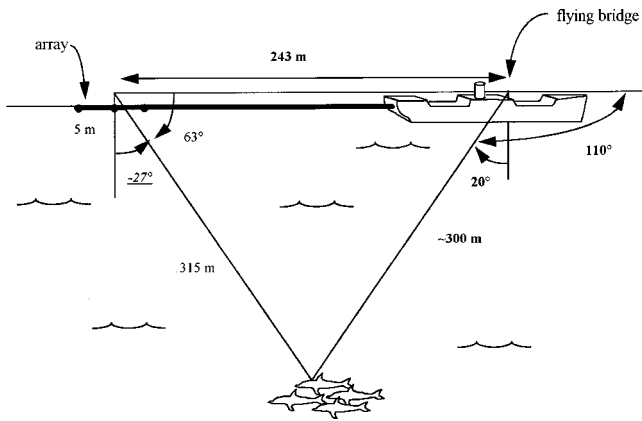


FIG. 1. Geometry used to convert a visual bearing into a relative array bearing for data in Fig. 2. Boldfaced numbers are coordinates measured by a visual observer on the flying bridge. The underlined number is the derived bearing, relative to the array.

have a nonzero value of  $m$ ) will change with the beamformed frequency. Although the grating lobe magnitudes are identical to that of the mainlobe, the latter is easily identified as the maxima that maintains a fixed bearing with frequency. Averaging multiple beams generated at different frequencies thus suppresses the grating lobes. This fact allows FD methods to be applied to three widely spaced elements, preserving array aperture and thus angular precision.<sup>4</sup> Frequency averaging is a common technique in many applications involving sparse arrays.<sup>6,7</sup>

## II. EXPERIMENTAL SETUP

The visual and acoustic data from dolphins were recorded in Nov./Dec. 1998 during a National Marine Fisheries Service dolphin survey cruise, near the coast of Panama.<sup>8</sup> A three-element array (built by Don Norris, SonaTech, Inc.) was deployed at  $\sim 5$  m depth, 200 m behind the fantail of the R/V ENDEAVOR, a 175-ft oceanographic research vessel. The individual elements were spaced 4 and 5 m apart, and had omnidirectional sensitivity and a nearly linear frequency response from 2–150 kHz. The signals from the three phones were fed into a Mackie CR1604-VLZ audio mixer, which was used as a 200-Hz high-pass filter. The signals were then passed through a TASCAM DA-38 digital recorder with a 48-kHz low-pass filter, and recorded for later analysis. Simultaneously, the TASCAM outputs were passed through the mixer a second time to allow further amplification and filtering (usually to emphasize the 8–20 kHz band), before being sampled into two Data Translation DT-3809 12-bit A/D cards within a Dell Optiplex 200-MHz Pentium computer.

The beamforming software was written in MATLAB 5.0. When executed, the program sampled the data at 48 kHz and wrote up to a minute of multichannel data to hard disk. The data from the first hydrophone were then read and displayed as a spectrogram. A simple mouse-driven graphical user interface allowed the user to trace the time-frequency contour of a whistle in order to analyze it. When selected, the program computed and averaged the spectra of the first 1024 points of the signal from each hydrophone. The frequency

with the highest average power within a certain frequency window was then selected for beamforming. The program then adjusted its selection window based on the contour trace, obtained the next data sample, and repeated the process. Using this technique, most of the frequency components of a modulated sweep could be beamformed, regardless of whistle contour or duration. A selection window of  $\pm 500$  Hz bandwidth was found sufficient to track all but the most rapidly modulated whistles. In cases where multiple whistles were recorded at the same time, the selection window was made narrower to avoid interference.

The individual beamformer outputs were then displayed as a frequency-azimuth plot before being incoherently averaged to yield a total beamformer output, where the mainlobe is revealed as the global maximum.

