

THE JOURNAL of the Acoustical Society of America

Vol. 106, No. 6

December 1999

SOUNDINGS SECTION

ACOUSTICAL NEWS—USA		3041
ACOUSTICAL NEWS—INTERNATIONAL		3045
International Meetings Calendar		3045
FORUM		
Underwater sound transmission and SI units	Clarence S. Clay	3047
Noise Control and SI Units	Robert Hickling	3048

GENERAL LINEAR ACOUSTICS [20]

Improvement of sound barriers using headpieces with finite acoustic impedance	Michael Möser, Rudi Volz	3049
Wideband quantitative ultrasonic imaging by time-domain diffraction tomography	T. Douglas Mast	3061
A Galerkin method for the numerical analysis of diffraction by a rectangular screen	Nathaniel J. de Lautour	3072
Acoustic response of a periodic layer of nearly rigid spherical inclusions in an elastic solid	Konstantin Maslov, Vikram K. Kinra	3081
Effect on ultrasonic signals of viscous pore fluids in unconsolidated sand	Patricia K. Seifert, Bruno Kaelin, Lane R. Johnson	3089
Computation of transient radiation in semi-infinite regions based on exact nonreflecting boundary conditions and mixed time integration	Lonny L. Thompson, Runnong Huan	3095
Characterization of multiple-sprung masses for wideband noise control	G. Maidanik, K. J. Becker	3109
Criteria for designing multiple-sprung masses for wideband noise control	G. Maidanik, K. J. Becker	3119
A class of expansion functions for finite elastic plates in structural acoustics	Richard V. Craster, Stefan G. Llewellyn Smith	3128

NONLINEAR ACOUSTICS [25]

Nonlinear propagation of laser-generated sound pulses in a water and granular medium	K. A. Naugolnykh, S. V. Egerev, I. B. Esipov, K. A. Matveev	3135
Nonlinear longitudinal waves in inhomogeneously predeformed elastic media	Arvi Ravasoo	3143
Multifrequency plane, nonlinear, and dissipative waves at arbitrary distances	Claes M. Hedberg	3150
Nonlinear oscillation of a spherical gas bubble in acoustic fields	Masaharu Kameda, Yoichiro Matsumoto	3156

(Continued)

CONTENTS—Continued from preceding page

AEROACOUSTICS, ATMOSPHERIC SOUND [28]

- Two-dimensional effects in the edge sound of vortices and dipoles G. Schouten 3167

UNDERWATER SOUND [30]

- Development of a velocity gradient underwater acoustic intensity sensor Kevin J. Bastyr, Gerald C. Lauchle, James A. McConnell 3178
- Low-frequency ambient sound in the North Pacific: Long time series observations Keith R. Curtis, Bruce M. Howe, James A. Mercer 3189
- Ambient noise imaging in warm shallow seas; second-order moment and model-based imaging algorithms John R. Potter, Mandar Chitre 3201
- Imaging in the ocean with ambient noise: the ORB experiments Chad L. Epifanio, John R. Potter, Grant B. Deane, Mark L. Readhead, Michael J. Buckingham 3211
- Automatic matched-field tracking with table lookup Homer Bucker, Paul A. Baxley 3226
- A three-dimensional propagation algorithm using finite azimuthal aperture Kevin B. Smith 3231
- Bottom reverberation in shallow water: Coherent properties as a function of bandwidth, waveguide characteristics, and scatterer distributions Kevin LePage 3240
- Broadband source localization in shallow water in the presence of internal waves Kwang Yoo, T. C. Yang 3255
- Range-dependent matched-field inversion of SWellEX-96 data using the downhill simplex algorithm Martin Musil, N. Ross Chapman, Michael J. Wilmut 3270

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

- Dispersion and absorption of sound in monatomic gases: An extended kinetic description Wilson Marques, Jr. 3282
- Internal circulation in a drop in an acoustic field Hong Zhao, S. S. Sadhal, Eugene H. Trinh 3289
- Shear interfacial waves in piezoelectrics A. N. Darinskii, V. N. Lyubimov 3296
- Bjerknes forces between two bubbles in a viscous fluid Alexander A. Doinikov 3305
- An acoustical helicoidal wave transducer with applications for the alignment of ultrasonic and underwater systems Brian T. Hefner, Philip L. Marston 3313

TRANSDUCTION [38]

- One-dimensional phenomenological model of hysteresis. I. Development of the model Jean C. Piquette, Stephen E. Forsythe 3317
- One-dimensional phenomenological model of hysteresis. II. Applications Jean C. Piquette, Stephen E. Forsythe 3328
- A basic concept of direct converting digital microphone Yoshinobu Yasuno, Yasuhiro Riko 3335

STRUCTURAL ACOUSTICS AND VIBRATION [40]

- Magnetic excitation and acoustical detection of torsional and quasi-flexural modes of spherical shells in water Brian T. Hefner, Philip L. Marston 3340
- Measurement and prediction of diffuse fields in structures M. J. Evans, P. Cawley 3348
- Reduced models for structures in the medium-frequency range coupled with internal acoustic cavities Christian Soize 3362
- Three-dimensional vibration analysis of a cantilevered parallelepiped: Exact and approximate solutions C. W. Lim 3375
- The sensitivity of structural acoustic response to attachment feature scale representation Kenneth A. Cunefare, Sergio De Rosa 3384

CONTENTS—Continued from preceding page

Near-field sensing strategies for the active control of the sound radiated from a plate	A. Berry, X. Qiu, C. H. Hansen	3394
Experimental demonstration of a band-limited actuator/sensor selection strategy for structural acoustic control	Robert L. Clark, David E. Cox	3407
NOISE: ITS EFFECTS AND CONTROL [50]		
Statistical structures of indoor traffic noise in a high-rise city	S. K. Tang, W. H. Au	3415
ACOUSTICAL MEASUREMENTS AND INSTRUMENTATION [58]		
A pulsed phase-sensitive technique for acoustical measurements	Vladimir F. Kozhevnikov, Dmitrii I. Arnold, Matthew E. Briggs, Salim P. Naurzakov, John M. Viner, P. Craig Taylor	3424
ACOUSTIC SIGNAL PROCESSING [60]		
Passive wideband cross correlation with differential Doppler compensation using the continuous wavelet transform	Brian G. Ferguson, Kam W. Lo	3434
Optimal array element localization	Stan E. Dosso, Barbara J. Sotirin	3445
PHYSIOLOGICAL ACOUSTICS [64]		
Responses of neurons to click-pairs as simulated echoes: Auditory nerve to auditory cortex	Douglas C. Fitzpatrick, Shigeyuki Kuwada, D. O. Kim, Kourosh Parham, Ranjan Batra	3460
Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio f_2/f_1	Manfred Mauermann, Stefan Uppenkamp, Peter W. J. van Hengel, Birger Kollmeier	3473
Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. II. Fine structure for different shapes of cochlear hearing loss	Manfred Mauermann, Stefan Uppenkamp, Peter W. J. van Hengel, Birger Kollmeier	3484
Phonemes, intensity and attention: Differential effects on the mismatch negativity (MMN)	Michael D. Szymanski, E. William Yund, David L. Woods	3492
PSYCHOLOGICAL ACOUSTICS [66]		
Loudness recalibration as a function of level	Dan Mapes-Riordan, William A. Yost	3506
Psychometric functions for gap detection in a yes–no procedure	Mary Florentine, Søren Buus, Wei Geng	3512
Informational masking by everyday sounds	Eunmi L. Oh, Robert A. Lutfi	3521
The auditory continuity phenomenon: Role of temporal sequence structure	Carolyn Drake, Stephen McAdams	3529
Effects of frequency and duration on psychometric functions for detection of increments and decrements in sinusoids in noise	Brian C. J. Moore, Robert W. Peters, Brian R. Glasberg	3539
Context dependence of fundamental-frequency discrimination: Lateralized temporal fringes	Hedwig Gockel, Robert P. Carlyon, Christophe Micheyl	3553
The influence of musical training on the perception of sequentially presented mistuned harmonics	Edward M. Burns, Adrianus J. M. Houtsma	3564
The effect of temporal placement on gap detectability	Karen B. Snell, Hue-Lu Hu	3571
The role of perceived spatial separation in the unmasking of speech	Richard L. Freyman, Karen S. Helfer, Daniel D. McCall, Rachel K. Clifton	3578

CONTENTS—Continued from preceding page

Auditory localization of nearby sources. III. Stimulus effects	Douglas S. Brungart	3589
Comparison of different forms of compression using wearable digital hearing aids	Michael A. Stone, Brian C. J. Moore, José I. Alcántara, Brian R. Glasberg	3603
SPEECH PRODUCTION [70]		
Determination of parameters for lumped parameter models of the vocal folds using a finite-element method approach	M. P. de Vries, H. K. Schutte, G. J. Verkerke	3620
SPEECH PERCEPTION [71]		
Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants	Stuart Rosen, Andrew Faulkner, Lucy Wilkinson	3629
A method to determine the speech transmission index from speech waveforms	Karen L. Payton, Louis D. Braida	3637
Training American listeners to perceive Mandarin tones	Yue Wang, Michelle M. Spence, Allard Jongman, Joan A. Sereno	3649
BIOACOUSTICS [80]		
Frequency dependence of ultrasonic backscatter from human trabecular bone: Theory and experiment	Keith A. Wear	3659
Simulation of ultrasonic pulse propagation, distortion, and attenuation in the human chest wall	T. Douglas Mast, Laura M. Hinkelman, Leon A. Metlay, Michael J. Orr, Robert C. Waag	3665
Excitation and propagation of surface waves on a viscoelastic half-space with application to medical diagnosis	T. J. Royston, H. A. Mansy, R. H. Sandler	3678
Low-frequency whale sounds recorded on hydrophones moored in the eastern tropical Pacific	Kathleen M. Stafford, Sharon L. Nieukirk, Christopher G. Fox	3687
Transmission beam pattern and echolocation signals of a harbor porpoise (<i>Phocoena phocoena</i>)	Whitlow W. L. Au, Ronald A. Kastelein, Tineke Rippe, Nicole M. Schooneman	3699
Modeling vibration and sound production in insects with nonresonant stridulatory organs	Gunther Tschuch, Denis J. Brothers	3706
Receiver operating characteristics and temporal integration in an insect auditory receptor cell	Jakob Tougaard	3711
Generation and maintenance of bubbles in small tubes by low-frequency ultrasound	Kwok W. Lam, James M. Drake, Richard S. C. Cobbold	3719
LETTERS TO THE EDITOR		
The force between two parallel rigid plates due to the radiation pressure of phonons [25]	O. Bschorr	3730
Structural radiation mode sensing for active control of sound radiation into enclosed spaces [50]	Ben S. Cazzolato, Colin H. Hansen	3732
Voice onset times and burst frequencies of four velar stop consonants in Gujarati [70]	Manish K. Rami, Joseph Kalinowski, Andrew Stuart, Michael P. Rastatter	3736

CONTENTS—Continued from preceding page

ACOUSTICS RESEARCH LETTERS ONLINE

Frequency acuity and binaural masking release in dyslexic listeners	Nicholas I. Hill, Peter J. Bailey, Yvonne M. Griffiths, Margaret J. Snowling	L53
Vibration of beads placed on the basilar membrane in the basal turn of the cochlea	Nigel P. Cooper	L59
Effect of acoustic dynamic range on phoneme recognition in quiet and noise by cochlear implant users	Qian-Jie Fu, Robert V. Shannon	L65
Consonant recordings for speech testing	Robert V. Shannon, Angela Jensvold, Monica Padilla, Mark E. Robert, Xiaosong Wang	L71
INDEX TO VOLUME 106		3739
SUBJECT INDEX TO VOLUME 106		3744
AUTHOR INDEX TO VOLUME 106		3767

NOTES CONCERNING ARTICLE ABSTRACTS

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

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

Frequency acuity and binaural masking release in dyslexic listeners

Nicholas I. Hill, Peter J. Bailey, Yvonne M. Griffiths, Margaret J. Snowling

*Department of Psychology, University of York, York YO10 5DD, United Kingdom
nih1@york.ac.uk pjb1@york.ac.uk ymg100@york.ac.uk mjs19@york.ac.uk*

Abstract: Two experiments compared auditory sensitivity in a group of 12 adult dyslexics and matched control listeners. The first experiment measured frequency discrimination and frequency modulation detection thresholds at both 1 and 6 kHz. Although thresholds were larger for the dyslexic group, the differences were not statistically reliable. The second experiment measured the binaural masking level difference for a 200 Hz pure tone in noise. Thresholds did not differ significantly between the two groups. The data provide little support for the hypothesis that dyslexic listeners are impaired in their ability to process information in the temporal fine structure of auditory stimuli.