## III. RESULTS

The beamforming algorithm was first calibrated on sperm whale clicks arriving from known bearings, identified using software developed by J. Barlow. Next, the methods were tested on over 100 whistles and burst pulses. On 16:36 Dec. 6, the system determined bearings from animals that were 3 km away, as determined by visual sighting methods. On another occasion the acoustic observers obtained bearings under Beaufort 5+ conditions, and assisted visual observers in locating a school of *Tursiops truncatus* during twilight, providing an example of how acoustic bearings can help supplement visual methods under adverse conditions.

Three specific examples of the system output are now presented. The first example demonstrated a match between a visual and acoustic observation on Dec. 5, 1998. On this date a compact group of 36–45 striped dolphins (*Stenella coeruleoalba*) approached the ship, rode the bow, and eventually passed by the beam of the Endeavor. The occurrence of an isolated, compact group of animals was an uncommon event—typically, dolphin encounters consisted of hundreds of animals widely scattered over an area. Between 10:00 and 10:08 an experienced marine mammal observer tracked the group from the flying bridge using 25x reticulated binoculars, recording the horizontal and vertical azimuth from the horizon approximately every 30 sec. At 10:02 the animals were at a horizontal azimuth of  $20^\circ$  past the ship's beam ( $110^\circ$  relative to the ship's bow), and about 300 m away from the flying bridge. The bearing of the group relative to the array was estimated to be near  $-26^\circ$  (as illustrated in Fig. 1).

A frequency-modulated (FM) whistle was recorded at 10:02, and is displayed in Fig. 2(a). Over 1 sec the whistle swept between a frequency of 9.8 to 18 kHz. The individual beams obtained from this whistle are plotted vs frequency in Fig. 2(b). The rapid frequency modulation precludes beamforming at every frequency bin within this band; however, the identity of the mainlobe is clear. Note that even if one were restricted to beamforming between 9–11 kHz, the mainlobe location would be readily distinguishable. In fact, signals with as little as 1 kHz modulation were successfully beamformed.

By averaging all the beams together [i.e., summing along the frequency axis in Fig. 2(b)], and subtracting the

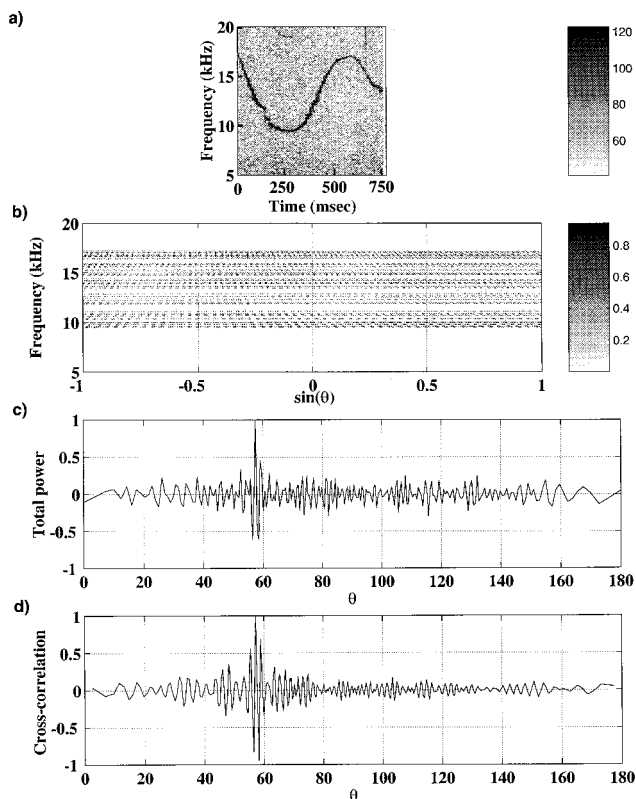


FIG. 2. (a) Spectrogram of dolphin whistle recorded at 10:02 am, Dec. 5, 1998. Fourier transform length is 1024 pts, sampling rate is 48 kHz, window overlap is 50%. (b) Plane-wave (PW) beams from (a), plotted as a function of frequency and the sine of the bearing. A negative value means the signal is arriving from the ship's bow. Dark colors represent stronger correlation. Note how all maxima shift positions with frequency, except for the main-lobe at position  $\sin(\theta) = -0.52$ . (c) Incoherent average of all beams vs bearing  $\theta$ . The whistle is arriving from a bearing of  $-33^\circ$ , or  $57^\circ$  away from the ship's heading. The mean value of the averaged beams has been subtracted and the pattern normalized by its maximum. (d) Cross correlation of the entire time series in (a), prefiltered between 9.4 and 17 kHz.

mean, a combined beamformer output can be displayed [Fig. 2(c)]. (Note that the horizontal axis is now in terms of  $\theta$ , not  $\sin \theta$ .) Finally, the result of cross correlating the first two hydrophone signals is shown in Fig. 2(d). For isolated whistles with high signal-to-noise ratio, both the time- and frequency-domain methods give similar results.

The beamformer estimates an acoustic bearing of  $-33^\circ$ , whereas the visual observer bearing estimate was  $-26^\circ$ . Given the disparity in timing between the visual and acoustic records ( $\sim 10$  sec), and the rapid motion of the dolphin group, the two measurements are in close agreement. A second visual observation at 10:08 gave a bearing of  $50^\circ$ , and the corresponding acoustic bearing estimate was  $41^\circ$ —again, an approximate match. Despite the fact that the range of the animals was less than 400 m, the plane-wave approximation seemed to work well, and modeling wave front curvature was not required.

The next example shows the advantage of using coherent gain from three hydrophones. Figure 3 illustrates a faint whistle recorded at 17:16 on Dec. 6, one day after the previous example. At this time multiple animals in a group were whistling; this example concerns the whistle indicated by the arrow in Fig. 3(a). As in Fig. 2, the individual beamformed

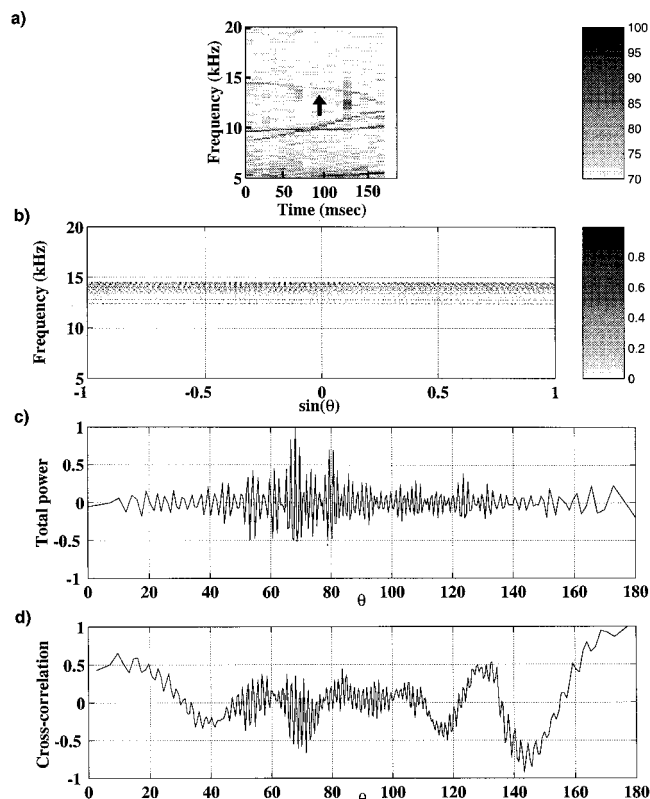


FIG. 3. (a) Spectrogram of dolphin whistle at 17:16, Dec. 6, 1998. Fast Fourier transform (FFT) length is 1024 pts, sampling rate is 48 kHz, window overlap is 50%. The arrow indicates the faint whistle to be processed. (b) PW beams plotted as a function of frequency and  $\sin(\theta)$ . (c) Incoherent summation of the beams displayed in (b), vs bearing  $\theta$ . (d) Cross correlation of the entire time series in (a), prefiltered between 12 and 15 kHz.