© 1999 Acoustical Society of America

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Fe

1. Introduction

Developmental dyslexia or specific reading impairment occurs in approximately 5% of the population and is defined typically as a discrepancy between an individual's actual reading ability and his or her predicted ability based on age and intelligence. Recently a number of psychoacoustic and physiological studies have reported results that suggest dyslexia may be linked to deficits in auditory processing (McAnally and Stein, 1996; Hari and Kiesila, 1996; Dougherty *et al.*, 1998; Baldeweg *et al.*, 1999). For example, McAnally and Stein (1996) compared the performance of a group of adult dyslexics with a group of matched control listeners on four auditory tasks: detection of a temporal gap in broadband noise, pure-tone frequency discrimination around 1 kHz, detection of a 1-kHz tone presented in a diotic noise masker when the tone was in phase at the two ears (NOS0), and when the tone was presented with opposite phase at the two ears (NOS π). Average thresholds for gap detection and detection of a tone in noise in the NOS0 condition did not differ significantly between the two groups. However, significant group differences were observed in the frequency discrimination task and the NOS π condition of the tone-in-noise detection task. In accounting for this pattern of results, McAnally and Stein (1996) hypothesized that dyslexics may be impaired in their ability to extract information about the temporal fine structure of auditory stimuli from the phase locking of auditory nerve firing patterns. Data consistent with this hypothesis have also been reported by Dougherty *et al.* (1998) and Baldeweg *et al.* (1999). There is also some evidence that dyslexic listeners have difficulty processing dynamic stimuli such as amplitude- and frequency-modulated tones (McAnally and Stein, 1997; Witton *et al.*, 1998).

The experiments reported here measured frequency difference limens (DLFs), frequency modulation detection limens (FMDLs) and binaural masking level differences (MLDs) for a group of adult dyslexic listeners and matched controls (see Moore, 1997 for a definition of these measures). The extent to which any deficits in the DLF and FMDL tasks were due to an inability to use information derived from phase locking was investigated by

measuring thresholds at a relatively low frequency, at which information derived from phase locking should be available, and at a higher frequency at which it should not.

2. Subject selection

Twelve listeners with a documented history of reading and spelling problems participated in the study; 11 were undergraduate students registered with the University disability officer at the University of York. Dyslexic adults are often able to compensate for their reading difficulties, so the IQ discrepancy criteria used to define dyslexia in samples of children are inappropriate for adults (Pennington *et al.*, 1986; Wolff and Melngailis, 1994). However, the persistence of an underlying phonological deficit in adulthood can be demonstrated using tests of nonword reading, spelling, phoneme deletion, phoneme fluency, and rapid naming (Bruck, 1992; Snowling *et al.*, 1997). Twelve normal readers were selected as controls to match the dyslexics in age and IQ. The dyslexic and control groups did not differ significantly on verbal or nonverbal IQ measures, but there were significant group differences on standardized reading and spelling tests and phonological tasks including digit span, nonword reading, phoneme deletion, spoonerisms, speech rate, nonword repetition, and rapid naming (see Table 1.). All listeners had audiometric pure-tone thresholds not exceeding 20 dB HL at octave frequencies in the range 250-8000 Hz.

Table 1. Subject characteristics (means and standard deviations)

	Dyslexic	Control	F(1,22)	MSE
Age	24.17 (7.8)	24.61 (10.1)	F<1	-
Vocabulary ¹	12.17 (2.2)	13.36 (2.5)	1.46	5.63
Block design ¹	12.92 (3.0)	14.00 (2.7)	F<1	-
Digit span ¹	9.67 (2.53)	11.75 (2.45)	4.18, p<0.05	6.22
WRAT reading ²	103.25 (6.1)	111.5 (5.0)	11.24, p<0.01	23.25
WRAT spelling ²	94.5 (6.3)	112.3 (6.6)	41.26, p<0.001	7.66
Nonword reading ³	10.55 (2.9)	13.75 (0.8)	11.53, p<0.01	4.51
Phoneme deletion ⁴	18.83 (2.82)	22.92 (2.15)	15.88, p<0.001	6.30
Spoonerisms ⁴	18.67 (7.73)	23.67 (1.15)	4.92, p<0.05	30.52
Speech rate ⁵	1.84 (0.29)	2.36 (0.36)	12.73, p<0.01	0.11
Word repetition ⁶	11.64 (0.67)	11.92 (0.29)	1.48	0.25
Nonword repetition ⁶	8.45 (3.24)	11.17 (1.19)	5.84, p<0.05	5.60
Rapid naming ⁷	1.31 (1.42)	-0.42 (0.74)	13.88, p<0.01	1.29

Notes: 1. Wechsler adult intelligence scale - revised, scaled score; 2. Wide range achievement test - III, standard score; 3. Max. score=15; 4. Max. score=24; 5. words/min; 6. Max. score=12; 7. digits and objects, composite z score. See Snowling *et al.* (1997) for details.

3. Experiment 1: DLFs and FMDLs

3.1. Stimuli and equipment

Two test frequencies were used, 1 and 6 kHz. In addition, a tone at 2.45 kHz (the geometric mean of the two test frequencies) was used for practice. All tones had a duration of 400 ms including 10-ms cosine-squared onsets and offsets and were presented at a level of 65 dB SPL. For the frequency modulation detection condition, the rate of modulation was 2.5 Hz, with the modulation envelope always in sine phase.

Stimuli were synthesized in real-time at a sampling rate of 50 kHz using custom software running on an IBM-compatible PC. The resulting waveforms were converted to voltages using a 16-bit digital-to-analogue converter (Tucker-Davis Technologies, model DD1). The level of the stimuli was controlled using a pair of attenuators (TDT model PA4). Stimuli were presented diotically over Sennheiser HD414 headphones. Listeners were run

individually in a sound-attenuating enclosure and responded using a standard computer keyboard.

3.2. Procedure

DLFs and FMDLs were estimated using a four-interval, two-alternative, forced-choice procedure with the signal presented in either the second or third interval. The first observation interval began 400 ms after the offset of a 400 ms visual alerting stimulus, with subsequent observation intervals separated from one another by a 400 ms silence. Listeners were given an unlimited time in which to respond. Immediately after responding, they were presented with feedback for 400 ms. The next trial began 800 ms after the termination of feedback.

The frequency difference or modulation depth (zero-peak), according to condition, was adjusted adaptively using a 2-down, 1-up rule, which targeted the threshold corresponding to 70.7% correct responses (Levitt, 1971). Each 60-trial run began with the frequency difference/modulation depth set to 2% of the center frequency. The frequency difference or modulation depth was adjusted logarithmically using an initial multiplier of 0.8, which was decreased to 0.64 following the first four reversals. DLFs and FMDLs at both 1 and 6 kHz were determined over two 1-hour sessions. Eight threshold estimates were obtained per session, two for each frequency and condition. The conditions were blocked such that in one-half of the session DLFs were estimated whereas in the other half, FMDLs were estimated. Within each condition, the four adaptive runs were presented in a random order. The experiment paused for a minimum of 30 seconds after the completion of a run, and listeners initiated the next run with a keypress. Before each block, listeners completed a single 60-trial practice run at 2.45 kHz. The order in which the conditions were presented was counterbalanced across sessions and listeners. Thresholds for an individual run were calculated by averaging the values of frequency difference/modulation depth across an even number of reversals, excluding the first three or four reversals as appropriate. The mean number of reversals contributing to the determination of threshold exceeded twelve. Reported thresholds correspond to the arithmetic mean of the individual threshold estimates.

3.3. Results and discussion

Individual and mean DLFs and FMDLs for the two groups of listeners, expressed as a percentage of center frequency, are shown in Figs. 1 and 2 respectively. The error bars associated with each mean denote plus/minus one standard deviation. Consider first the data for the frequency discrimination condition (Fig. 1.). Mean thresholds for the dyslexic group were larger than those of the control group at both frequencies. There were large variations in thresholds, particularly among the dyslexic sample, with an order of magnitude difference between the lowest and highest DLFs at both 1 and 6 kHz. The statistical reliability of group differences was assessed using Mann-Whitney U tests (level of significance: 5%).

The difference in performance between the two groups was not significant at either frequency. The differences were due almost entirely to four members of the dyslexic sample (H, I, J and K). The lack of a significant group difference in the 1-kHz condition contrasts with the findings of McAnally and Stein (1996) and Baldeweg *et al.* (1999). One possibility, given the large degree of intersubject variation, is that their studies included a larger proportion of dyslexics like H, I, J and K. Even for these listeners, the fact that DLFs were elevated at both 1 and 6 kHz suggests that their poorer performance cannot be attributed simply to an inability to process temporal fine structure information.

In the FMDL task, mean thresholds were also larger for the dyslexic group than the control group at both frequencies, but, again, the difference in performance was not statistically reliable. There was somewhat less variability in thresholds across listeners,

although there were still three listeners with FMDLs more than two standard deviations from the mean (J and W at 1 kHz, H and W at 6 kHz). With these outliers removed, there is a small but significant difference in performance between the two groups at 1 kHz ($U=34$, $p=0.04$), but not at 6 kHz. This finding is consistent with Witton *et al.*'s observation that dyslexics had elevated FMDLs at 1 kHz using a 2 Hz rate of modulation. However, the fact that the difference between the groups was not significant at 6 kHz and that listeners such as H and D had thresholds well below the mean at 1 kHz and well above it at 6 kHz makes it difficult to account for the present data simply in terms of Witton *et al.*'s suggestion that dyslexics have difficulty processing dynamic stimuli. The four dyslexic listeners who performed worse in the DLF task also performed relatively poorly at at least one frequency in the FMDL task.

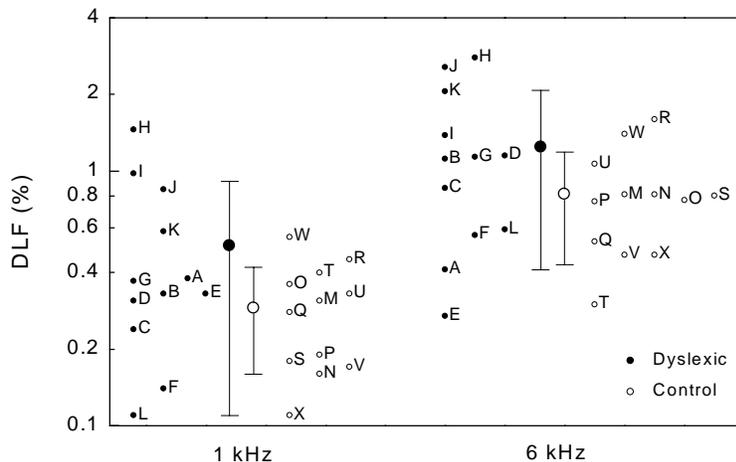


Fig. 1. Frequency difference limens (DLFs) at 1 and 6 kHz.