outputs are shown in Fig. 3(b) and the averaged beamformer output between 12 and 15 kHz is plotted in Fig. 3(c) For comparison, the cross correlation between the first two phones is provided in Fig. 3(d). In this last plot the data have been digitally bandpass filtered using a finite impulse response (FIR) filter, designed to bandpass the same frequency range displayed in Fig. 3(b). The FD result yields a bearing of  $68^\circ$ , whereas cross correlation fails due to the low signal-to-noise ratio. If the original signal is digitally refiltered using different choices for passband, the cross correlation still fails.

The final example was recorded at 17:02 Dec. 6, and is shown in Fig. 4. It illustrates a common situation, wherein two whistles by two separate animals overlap in frequency and time, such that the signals cannot be isolated by simple bandpass filtering. One whistle has the upward U-shaped contour labeled "1" in Fig. 4(a) and the second has the relatively longer FM sweep labeled "2" in the same spectrogram. When applying the beamformer algorithm, a frequency selection window of 100 Hz was used to ensure that the beamformer sampled only the energy from whistle 2, whose output is shown in Fig. 4(b) and (c). The mainlobe position indicates that whistle 2 is arriving from the bow of the ship.

When time-domain techniques are applied to two overlapping signals, one obtains two separate peaks if the signals are uncorrelated, with no means of telling which peak was generated by which signal. If the signals are correlated (i.e.,

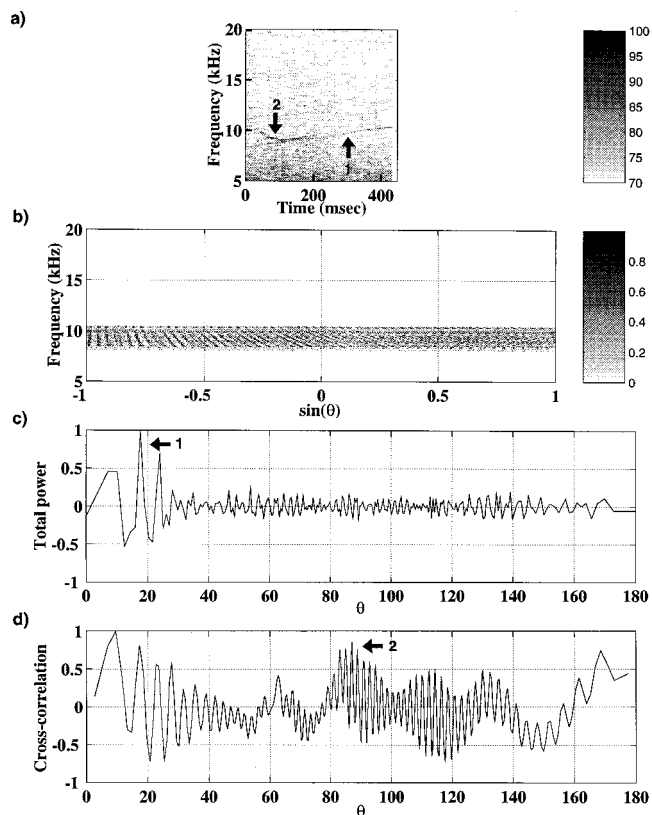


FIG. 4. (a) Spectrogram of overlapping dolphin whistles at 17:02, Dec. 6, 1998. For this case the search window is narrowed to 0.1 kHz around the frequency contour of whistle #1. (b) PW beams from whistle 1, plotted as a function of frequency and  $\sin(\theta)$ . (c) Incoherent summation of the beams displayed in (b), vs bearing  $\theta$ . The arrow indicates the bearing of the whistle 1. (d) Cross correlation of the time series in (a), prefiltered between 8 and 11 kHz. As it is impossible to filter out the effects of whistle 2, which is arriving from broadside, the output is corrupted and the bearing of whistle 1 is not obtained. The arrow indicates the contamination from whistle 2.

both share similar frequency contours) then the bearing estimates can be compromised, as in the case in Fig. 4(d). Here the time-domain output shows the combined effects of whistle 2, which is arriving from broadside, and whistle 1. The bearing estimate of the latter whistle has been corrupted due to the fact that both signals are slightly correlated. The FD beamformer avoided this problem and yielded separate bearings for each whistle. The ability to obtain bearings from overlapping whistles was useful when trying to separate distant animals ahead of the ship from nearby animals the ship had already passed.

#### IV. CONCLUSION

Successful frequency-domain beamforming on dolphin whistles has been demonstrated on a three-element sparse

array, in which the array elements are spaced apart by more than 50 times a typical acoustic wavelength. FD beamforming is demonstrated to be more effective than temporal cross correlation in cases where the signal is faint and/or overlapping with other signals. The sparse-array beamforming methods discussed here and in Ref. 3 should work on most broadband marine mammal sounds, including the majority of tonal sounds produced by odontocetes.<sup>9</sup>

#### ACKNOWLEDGMENTS

The authors would like to thank the crew and the marine mammal observers of the R/V ENDEAVOR (particularly Laura Morse, Meghan Donahue, and Jorge Rodríguez) for their cooperation and support during the field testing. They would also like to thank Patrick Miller and Peter Tyack of Woods Hole Oceanographic Institute for advice about the manuscript. This work was funded under U.S. Department of Commerce Requisition #NFFR2100800295. Equipment support was provided by the National Marine Fisheries Service and the Office of Naval Research. Tim Gerrodette was Chief Scientist for the dolphin survey.

- <sup>1</sup>L. E. Freitag and P. L. Tyack, "Passive acoustic localization of the Atlantic bottlenose dolphin using whistles and echolocation clicks," *J. Acoust. Soc. Am.* **93**, 2197–2205 (1993).
- <sup>2</sup>W. A. Watkins and W. E. Schevill, "Sound source location by arrival times on a non-rigid three-dimensional hydrophone array," *Deep-Sea Res.* **19**, 691–706 (1972).
- <sup>3</sup>P. J. Miller and P. L. Tyack, "A small towed beamforming array to identify vocalizing resident killer whales (*Orcinus Orca*) concurrent with focal behavioral observations," *Deep-Sea Research II* **45**, 1389–1405 (1998).
- <sup>4</sup>D. J. DeFatta, J. G. Lucas, and W. S. Hodgkiss, "Conventional beamforming, Appendix A11," in *Digital Signal Processing: A System Design Approach* (Wiley, New York, 1988), pp. 629–649.
- <sup>5</sup>J. N. Oswald, "Delphinid whistles recorded in the eastern tropical Pacific Ocean," National Marine Fisheries Service, Southwest Fisheries Science Center, Administrative LJ-99-07C, July 1999.
- <sup>6</sup>M. J. Hinich, "Processing spatially aliased arrays," *J. Acoust. Soc. Am.* **64**, 792–794 (1978).
- <sup>7</sup>F. Anderson, W. Christensen, L. Fullerton, and B. Kortegaard, "Ultra-wideband beamforming in sparse arrays," *IEE Proc., Part H: Microwaves, Antennas, Propag.* **138**, 342–346 (1991).
- <sup>8</sup>P. Olson and T. Gerrodette, "Report of the meeting to review the preliminary estimates of Eastern Tropical Pacific dolphin abundance in 1998," National Marine Fisheries Service, Southwest Fisheries Science Center, Administrative LJ-99-03, Jan. 21, 1999.
- <sup>9</sup>J. N. Mathews, L. E. Rendall, J. C. D. Gordon, and D. W. Macdonald, "A review of frequency and time parameters of cetacean tonal calls," *Bioacoustics* **10**, 47–71 (1999).