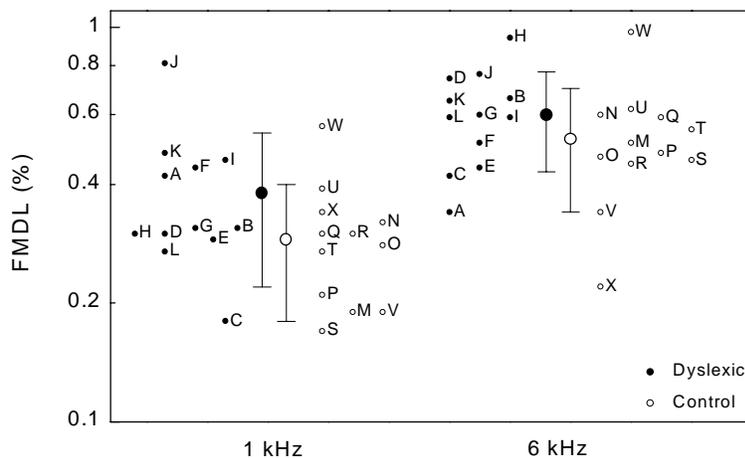


Fig. 2. Frequency modulation detection limens (FMDLs) at 1 and 6 kHz.

4. Experiment 2: MLDs

4.1. Stimuli and Equipment

The signal was a 200 Hz pure tone having a total duration of 200 ms. The signal, when presented, was gated on 200 ms into a 500 ms low-pass Gaussian noise masker having a cut-

off frequency of 1000 Hz. Both signal and masker durations included 50-ms cosine-squared onsets and offsets. The overall level of the noise was 77 dB SPL. In the NOS0 condition, both the noise and signal were identical in the two ears whereas in the NOS π condition, the signal was presented with an interaural phase difference of π radians.

Signal and masker were synthesized in real-time at a sampling rate of 44.1 kHz using custom software running on an IBM-compatible PC. The resulting waveforms were converted to voltages by means of a 24-bit audio card (LynxONE). The level of the stimuli was controlled in software. Stimuli were presented binaurally over Sennheiser HD414 headphones. Listeners were run individually in a sound-attenuating enclosure and responded using a standard computer keyboard.

4.2. Procedure

Thresholds for detecting the 200 Hz tone were measured in two conditions: with the tone and noise identical in the two ears (NOS0), and with the tone inverted in one ear (NOS π). The binaural masking level difference is the difference in threshold between these two conditions. Thresholds were determined using a two-interval, two-alternative, forced-choice procedure. The first observation interval began 500 ms after the offset of a 200 ms visual alerting stimulus, with a 500 ms silence separating intervals one and two. Listeners were given an unlimited time in which to respond. Immediately after responding they were presented with feedback for 400 ms. The next trial began 1s after the termination of feedback.

The level of the tone was adjusted adaptively using a 2-down, 1-up rule that targeted the level corresponding to 70.7% correct responses. The initial step size was 4 dB; this was reduced to 2 dB following the first four reversals. Each 50-trial run began with the level of the tone set to 72 dB SPL. Data were collected over two identical blocks with four runs (two for each condition) completed per block. The four runs were randomly interleaved with the signal level determined separately for each run. During each block, the experiment was paused for a minimum of 30 seconds after the 67th and 134th trial and restarted by a keypress. Each block lasted an average of 15 minutes. Prior to data collection, all listeners were presented with a few trials in the presence of the experimenter to ensure that they were familiar with the task and stimuli. Thresholds for each run were defined as the average signal level across the even number of reversals remaining after the first three or four reversals were discarded. The average number of reversals contributing to the determination of threshold exceeded 10.

4.3. Results and Discussion

Table 2 shows the mean signal level at threshold and the standard deviation of the mean for the NOS0 and NOS π conditions. The estimated binaural masking level difference for both groups was around 18 dB with no significant difference between groups in either the NOS0 or NOS π conditions (one-tailed t-test, $p > 0.05$).

Table 2. Threshold signal levels (dB SPL)

Group	NOS0	NOS π	MLD (dB)
Dyslexic	64.9 (2.0)	47.0 (1.7)	17.9
Control	63.9 (1.8)	45.9 (2.0)	18.0

The fact that MLDs for the two groups did not differ contrasts with the findings of McAnally and Stein (1996), who found that for their dyslexic group, MLDs were about 4 dB lower than controls. One difference between the studies is the choice of signal frequency. In McAnally and Stein's study the signal frequency was 1000 Hz, resulting in an average MLD of around 8 dB for the control group. A lower frequency signal was used here in an effort to

increase the magnitude of the MLD. We are currently investigating the possibility that dyslexic listeners show a more rapid decline in sensitivity to interaural timing differences with increasing frequency than normal listeners.

5. Summary and conclusion

DLFs did not differ significantly between the dyslexic and control group, although there was a small subgroup of dyslexic listeners with high thresholds at both 1 and 6 kHz. The fact that the same listeners were impaired at both frequencies suggests that the threshold elevation was not due simply to an inability to process temporal fine structure information. FMDLs at 1 kHz were significantly larger for the dyslexic group when outliers were removed. However, the small magnitude of the difference and the fact that the difference was not significant at 6 kHz indicate that more data are required before this result can reliably be attributed to dyslexic listeners' difficulty with dynamic stimuli. MLDs were almost identical for the two groups, suggesting that, at least at low frequencies, dyslexics are not impaired in their ability to exploit interaural temporal differences. The present data provide little support for the hypothesis that the core phonological deficit in dyslexia can be traced to a low-level auditory deficit. As a further test of this hypothesis, JM, a severe phonological dyslexic, whose development has been followed longitudinally (e.g., Hulme and Snowling, 1992) was tested using the present protocol. Only in the 6 kHz DLF condition did his threshold exceed the mean for the control group, and, even then, his performance on 2 of the 4 runs was normal, perhaps suggesting he had experienced difficulty settling in to the task.

Acknowledgments

We are grateful to Dr. Kate Nation for assistance with data collection. This work was supported by The Wellcome Trust and the University of York's Innovation and Research Priming Fund.

References and links

- Baldeweg, T., Richardson, A., Watkins, S., Foale, C., and Gruzelier, J. (1999). "Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials," *Ann. Neurol.* **45**, 495-503.
- Bruck, M. (1992). "Persistence of dyslexics' phonological awareness deficits," *Dev. Psychol.* **28**, 874-886.
- Dougherty, R. F., Cynader, M. S., Bjornson, B. H., Edgell, D., and Giaschi, D. E. (1998). "Dichotic pitch: a new stimulus distinguishes normal and dyslexic auditory function," *NeuroReport* **9**, 3001-3005.
- Hari, R., and Kiesila, P. (1996). "Deficit of temporal auditory processing in dyslexic adults," *Neurosci. Lett.* **205**, 138-140.
- Hulme, C., and Snowling, M. (1992) "Deficits in output phonology: an explanation of reading failure?" *Cognitive Neuropsych.* **9**, 47-72.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 167-177.
- McAnally, K. I., and Stein, J. F. (1996). "Auditory temporal coding in dyslexia," *Proc. R. Soc. Lond. B.* **263**, 961-965.
- McAnally, K. I., and Stein, J. F. (1997). "Scalp potentials evoked by amplitude-modulated tones in dyslexia," *J. Speech Lang. Hear. Res.* **40**, 939-945.
- Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing* (Academic, New York), 4th ed.
- Pennington, B. F., McCabe, L. L., Smith, S. D., Lefly, D. L., Bookman, M. O., Kimberling, W. J., and Lubs, H. A. (1986). "Spelling errors in adults with a form of familial dyslexia," *Child Dev.* **57**, 1001-1013.
- Snowling, M., Nation, K., Moxham, P., Gallagher, A., and Frith, U. (1997). "Phonological processing skills of dyslexic students in higher education: a preliminary account," *J. Res. Read.* **20**, 31-41.
- Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., Stein, J. F., and Green, G. G. R. (1998). "Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers," *Curr. Biol.* **8**, 791-797.
- Wolff, P. H., and Melngailis, I. (1994). "Family patterns of developmental dyslexia: Clinical findings," *Am. J. Med. Genet.* **54**, 122-131.

Vibration of beads placed on the basilar membrane in the basal turn of the cochlea

Nigel P. Cooper

*Department of Physiology, University of Bristol, Bristol, BS8 1TD, United Kingdom.
nigel.cooper@bristol.ac.uk*

Abstract: Interferometric recordings of sound-evoked vibrations in the basal turn and hook regions of the guinea-pig cochlea are used to show that reflective microbeads (i) follow the motion of the structures on which they are placed, and (ii) do not affect this motion dramatically. Extrapolating these findings to other types of reflective or radioactive material lends support to the findings of numerous studies of cochlear mechanics.

©1999 Acoustical Society of America

PACS numbers: 43.64.Kc, 43.64.Nf, 43.64. Pg, 43.64.Yp

1. Introduction

Experimental studies of cochlear mechanics have provided insight into the manner in which sounds are processed in the peripheral auditory system. Most of the observations in these studies have been made at the level of the basilar membrane, and almost all have been facilitated by the placement of small objects (either radioactive sources, reflective microbeads, or mirrors) on the cochlear partition. The effects that these objects might have on the vibrations of the cochlear partition have often been questioned, but have very rarely been tested [Kliauga and Khanna, 1983; Sellick *et al.*, 1983].

One recent report has highlighted the potential problems of using reflective microbeads in studies of cochlear mechanics; Khanna *et al.* (1998) reported that beads that had been placed on the basilar membrane did not follow the motion of the cochlear partition. Moreover, the presence of the beads affected the motion of the underlying partition quite dramatically. These observations were made in the apical turns of the cochlea in an isolated temporal bone preparation, but there is little reason to prevent their extrapolation to other regions of the cochlea. Such an extrapolation would clearly cast doubt on the relevance of almost all other observations in the field of cochlear mechanics.

The present study was designed to investigate the issue of using beads in the more basal turns of the living cochlea, where the vast majority of previous observations have been made. The principle aims of the study were to find out whether the reflective microbeads follow the motion of the basilar membrane, and/or whether they affect this motion.

2. Methods

2.1 Surgical preparation

Ten young, pigmented guinea-pigs provided substantial data in this study. The animals were anesthetized using a combination of pentobarbital sodium (25-30 mg/kg i.p.) and Hypnorm (6 mg/kg fluanisone + 200 µg/kg fentanyl i.m.). Supplementary doses of each agent were given as needed to maintain a state of deep areflexia, and the animals were overdosed with pentobarbital on completion of the experimental procedures. Tracheotomies were performed, and end-tidal CO₂ concentrations were monitored using an infra-red gas detector. Artificial ventilation was provided when necessary. Core temperatures were maintained near 37.6°C using a thermostatically-controlled heating blanket monitored by a rectal probe thermistor.

The animal's scalp and right pinna were retracted, and the skull was fixed into a combined head-holder and earpiece. The postero-lateral bulla was opened to gain access to the cochlea, and a small opening was made into the scala tympani of the cochlea's basal turn. The opening was made either (i) by gently shaving through the bone ~3.5 mm from the basal end of the cochlear partition (n=1), (ii) by picking through the bone from the apical border of the round window ~2 mm from the basal end of the partition (n=2), or (iii) by carefully tearing through the round window membrane between 1 and 2 mm from the basal end of the partition (n=7). The temperature inside the bulla was maintained near 37.6°C using heat from a thermostatically-controlled light bulb.

Gold-coated polystyrene microbeads (15 or 25 μm diameter, mean density range 1.1 to 1.5 $\text{g}\cdot\text{cm}^{-3}$) were introduced onto the perilymphatic meniscus using a stainless-steel pick. The beads were manipulated to fall through the meniscus and onto the underlying basilar membrane. Once an acceptable bead placement had been made, the perilymphatic meniscus was covered with a small glass cover-slip. The cover-slip minimized any optical interference caused by movements of the meniscus, but it did not form an acoustic seal [Cooper and Rhode, 1992].

2.2 Stimulus generation and control

Stimuli were generated using a computer-controlled DAC and attenuators, with a reverse-driven condenser microphone cartridge serving as a loudspeaker. Multiple presentations of a 30-ms-duration tone-pip were made with repetition periods of >100 ms. The tone-pips were gated on and off using half-periods of a raised-cosine envelope (1-ms duration).

Stimuli were coupled into the dissected ear canal through an earpiece that formed part of the animal's head-holder. Closed-field sound pressures were monitored less than 1 mm from the tympanic membrane using a condenser microphone equipped with a calibrated 1-mm-diameter probe tube. All experiments were performed on a vibration-isolated workstation in a soundproof chamber.

2.3 Measurement techniques

Vibrations were measured using displacement-sensitive heterodyne laser interferometry [Cooper, 1999a]. This technique has been shown to be linear over a wide dynamic range (object velocities from 0 to 13 $\text{mm}\cdot\text{s}^{-1}$) and was sensitive enough to measure low-level displacements from the untreated cochlear partition (the system's noise-floor was ~10 $\text{pm}\cdot\text{Hz}^{-0.5}$ for objects that reflected just 1 ppm of the incident light). In the present study, displacement responses were averaged across 8-64 presentations of the 30-ms-long stimuli. The steady-state portions of the averaged responses (between 1 and 29 ms peri-stimulus-time) were Fourier analyzed to determine the magnitude and phase of the response component at the stimulus frequency.

Compound action potentials (CAPs) were used to monitor the physiological condition of the cochlea. These were recorded from a silver wire electrode placed in the round window niche. CAPs were amplified (x1000) and band-pass filtered (0.1-10 kHz) before being displayed on an oscilloscope and digitized and averaged in a computer. Baseline CAP audiograms [Johnstone *et al.*, 1979] were determined prior to opening the cochlea in each experiment. Subsequent checks on the CAPs were made whenever there was reason to suspect a change in the physiological condition of the cochlea.

3. Results

Comparisons were made between the responses observed (i) on and (ii) within 20-190 μm of individual microbeads in six living cochleae and two dead cochleae. The responses from each site were stimulus-dependent and varied with the physiological condition of the

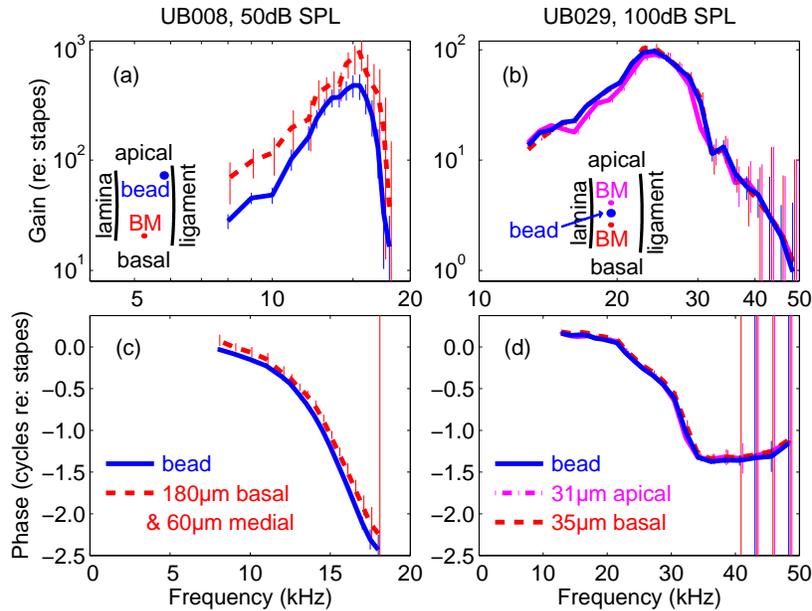


Fig. 1. Tuning characteristics measured on and around individual microbeads in the basal turn and hook regions of the guinea-pig cochlea. Response amplitudes (a,b) and phases (c,d) are expressed with respect to those measured from the stapes in each experiment. Solid lines show bead responses; dashed lines show responses measured from adjacent sites on the cochlear partition. Vertical bars indicate 95% confidence intervals. Bead locations were ~ 3.5 and ~ 1.8 mm from the basal end of the cochlear partition, ~ 140 and 56 μm from the osseous spiral lamina, in UB008 and UB029, respectively. Relative locations of other recording sites are shown in the insets of (a) and (b) and specified in (c) and (d).

preparation, but the comparisons between the responses on and around the beads were always similar. Typical results are illustrated in Fig. 1: in each panel of this figure, the tuning curves observed on and around the beads have very similar shapes. To be more specific, more than 90% of the variance observed in the bead responses could be accounted for by the variance observed at adjacent sites on the cochlear partition. (This statement applies to the findings in every experiment.) Systematic differences between the absolute sensitivities of the individual curves were often observed (the largest difference is illustrated in Fig. 1a), but these could always be accounted for by differences in the radial locations of the recording sites [Cooper, 1999b]. On three occasions, there were also systematic differences between the response phases observed on and around the beads (the largest of these is illustrated in Fig. 1c). These phase differences were always consistent with differences in the longitudinal locations of the sites studied, assuming that the displacement waves travel from the base to the apex of the cochlea. In general, the closest matching responses on and around the beads were observed when the distance between the two recording sites was minimized and the recordings were made at similar radial positions (cf. Figs. 1 b, d).

Attempts to make recordings before and after placing single reflective microbeads in individual preparations were unsuccessful (the final locations of the microbeads were too difficult to predict). However, two lines of indirect evidence suggest that the actual placement of the microbeads had little effect on the mechanics of the cochlear partition. First, as shown in Fig. 2, comparisons between preparations that were known to be in good physiological condition revealed very similar characteristics when microbeads were used

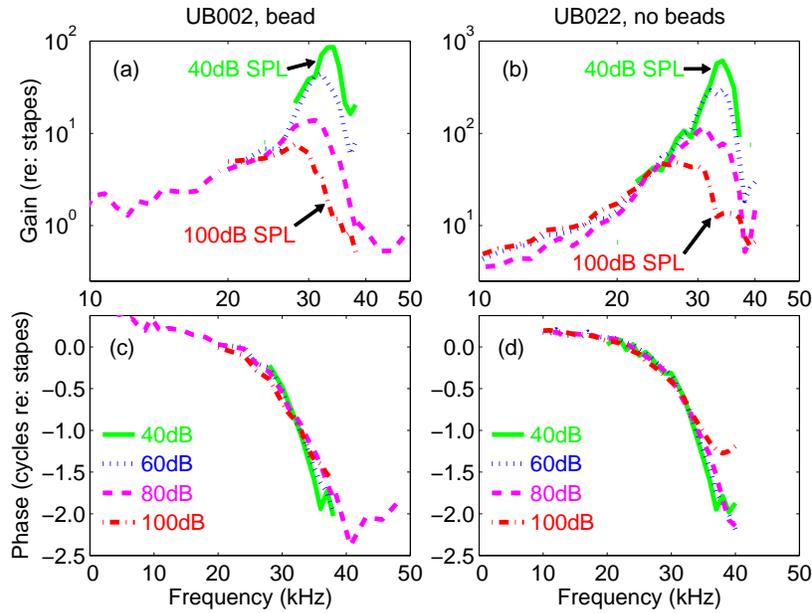


Fig. 2. Level-dependent tuning characteristics measured with (a,c) and without (b,d) the use of microbeads in two separate experiments. Sound pressure levels (dB re: 20 μ Pa) ranged from 40 to 100 dB in 20 dB steps. Recording were made \sim 1.5 mm from the basal end of the cochlear partition, \sim 20 and \sim 60 μ m from the osseous spiral lamina in UB002 and UB022, respectively.

(Figs. 2 a, c) and were not used (Figs. 2 b, d). Once again, there were differences (in this case on the order of 20 dB) between the absolute sensitivities of the different preparations, but these could probably be accounted for by interanimal variability (including variations in the radial locations of the recording sites on the basilar membrane). The most important features to note in Fig. 2 are the similarities between the *shapes* of the tuning curves and their nonlinear variation with sound pressure level. These similarities are consistent with the suggestion that the beads have little effect on the ‘normal’ operation of the cochlea (see discussion).

The second line of evidence that the beads had little effect on the operation of the cochlea is illustrated in Fig. 3: very little difference (typically $< \pm 3$ dB) was observed between the CAP thresholds recorded before and after placement of the beads on the basilar membrane. The supra-threshold amplitudes of the CAP and other potentials (e.g., summing potentials) often varied more than the CAP thresholds themselves (although not in the example illustrated in Fig. 3). The most likely reason for this is that the current paths in the cochlea changed when the cochlea was opened and/or the perilymph was manipulated.

An overall comparison between the sharpness of tuning observed with and without the use of reflective microbeads is shown in Fig. 4. The Q_{10dB} values in this figure quantify tuning sharpness by dividing each preparation’s peak frequency by its bandwidth at points 10dB below the peak. The fact that the reference line (showing $Q_{10dB \text{ bead}} = Q_{10dB \text{ BM}}$) intersects all of the error bars in Fig. 4 indicates that the data are consistent with a hypothesis that motion of the beads follows that of the basilar membrane.

4. Discussion

The results of this investigation indicate that reflective microbeads (i) follow the motion of the structures on which they are placed, and (ii) do not affect this motion dramatically. These

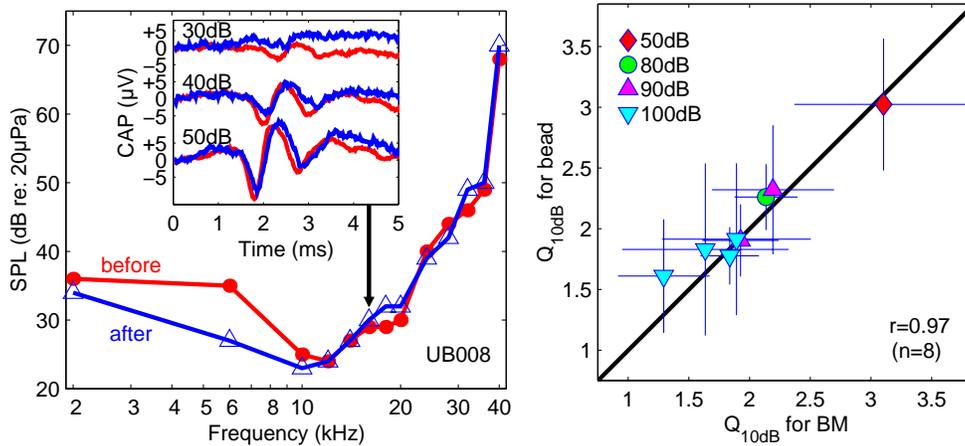


Fig. 3 (left). CAP audiograms and responses to 16 kHz tones (inset) before and after placing a microbead ~3.5 mm from the basal end of the cochlear partition (cf. Fig. 1 a, c).

Fig. 4 (right). Sharpness of tuning on the basilar membrane (BM) and on individual microbeads in 8 experiments. Q_{10dB} values quantify the sharpness of tuning by dividing peak frequencies by the bandwidths 10 dB below the peak. Error bars depict 95% confidence intervals (cf. Fig. 1). Sound pressure levels (see key) are expressed in dB re: 20 μPa.

findings lend direct support to the findings of previous studies that have used similar types of microbead [e.g., Cooper and Rhode, 1997]. The consistencies between these studies and studies that have used different types of microbead [e.g., Ruggero *et al.*, 1997], or even radioactive sources [e.g., Rhode, 1971; Sellick *et al.*, 1982], allow this support to be extended to numerous previous observations. Perhaps the most important of these observations are that basilar membrane responses are sharply tuned, nonlinear, and physiologically vulnerable.

The major weakness of the data in this report is the lack of ‘before and after’ observations from individual preparations that were known to be in good physiological condition. This is unfortunate, and may well be overcome in subsequent experiments. In the meantime, the strongest evidence that the beads have little effect even under near-optimal physiological conditions (Figs. 2 and 3) is indirect. It should be noted that the data of Fig. 2 (either with or without beads) demonstrate much more nonlinearity and better tuning than those in previous studies of the hook region basilar membrane [e.g., Cooper and Rhode, 1992]. One reason for this is that these preparations were in much better condition than those reported previously; their CAP thresholds of 35-40 dB SPL at ~34 kHz were 10-15 dB lower than those observed in most experiments (cf. Fig. 3). The fact that the beads led to little or no deterioration in the CAP thresholds when the thresholds were so low to begin with is the strongest evidence to date that the beads have little effect on the operation of the cochlea.

4.1 Comparison with previous studies

The conclusions of the present study diametrically oppose those of the only other investigation of this kind to be performed to date [Khanna *et al.*, 1998]. Various methodological differences may mean that the two sets of conclusions are not actually inconsistent with one another, however. In the view of the present author, the most important of these differences are likely to relate to the longitudinal positions of the sites studied. This issue will be expanded below. Other methodological differences, such as the

exact type of microbead used, the orientation of the recording sites when the vibration measurements are made, and the physiological condition of the preparations are likely to be less important. (This statement is based on the author's own observations using various types of microbead in a wide range of cochlear preparations.)

The most obvious difference between the studies of Khanna *et al.* (1998) and those reported here is that they were performed in different parts of the cochlea. This may have profound consequences. The placement of beads in the apical turn of the cochlea involves contact with the luminal surfaces of specialized epithelial cells known as Claudius cells. Both Khanna *et al.* (1998) and Rhode and Cooper (1996) have commented on the difficulty of getting a microbead to 'stick' to these cells. In contrast, the placement of beads in the more basal turns of the cochlea involves contact with cells that line the tympanic face of the basilar membrane. Anecdotal evidence suggests that these cells provide much more grip for the microbeads. It is rare for a bead to slip or drift away from its initial landing place in the basal turns of the cochlea, whereas it is rare for one *not* to run away in the apical turns. Another difference between the approaches to the basilar membrane in the apical and basal turns of the cochlea concerns optical access. It is relatively simple to obtain near perpendicular access to the basilar membrane in the basal turn of the cochlea, whereas perpendicular access in the apex is hindered by the highly reflective surfaces of the lipid droplets in the Hensen's cells just above the basilar membrane. The presence and orientation of the bony septum just below the basilar membrane can also be problematical in the apical turn of the cochlea. These features make unambiguous recordings of basilar membrane motion very difficult to achieve (without the use of microbeads) in the apical turn of the cochlea, and prevent the present author from making direct comparisons with the data of Khanna *et al.* (1998).

Acknowledgments

Supported by The Royal Society and The Wellcome Trust.

References and links

- Cooper, N.P. (1999a). "An improved heterodyne laser interferometer for use in studies of cochlear mechanics," *J. Neurosci. Meth.* **88**, 93-102.
- Cooper, N.P. (1999b). "Radial variation in the vibrations of the cochlear partition," in *Recent Developments in Auditory Mechanics*, edited by H. Wada and T. Takasaka (World Scientific, Singapore).
- Cooper, N.P., and Rhode, W.S. (1992). "Basilar membrane mechanics in the hook region of cat and guinea-pig cochleae: sharp tuning and nonlinearity in the absence of baseline position shifts," *Hear. Res.* **63**, 163-190.
- Cooper, N.P., and Rhode, W.S. (1997). "Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea," *J. Neurophys.* **78**, 261-270.
- Johnstone, J.R., Alder, V.A., Johnstone, B.M., Robertson, D., and Yates, G.K. (1979). "CAP threshold and single unit thresholds," *J. Acoust. Soc. Am.* **65**, 254-257.
- Khanna, S.M., Ulfendahl, M., and Steele, C.R. (1998). "Vibration of reflective beads placed on the basilar membrane," *Hear. Res.* **116**, 71-85.
- Kliauga, P., and Khanna, S.M. (1983). "Dose rate to the inner ear during Mossbauer experiments," *Phys. Med. Biol.* **28**, 359-366.
- Rhode, W.S. (1971). "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mossbauer technique," *J. Acoust. Soc. Am.* **49**, 1218-1231.
- Rhode, W.S., and Cooper, N.P. (1996). "Nonlinear mechanics in the apical turn of the chinchilla cochlea *in vivo*," *Auditory Neuroscience* **3**, 101-121.
- 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.
- Sellick, P.M., Patuzzi, R.B., and Johnstone, B.M. (1982). "Measurement of basilar membrane motion in the guinea pig using the Mossbauer technique," *J. Acoust. Soc. Am.* **72**, 131-141.
- Sellick, P.M., Yates, G.K., and Patuzzi, R. (1983). "The influence of Mossbauer source size and position on phase and amplitude measurements of the guinea pig basilar membrane," *Hear. Res.* **10**, 101-108.

Effect of acoustic dynamic range on phoneme recognition in quiet and noise by cochlear implant users

Qian-Jie Fu and Robert V. Shannon

*Department of Auditory Implants and Perception, House Ear Institute,
2100 West Third Street, Los Angeles, CA 90057
qfu@hei.org shannon@hei.org*

Abstract: The present study measured phoneme recognition in cochlear implant users when the dynamic range of the input speech signals was reduced by either peak clipping or center clipping. In quiet, reducing the acoustic dynamic range to 30 dB still provides sufficient speech information for phoneme identification in cochlear implant users with 4-channel continuous interleaved sampler strategy and normal-hearing listeners listening to correspondingly degraded speech. Phoneme recognition decreased when the dynamic range was reduced below 30 dB: peak clipping was more detrimental to vowel recognition and center clipping was more detrimental to consonant recognition. However, in background noise, center clipping produced a small increase in speech recognition for cochlear implant listeners.

©1999 Acoustical Society of America

PACS numbers: 43.71.Es, 43.71.Ky, 43.66.Ts

1. Introduction

In cochlear implant speech processor design, one of the major concerns is to properly transform acoustic amplitudes to electric currents. Normal acoustic hearing can process sounds over a range of 120 dB, and instantaneous amplitudes in normal speech cover a 30 to 60 dB range [Boothroyd *et al.*, 1994]. However, implant listeners typically have dynamic ranges of only 6 to 15 dB in electric current, requiring the larger acoustic range to be compressed into the smaller electric range. The mapping between acoustic amplitudes and electric currents is affected by the range of acoustic amplitudes, the range of electric currents, and the mapping function. In general, the range of electric currents is determined by the clinically measured threshold (T-level) and comfortable loudness level (C-level) for each individual cochlear implant user, which is then fixed for each subject. Once the electrical dynamic range is fixed, the mapping is completely determined by the range of acoustic amplitudes and the mapping function.

Fu and Shannon [1998] systematically investigated the effects of a power-law mapping function on vowel and consonant recognition in both cochlear implant users and normal-hearing listeners. In their study, the range of acoustic amplitudes was fixed at 40 dB, and the range of electric currents was predetermined by the psychophysically measured T-level and C-level for each subject. Because the exponent of the power-law mapping function was varied, they found that, for normal-hearing listeners, the best performance was obtained, as expected, when the normal loudness growth was preserved, which was obtained with a linear mapping ($p=1.0$). However, for cochlear implant users, the best performance was achieved with a compressive exponent ($p=0.2$). This exponent is consistent with the restoration of normal loudness growth in electrical stimulation. Performance deteriorated only slightly in both acoustic and implant listeners when the mapping function was either more

compressive or less compressive than the one that preserved the normal loudness growth. Fu and Shannon concluded that the mapping function has only a minor effect on phoneme recognition in quiet. Similar results were reported by Zeng and Galvin [1999].

Once the electric dynamic range and the loudness exponent have been fixed, the transformation from acoustic amplitude to electric current is dependent on the input acoustic dynamic range. The long-term average speech spectrum has been widely used as a basis to estimate the dynamic range of acoustic amplitudes. Based on a combination of statistical and behavior measures, the dynamic range of speech is usually estimated to be 30 dB. However, when inter-talker differences are considered, the overall dynamic range of acoustic amplitudes can be as high as 53 dB [Boothroyd *et al.*, 1994]. It is well established [e.g., Licklider and Pollack, 1948; Drullman, 1995] that speech recognition is highly robust to peak and center clipping when full spectral cues are available. It is theoretically as well as practically important to understand whether the same robustness occurs with reduced spectral resolution, or whether a larger dynamic range of acoustic amplitudes is required for speech recognition to compensate for the loss of spectral resolution. In the present study, phoneme recognition was measured as a function of the dynamic range of acoustic amplitudes both in cochlear implant users with 4-channel continuous interleaved sampler strategy and in normal-hearing subjects listening to the spectrally degraded speech processed by a 4-channel noise vocoder [Shannon *et al.*, 1995]. Both peak clipping and center clipping were used to limit the dynamic range to explore potential differential effects of dynamic range on strong vowels and weak consonants.

Cochlear implant users are highly sensitive to background noise due to the reduced spectral resolution inherent to implant devices. Theoretically, spectral resolution could be improved by increasing the number of electrodes, but relatively simple amplitude manipulations might also improve performance in noise. Fu and Shannon [1999] found that when background noise was introduced the effect of altering the exponent of the amplitude mapping function was dramatic and asymmetric. Performance declined mildly in noise with expansive mappings but dramatically in noise with compressive mappings. Fu and Shannon concluded that, although an expansive mapping distorts the loudness growth, it increases the effective S/N ratio by mapping the low-amplitude noise to low electrical levels near threshold. A similar effect could be achieved by center clipping. Although center clipping will remove some speech information, this loss might be more than compensated by the effective increase in SNR levels. The present study measured the trade-off between speech information and S/N ratio in three cochlear implant listeners with 4-channel continuous interleaved sampler (CIS) strategy. Vowel and consonant recognition were measured as a function of the amount of center clipping under three different S/N ratios.

2. Methods

2.1 Subjects

Three cochlear-implant (CI) users and four normal-hearing (NH) listeners participated in this experiment. All were native speakers of American English. The NH subjects, aged 25 to 35, had thresholds better than 15 dB HL at audiometric test frequencies from 250 to 8000 Hz. Cochlear implant subjects were three postlingually deafened adults using the Nucleus-22 device. All had at least four years experience utilizing the SPEAK speech processing strategy and all were native speakers of American English. The Nucleus processor with the SPEAK strategy divides the input acoustic signal into 20 frequency bands, extracts the amplitude envelope from each band, and stimulates the electrodes corresponding to the 6 to 10 bands with the maximum amplitudes [McDermott *et al.*, 1992]. The frequency allocation table specifies the frequency range covered by the speech processor. Two subjects (N4 and N7) used table 9 (150-10,823 Hz), and one subject (N3) used table 7 (120-8,658 Hz). All implant subjects had 20 active electrodes available for use. Table 1 contains relevant information for

the three subjects, including their most recent scores on the HINT sentence test, multitalker 12-vowel and 16-consonant identification test with their clinical SPEAK processor.

Table 1. Subject information on three Nucleus-22 cochlear implant listeners in present study

Subject	Age	Gender	Cause of Deafness	Duration Of use	Freq. Table	Score (HINT)	Vowel Score	Cons. Score
N3	56	M	Trauma	7 years	7	96.2%	69.5%	51.8%
N4	40	M	Trauma	5 years	9	100.0%	81.1%	71.8%
N7	55	M	Unknown	5 years	9	100.0%	64.5%	76.8%

2.2 Test materials and procedures

Speech recognition was assessed for medial vowels and consonants. Vowel recognition was measured in a 12-alternative identification paradigm, including 10 monophthongs (/i ɪ e æ ɑ ɔ u ʌ ɜ/) and 2 diphthongs (/e o/), presented in a /h/-vowel-/d/ context. The tokens for these closed-set tests were digitized natural productions from 5 men, 5 women, and 5 children drawn from the material collected by Hillenbrand *et al.* [1995]. Consonant recognition was measured in a 16-alternative identification paradigm for the consonants /b d g p t k l m n f s ʃ v z θ dʒ/ presented in an /a/-consonant-/a/ context. Two exemplars of each of the 16 consonants were produced by three speakers (1 male, 2 female) for a total of 96 tokens (16 consonants * 3 talkers * 2 exemplars).

Each test block included 180 tokens (12 vowels * 15 talkers) for vowel recognition or 192 tokens (16 consonants * 3 talkers * 2 exemplars * 2 repeats) for consonant recognition. A stimulus token was randomly chosen from all 180 tokens in vowel recognition and from 192 tokens in consonant recognition and presented to the subject. Following the presentation of each token, the subject responded by pressing one of 12 buttons in the vowel test or one of 16 buttons in the consonant test, each marked with one of the possible responses. The response buttons were labeled in a /h/-vowel-/d/ context (heed, hawed, head, who'd, hid, hood, hud, had, heard, hoed, hod, hayed) for the vowel identification task and a /a/-consonant-/a/ context following an example word for the consonant identification task.

2.3 Signal processing

The speech signal was mixed with simplified speech spectrum-shaped noise (constant spectrum level below 800 Hz and 10-dB/octave roll-off above 800 Hz). The signal-to-noise ratio (S/N) was defined as the difference in decibels between the root-mean-square (RMS) levels of the whole speech token and the noise.

In electric hearing, all signals were processed by a custom 4-channel continuous interleaved sampler (CIS) speech processor [Wilson *et al.*, 1991] and presented at comfortable audible levels through a custom implant interface system [Shannon *et al.*, 1990]. The 4-channel CIS processor was implemented as follows. The signal was first pre-emphasized using a first-order Butterworth high-pass filter with a cutoff frequency of 1200 Hz and then band-pass filtered into four broad frequency bands using 8th-order Butterworth filters. The five corner frequencies of the four bands were at 300 Hz, 713 Hz, 1509 Hz, 3043 Hz, and 6000 Hz. The envelope of the signal in each band was extracted by half-wave rectification and low-pass filtering (8th-order Butterworth) with a 160 Hz cutoff frequency. For each S/N level, the amplitude histogram in each band was computed for the test materials presented at 70 dB SPL. The maximum amplitude used in the acoustic range (A_{max}) was set to the 99th percentile of all amplitude levels in all channels. The minimum amplitude (A_{min}) used in the control condition was set to $0.01 * A_{max}$ so that the dynamic range was fixed at 40 dB.

The dynamic range was reduced by either center clipping or peak clipping. Center clipping was implemented by setting the maximum amplitude to A_{max} and changing the

minimum amplitude from $0.01 \cdot A_{\max}$ in the control condition to $0.4 \cdot A_{\max}$ ($0.2 \cdot A_{\max}$ for consonant recognition in quiet) in six steps. Correspondingly, the input dynamic range was gradually reduced from 40 dB to 8 dB (14 dB for consonant recognition in quiet). Peak clipping was implemented by setting the minimal amplitude to A_{\min} and changing the maximum amplitude from A_{\max} to $2.5 \cdot A_{\min}$, resulting in a reduction of the input dynamic range from 40 dB to 8 dB. The current level (E) of electric stimulation in the i^{th} band was set to the acoustic envelope value (A) raised to a power. The exponent of the power function was set to 0.2 (Fu and Shannon, 1998) regardless of the acoustic dynamic range. This transformed amplitude was used to modulate the amplitude of a continuous, 500-pulse/sec. biphasic pulse train with a 100- μs /phase pulse duration. The stimulus order of the 4 channels was 1-3-2-4 for electrode pairs (18,22), (13,17), (8,12), and (3,7), respectively.

In acoustic hearing, the speech signal was spectrally degraded using a four-band modulated noise processor [Shannon *et al.*, 1995]. Envelope extraction in each band was the same as in electric stimulation. The following manipulation of the input dynamic range was similar to that in electric stimulation. The output dynamic range was fixed at 40 dB, and a linear transformation was applied between the input and output amplitudes. Then, the envelope waveform in each band was used to modulate wideband noise that was subsequently spectrally limited by a bandpass filter with the same characteristics used for the analysis filter band. The outputs from all modulated noise bands were then summed, and the processed speech tokens were equated in terms of RMS energy. All processed speech stimuli were stored on computer disk and presented via custom software to a 16-bit D/A converter (TDT DD1) at a 16-kHz sampling rate. Stimuli were presented to the listeners through Sennheiser HDA200 headphones at 70 dB on an A-weighted scale.

All subjects were well familiarized with the processed speech provided by the speech processor, the test materials, and the test procedure from prior experiments. The order of the clipping level conditions and of the vowel and consonant tests were counterbalanced across subjects. Within each test, speech stimuli were presented in random order, and test conditions were pseudo-randomized for each subject. No feedback was provided.

3. Results and discussion

Figure 1 shows the individual and mean scores of vowel and consonant recognition in quiet as a function of the range of acoustic amplitudes in both CI users and NH listeners. The open symbols show the individual scores, and the solid lines show the mean scores.

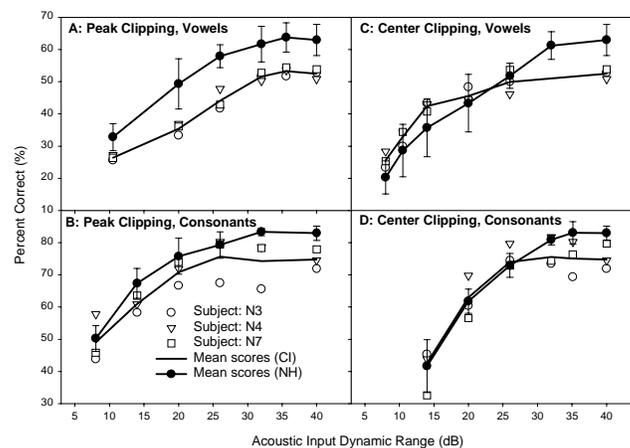


Fig. 1. Recognition scores of vowels and consonants as a function of acoustic input dynamic range in three cochlear implant users and four normal-hearing subjects. Recognition scores of: (A) peak-clipped vowels; (B) peak-clipped consonants; (C) center-clipped vowels; (D) center-clipped consonants.

Preliminary repeated measures analyses of variance for the overall performance used factors of clipping type, dynamic range, and subject. These revealed the significant effect of clipping type, dynamic range, and subject for all measures. There were no significant interactions between subject and dynamic range, and the subject factor was henceforth ignored. A second series of repeated measures ANOVAs tested the effect of dynamic range for peak clipping and center clipping separately. Peak clipping produced a significant effect on vowel recognition [$F(5,36)=16.84$, $p<0.001$] and consonant recognition [$F(5,36)=32.68$, $p<0.001$]. Post-hoc Bonferroni tests showed that vowel scores dropped significantly only when the acoustic dynamic range was 20 dB or lower. No significant drop in consonant recognition was observed until the acoustic dynamic range was reduced to 14 dB or lower. Similarly, the center clipping also produced a significant effect on vowel recognition [$F(5,36)=34.18$, $p<0.001$] and consonant recognition [$F(5,36)=59.92$, $p<0.01$]. Post-hoc Bonferroni tests showed that both vowel and consonant scores dropped significantly when the dynamic range was 20 dB or lower.

One important finding from these results is that, even with reduced spectral resolution in acoustic and electric hearing, a 30-dB input range provides sufficient speech information for phoneme identification. There is a differential effect of peak clipping and center clipping on vowel and consonant recognition. Vowels are more tolerant of center clipping, whereas consonants are more tolerant of peak clipping. This differential effect might be caused primarily by the different amplitude distribution of strong vowels and weak consonants within the dynamic range of speech.

Figure 2 shows the individual and mean vowel and consonant recognition scores in noise as a function of center clipping for cochlear implant listeners. The differences between the results in quiet (dashed lines) and in noise (solid and dotted lines) have important implications for signal processing for cochlear implants.

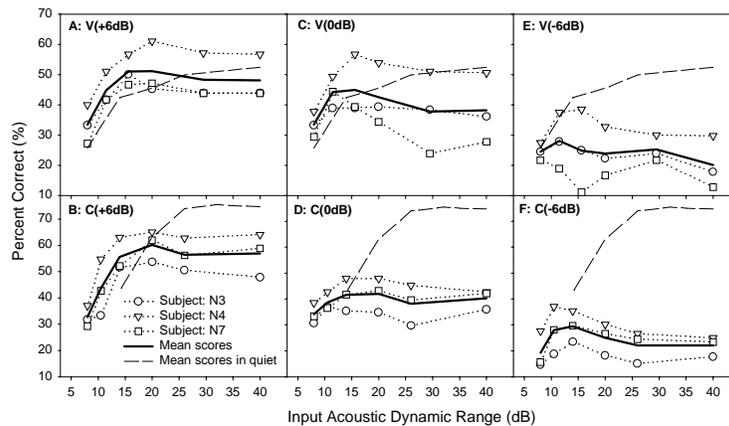


Fig. 2. Recognition scores of vowels and consonants as a function of the range of the input speech signals, produced by center clipping in three implant users. (A) Vowel scores in +6 dB S/N; (B) Consonant scores in +6 dB S/N; (C) Vowel scores in 0 dB S/N; (D) Consonant scores in 0 dB S/N; (E) Vowel scores in -6 dB S/N; (F) Consonant scores in -6 dB S/N.

Statistical analysis showed no significant performance drop (see Fig. 2, solid lines) as the acoustic dynamic range was reduced from 40 to 15 dB for all noise levels and test materials. In quiet conditions, a reduction of the dynamic range by center clipping results in the loss of speech information and causes a monotonic drop in phoneme recognition (see Fig. 2, dashed lines). However, in noisy conditions, a reduction of dynamic range by center clipping can also increase the effective S/N ratio by removing the low-amplitude noise components, similar to a squelch adjustment in radio transmission. This increase of the effective S/N ratio can compensate for the loss of speech information, resulting in a non-

monotonic function in performance (see Fig. 2, solid lines) as dynamic range was reduced. In all conditions the highest mean score was observed with a reduced acoustic dynamic range (typically below 20 dB) rather than at the largest 40-dB range. The amount of improvement (the highest score minus the score from a 40-dB range) was relatively small (2%-8%) and highly depended on the S/N levels (e.g. panel A vs. C), test materials (e.g., panel C vs. D) as well as subjects (e.g., panel E, N4 vs. N7). However, statistical analysis (student t-test) did reveal a significant improvement in vowel recognition at 0dB SNR (panel C) and in consonant recognition at -6dB SNR (panel F). Another benefit by center clipping is that all subjects reported an improvement of the sound quality and comfort of the processor, especially when listening at very low S/N levels (e.g. -6dB). The results suggest that cochlear implant users with the CIS strategy can significantly benefit from an appropriate reduction of the acoustic range by center clipping in a noisy environment.

Stochastic resonance (SR) might be another factor contributing to improved speech recognition when the acoustic input dynamic range was limited (e.g. < 20 dB). In SR, the addition of an appropriate amount of noise to the input of a nonlinear system can actually improve speech recognition [Morse and Evans, 1996]. In the present results, some center-clipping conditions (e.g., Fig. 2, panels A and B, dynamic range<15dB) produced significantly better performance in noise ($p<0.05$, student t-test) than the same dynamic range in quiet.

4. Conclusions

In summary, a 30-dB dynamic range is sufficient for phoneme recognition in both acoustic hearing and electric hearing, even with highly reduced spectral resolution. A further reduction of the acoustic range by center clipping can be beneficial when listening in noisy conditions and should be considered as a processing option for cochlear implants.

Acknowledgments

We wish to thank Prof. James Hillenbrand for allowing us to use the multitalker vowel test materials. The research was supported by grants from the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health.

References and links

- Boothroyd, A., Erickson, F. N., and Medwetsky, L. (1994). "The hearing aid input: a phonemic approach to assessing the spectral distribution of speech," *Ear and Hearing* **15**, 432-442.
- Drullman, R. (1995). "Temporal envelope and fine structure cues for speech intelligibility", *J. Acoust. Soc. Amer.*, **97**, 585-592.
- Fu, Q.-J., and Shannon, R. V. (1998). "Effects of amplitude nonlinearity on speech recognition by cochlear implant users and normal-hearing listeners," *J. Acoust. Soc. Am.* **104**, 2571-2577.
- Fu, Q.-J., and Shannon, R. V. (1999). "Phoneme recognition by cochlear implant users as a function of signal-to-noise ratio and nonlinear amplitude mapping," *J. Acoust. Soc. Am.* **106**, L18-L23.
- Hillenbrand J., Getty, L. A., Clark, M. J., and Wheeler K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**, 3099-3111.
- Licklider, J. C. R., and Pollack, I. (1948). "Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech," *J. Acoust. Soc. Am.* **20**, 42-51.
- McDermott, H. J., McKay, C. M., and Vandali, A.E. (1992). "A new portable sound processor for the University of Melbourne/Nucleus Limited Multichannel cochlear implant," *J. Acoust. Soc. Am.* **91**, 3367-3371.
- Morse, R. P., and Evans, E. F. (1996). "Enhancement of vowel coding for cochlear implants by addition of noise," *Nature Medicine* **2**, 928-932.
- Shannon, R. V., Adams, D. D., Ferrel, R. L., Palumbo, R. L., and Grandgenett, M. (1990). "A computer interface for psychophysical and speech research with the Nucleus cochlear implant," *J. Acoust. Soc. Am.* **87**, 905-907.
- Shannon, R. V., Zeng, F-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303-304.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (1991). "Better speech recognition with cochlear implants," *Nature* **352**, 236-238.
- Zeng, F.-G., and Galvin, J. (1999). "Amplitude mapping and phoneme recognition in cochlear implant listeners," *Ear Hear* **20**, 60-74.

Consonant recordings for speech testing

Robert V. Shannon^{1,2}, Angela Jensvold², Monica Padilla², Mark E. Robert¹ and Xiaosong Wang¹

¹House Ear Institute, 2100 W. Third St., Los Angeles, CA 90057 and

²Biomedical Engineering Dept., University of Southern California, Los Angeles, CA 90089

Shannon@hei.org

Abstract: Initial and medial consonants were recorded in three vowel contexts for use in speech recognition experiments. Five male and five female talkers were recorded producing the twenty-five consonants /b, d, g, p, t, k, m, n, ŋ, l, r, f, v, θ, ð, s, z, ʃ, tʃ, dʒ, ʒ, j, w, ʌ, h/ in medial (v/C/v) and initial (C/v) positions using vowels /a/ ("hod"), /i/ ("heed"), and /u/ ("who'd"). The sampling rate for these recordings was 44.1kHz. Representative tokens of each consonant were amplitude normalized to the steady-state portion of the vowel. Listening tests were conducted with normal-hearing listeners on a subset of twenty consonants in all three vowel contexts and in initial and medial positions. The results showed that the consonants were clearly recognized with only a few minor confusions, primarily between /v/ and /ð/. The full set of recordings is available for research use.

© 1999 Acoustical Society of America

PACS Numbers: 43.71.Es, 43.66.Yw

1. Introduction

Speech recognition tests are commonly used to assess performance with prosthetic devices (like cochlear implants and hearing aids) or in the assessment of communication channels. Depending on the number of conditions to be tested, these materials could be repeated to the same group of listeners hundreds of times. Sentences-length materials and even single words cannot be used in these testing situations because of the differences in linguistic difficulty and the limited number of test sets. Even phonemes with a limited number of exemplars or with single talkers are susceptible to incidental learning (Uchanski *et al.*, 1991). Hillenbrand *et al.* (1995) recorded 45 men, 48 women, and 46 children each producing three repetitions of twelve vowels in /hVd/ context. These well-recorded and well-characterized stimuli can be used to test vowel recognition in listeners of standard American English. To complement this set of vowel materials, we have recorded multiple productions of twenty consonants in medial (v/C/v) and initial (C/v) positions in three vowel contexts (/a/, /i/ and /u/) spoken by five women and five men. Measurements were made of fundamental and formant frequencies. After recording, the signals were presented to six listeners to test identification. This paper describes the recording and testing procedures for these consonant tokens.

2. Consonant Recording and Analysis

2.1 Talkers

A total of ten talkers, five men and five women, were selected to record the consonant materials. Talkers were chosen who had no noticeable regional accent (standard American Midwest dialect).

2.2 Recordings

A set of twenty-five consonants /b, d, g, p, t, k, m, n, ŋ, l, r, f, v, θ, ð, s, z, ʃ, tʃ, dʒ, ʒ, j, w, ʌ, h/ was recorded in the medial (v/C/v) and initial (C/v) position in three vowel contexts: /a/, /i/ and /u/. Ten repetitions of each token were recorded at a sampling rate of 44.1 kHz. The recordings were made in a double-walled, sound-treated booth (IAC), using a 1" free field microphone (B&K 4144) attached to a preamplifier (B&K 2619) and amplifier (B&K 2609). Sound pressure level was monitored with a SA0180 meter, which was calibrated with a 94dB SPL tone at 1kHz. The amplifier output was digitized at 44.1 kHz on a Tucker-Davis Technologies (TDT) 24-bit analog to digital converter (DD1) and stored in 16-bit format.

Talkers were instructed to produce ten repetitions of each item at a normal speaking rate without overarticulation. Talkers practiced until their productions were consistent in duration and amplitude across the ten repetitions. Recording gain was adjusted to achieve the maximum digitized levels without clipping. If clipping was detected, the item was repeated.

2.3 Acoustic measurements

Measurements of the fundamental voicing frequency and the first two formant frequencies were made for each talker and each token. The power spectrum of the steady state portion of the initial vowel was used for tokens in the v/C/v format, and the steady state of the final vowel for tokens in the C/v format. Before the power spectrum was calculated, the signals were filtered to include only the band where the formant should be. The power spectrum was not smoothed, and the peak values of the spectrum were chosen as the formant frequencies, being careful not to select harmonics of the fundamental frequency as the first and second formants. First, the signal was filtered with a low-pass filter with cutoff frequency of 270 Hz. The peak in the power spectrum of the resulting signal was selected as the fundamental frequency. Low-pass and high-pass filters were used to filter the signal prior to calculating the first (cutoff frequencies of 480 Hz and 850 Hz respectively) and second (1.2 and 2.5 kHz) formants. The power spectrum of each new filtered signal was calculated, and the frequencies where the peak values of the signal were located were chosen as the formant frequencies.

The average signal-to-noise level of the recordings was measured from 230 samples selected at random from the 1200 tokens. The rms level of the steady state portion of the vowel was computed and compared to the rms level of the silent interval prior to the onset of the token. The average difference in rms level between the background noise and the vowel was 45.6 dB with a standard deviation across tokens of 3.23 dB.

2.4 Results

The values obtained for the fundamental and formant frequencies using the power spectrum of the signal were similar to previous measurements by Hillenbrand *et al.* (1995) and Flege and Munro (1994). The average formant values found by Flege and Munro are lower than the ones found here, and the values found in the Hillenbrand study are somewhat higher. However, the relative difference between the formant frequencies of different vowels is similar to previous measurements. The minimum, average, and maximum formant frequency values for each gender and each vowel are given in Table 1.

Table 1. Formant frequency statistics for the vowels /a/, /i/ and /u/.

Formant	Gender	"a"			"i"			"u"		
		low	mean	high	low	mean	high	low	mean	high
F0	M	86.1	113.3	172.3	86.1	134.0	193.8	86.1	130.4	236.9
	F	172.3	202.8	258.4	172.3	225.9	279.9	172.3	232.2	323.0
F1	M	559.9	719.0	839.8	301.5	410.1	581.4	301.5	394.7	516.8
	F	559.9	745.9	1012.1	366.1	466.0	710.6	387.6	466.2	624.5
F2	M	947.5	1214.8	1485.8	1830.3	2305.5	2820.4	969.0	1298.9	2304.1
	F	990.5	1366.8	1873.4	2110.3	2726.9	3251.4	882.9	1551.8	2454.8

3. Consonant recognition

To ensure that the stimuli chosen were reliable, six normal-hearing listeners were presented with the recorded consonants for speech recognition.

3.1 Stimuli

From the ten samples recorded, a single token was selected that best represented the talker's output considering duration, signal amplitude, and overall clarity. After tokens were selected, they were equated in rms amplitude for the steady-state portion of the vowel segment. Only 20 of the 25 consonants recorded were chosen to be included in the formal listening tests. Five consonants (/ŋ, θ, ʌ, h, ʒ/) were dropped because of pronunciation errors across the ten talkers and perception difficulties in pilot listening experiments. The consonants included in the final database are /b, d, g, p, t, k, m, n, l, r, f, v, s, z, ʃ, tʃ, ð, dʒ, w, j/. The amplitude of the signal was calibrated so that the vowel of each token was presented at 70 dB on an A-weighted scale.

3.2 Procedures

Listeners were tested in a sound-treated booth (IAC) with the consonants presented diotically over headphones (TDH-49). Following each stimulus presentation, subjects indicated which consonant they heard from a matrix of 20 consonants presented on the screen. Each consonant was represented on the screen by an example word. All 20 consonants were presented in both initial and medial position in all three vowel contexts. All stimuli were presented in random order. Male and female talkers were presented in different blocks with three repetitions of each token for a total of 3600 presentations (20 consonants * 2 positions * 3 vowel contexts * 10 talkers * 3 repeats).

3.2 Results

Confusion matrices were compiled for each talker. Consonants were recognized at 97.5% correct from male talkers and 97.0% correct from female talkers (Table 2). The lowest average recognition score for a single talker was for female 2 whose consonant productions were recognized at 92.3% correct. Most of the errors were confusions between /v/ and /ð/.

A confusion matrix was computed for all listeners over all the talkers, consonant positions, and vowel contexts and is shown in table 3. A total of 1080 (3 repetitions * 2 positions * 3 vowels * 10 talkers * 6 listeners) tokens were sent of each consonant.

Table 2. Average recognition performance by six normal-hearing listeners

Talker	Recognition Performance (%)	Talker	Recognition Performance (%)
Male 1	98.9	Female 1	98.9
Male 2	96.8	Female 2	92.3
Male 3	96.1	Female 3	97.0
Male 4	98.4	Female 4	98.9
Male 5	97.3	Female 5	97.8

Table 3. Confusion matrix of errors over all talkers for both initial and medial consonants in all three vowel contexts

Consonant		Heard																			
		b	d	g	p	t	k	m	n	l	r	f	v	s	z	ʃ	tʃ	ð	dʒ	w	j
Sent	b	1073	2		4	1															
	d		1071	6			1										2				
	g	1		1067			12														
	p				1075		3	2													
	t				1	1078															1
	k						1078												2		
	m							1060	20												
	n				1			16	1062						1						
	l					1				1064								14			1
	r				1					1	1077	1									
	f									1		1029	4					46			
	v	4	2							4			986					80	1		3
	s										1		1065	1				13			
	z												17	1041				19	2		1
	ʃ													3		1072	4		1		
	tʃ					15										3	1055		7		
	ð	1	5			1				18	1		89					963	2		
	dʒ		47	37					2						22	8	18		945		1
	w												1			1				1074	4
	j					2				2					1				3		1072

4. Summary

A database of 25 initial and 25 medial consonants were recorded from ten speakers in three vowel contexts: /a/, /i/ and /u/. On a subset of 20 consonants, six normal-hearing subjects obtained an average of 97.3% correct identification of these tokens. Both the edited subset of twenty consonants and all original recordings of the 25 initial and medial consonants recorded are available from the first author at shannon@hei.org. All recorded files are in *.wav file format, which is compatible with most PC-based sound cards. A Windows™-based program for presenting stimuli and scoring results is also available.

References and links

Flege, J.E., and Munro, M.J. (1994). "Auditory and categorical effects on cross-language vowel perception", J. Acoust. Soc. Am. **95** (6), 3623-3641.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels", J. Acoust. Soc. Am. **97** (5), Pt. 1, 3099-3110.

Uchanski, R.M., Millier, K.M., Reed, C.M., and Braida, L.D. (1991). "Effects of token variability on resolution for vowel sounds," in *The Auditory Processing of Speech: From Sounds to Words*, edited by M.E.H. Schouten, (Mouton-Gruyer, New York).

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 1NO1, 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.

Report to the Auditor

Published herewith is a condensed version of our auditors' report for calendar year ended 31 December 1998.

Independent Auditors' Report

To the Executive Council
Acoustical Society of America

We have audited the accompanying statements of financial position of the Acoustical Society of America as of December 31, 1998 and the related statements of activity and cash flows for the year then ended. These financial statements are the responsibility of the Society's management. Our responsibility is to express an opinion on the financial statements based on our audit.

We conducted our audit in accordance with generally accepted auditing standards. Those standards require that we plan and perform the audit to obtain reasonable assurance about whether the financial statements are free of material misstatement. An audit includes examining, on a test basis, evidence supporting the amounts and disclosures in the financial statements. An audit also includes assessing the accounting principles used and significant estimates made by management, as well as evaluating the overall financial statement presentation. We believe that our audit provides a reasonable basis for our opinion.

In our opinion, the financial statements referred to above present fairly, in all material respects, the financial position of the Acoustical Society of America as of December 31, 1998 and the changes in its net assets and its cash flows for the year then ended in conformity with generally accepted accounting principles.

CONROY, SMITH & CO.
18 June 1999
New York, NY

**ACOUSTICAL SOCIETY OF AMERICA
STATEMENTS OF FINANCIAL POSITION
AS OF 31 DECEMBER 1998
(With Comparative Totals For 1997)**

	1998	1997
Assets:		
Cash and cash equivalents	\$2,641,647	\$ 291,383
Accounts receivable.....	378,367	227,177
Marketable securities.....	3,914,155	5,134,080
Furniture, fixtures and equipment—net	102,218	101,165
Other assets.....	246,546	270,367
Total assets	<u>\$7,282,933</u>	<u>\$6,024,172</u>
Liabilities:		
Accounts payable and accrued expenses	\$ 205,838	\$ 234,484
Deferred revenue	1,146,724	836,905
Deferred rent liability	46,394	46,228
Total liabilities	<u>\$1,398,956</u>	<u>\$1,117,617</u>
Net assets:		
Unrestricted.....	\$4,758,932	\$3,882,375
Temporarily restricted	528,825	472,038
Permanently restricted	596,220	552,142
Total net assets	<u>\$5,883,977</u>	<u>\$4,906,555</u>
Total liabilities and net assets	<u>\$7,282,933</u>	<u>\$6,024,172</u>

ACOUSTICAL SOCIETY OF AMERICA
 STATEMENTS OF ACTIVITY
 FOR THE YEAR ENDED 31 DECEMBER 1998
 (With Comparative Totals For 1997)

	1998	1997
Unrestricted net assets:		
Revenue		
Dues	\$ 693,123	\$ 522,273
Publishing	2,158,365	2,204,176
Standards.....	352,200	270,506
Meetings.....	786,311	295,101
Interest and dividends.....	206,691	168,799
Unrealized gain (loss).....	98,442	161,509
Other	71,545	90,786
Realized gain (loss).....	<u>222,185</u>	<u>340,162</u>
Total unrestricted revenue	<u>\$4,588,862</u>	<u>\$4,053,312</u>
Expenses:		
Publishing	\$1,682,237	\$1,820,558
Standards.....	448,058	428,281
Administrative and general	493,500	485,402
Meetings.....	690,157	452,926
Other expenses.....	<u>406,989</u>	<u>263,711</u>
Total expenses	<u>\$3,720,941</u>	<u>\$3,450,878</u>
Increase in net assets.....	\$ 867,921	\$ 602,434
Net assets released from restrictions:		
Satisfaction of program restrictions.....	<u>8,636</u>	<u>7,764</u>
Increase in unrestricted net assets and reclassifications.....	<u>\$ 876,557</u>	<u>\$ 610,198</u>
Temporarily restricted net assets:		
Contributions	\$ 1,309	\$ 2,861
Investment income.....	52,145	65,004
Unrealized gain (loss).....	11,969	20,461
Release of restrictions	<u>(8,636)</u>	<u>(7,764)</u>
Increase (decrease) in temporarily restricted net assets	<u>\$ 56,787</u>	<u>\$ 80,562</u>
Permanently restricted net assets:		
Investment income.....	\$ 60,994	\$ 80,680
Unrealized gain (loss).....	14,000	25,377
Expenses	<u>(30,916)</u>	<u>(39,816)</u>
Increase (decrease) in permanently restricted net assets	<u>\$ 44,078</u>	<u>\$ 66,241</u>
Increase in net assets.....	\$ 977,422	\$ 757,001
Net assets, beginning of year.....	<u>4,906,555</u>	<u>4,149,554</u>
Net assets, end of year	<u><u>\$5,883,977</u></u>	<u><u>\$4,906,555</u></u>

ACOUSTICAL SOCIETY OF AMERICA
 STATEMENTS OF CASH FLOWS
 FOR THE YEAR ENDED 31 DECEMBER 1998
 (With Comparative Totals For 1997)

	Total All Funds	
	1998	1997
Operating Activities		
Increase in net assets	\$ 977,422	\$ 757,001
Adjustments to reconcile net income to net cash provided by operating activities:		
Depreciation and amortization	28,964	26,919
Unrealized (gain) loss on marketable securities.....	(124,411)	(207,347)
Changes in operating assets and liabilities:		
(Increase) decrease in accounts receivable.....	(151,190)	58,308
Decrease (increase) in other assets	23,821	39,389
Increase (decrease) in accounts payable and accrued expenses	(28,646)	105,455
Increase in deferred rent liability.....	166	4,902
Increase (decrease) in deferred revenue	309,819	(211,133)
Net cash flows provided by operating activities	<u>\$1,035,945</u>	<u>\$ 573,494</u>
Investing Activities		
Purchase of furniture, fixtures, equipment and leasehold improvements	(\$ 30,017)	(\$ 15,557)
Proceeds from sale of securities	4,623,945	4,564,188
Purchase of securities	(3,279,609)	(5,511,967)
Net cash (used in) provided by investing activities.....	<u>\$1,314,319</u>	<u>(\$ 963,336)</u>
Increase (decrease) in cash and cash equivalents	\$2,350,264	(\$ 389,842)
Cash and cash equivalents, beginning of year	<u>291,383</u>	<u>681,225</u>
Cash and cash equivalents, end of year	<u><u>\$2,641,647</u></u>	<u><u>\$ 291,383</u></u>

ACOUSTICAL NEWS—INTERNATIONAL

Walter G. Mayer

Physics Department, Georgetown University, Washington, DC 20057

EAA Secretariat moves to the UK

The work of the Institute of Acoustics (UK) continues; it was announced that the Secretariat of the European Acoustics Association is now located in the Institute's offices in St Albans. It is expected that the IOA will play a major role in the development of the EAA, moreover, it is hoped that this arrangement will insure that the members of the IOA can effectively compete for work throughout the European Union. This would mean common standards and cross recognition of professional qualifications. Although these matters advance slowly due to lengthy consultations and the different systems used throughout Europe, nonetheless the Institute has made a start. (Based on a President's Letter published in the "Acoustics Bulletin" of the IOA.)

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*.

December 1999

2–4 **International Conference and Exhibition on Ultrasonics**, New Delhi. (Fax: +91 11 575 2678; e-mail: iceau99@csnpl.ren.nic.in) 10/99

15–17 **International Conference on Stochastic Volume and Surface Scattering**, Cambridge. (Fax: +44 1727 850 553; e-mail: ioa@ioa.org.uk) 10/99

January 2000

20–21 ***European Symposium on Ultrasound Contrast Imaging**, Rotterdam, The Netherlands. (Mrs. C. Eefting, Erasmus University Thoraxcentre, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands; Fax: +31 10 408 9445; Web: www.eur.nl/fgg/thorax/contrast)

February 2000

17–18 **Measuring Noise Outdoors**, Home Counties Venue. (Fax: +44 1727 850 553; e-mail: ioa@ioa.org.uk)

March 2000

15–17 **Acoustical Society of Japan Spring Meeting**, Tokyo. (Fax: +81 3 3379 1456; e-mail: kym05145@nifty.ne.jp) 8/99

19–22 **25th International Acoustical Imaging Symposium**, Bristol. (Web: www.bris.ac.uk/depts/medphys) 10/99

20–24 **Meeting of the German Acoustical Society (DAGA)**, Oldenburg. (Fax: +49 441 798 3698; e-mail: dega@aku.physik.uni-oldenburg.de) 10/98

April 2000

3–4 **Structural Acoustics '2000**, Zakopane, Poland. (Fax: +48 12 423 3163; Web: www.cyf-kr.edu.pl/ghpanusz) 8/99

May 2000

17–19 **9th International Meeting on Low Frequency Noise and Vibration**, Aalborg. (Fax: +44 1277 223 453) 6/99

23–26 **Russian Acoustical Society Meeting**, Moscow. (Fax: +7 095 126 8411; e-mail: ras@akin.ru) 10/99

24–26

Joint International Symposium on Noise Control & Acoustics for Educational Buildings (24–25 May) and **5th Turkish National Congress on Acoustics** (25–26 May), Istanbul. (Fax: +90 212 261 0549; Web: www.takder.org) 10/99

June 2000

5–9

International Conference on Acoustics, Speech and Signal Processing (ICASSP-2000), Istanbul. (Fax: +1 410 455 3969; Web: 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) 6/99

14–17

IUTAM Symposium on Mechanical Waves for Composite Structures Characterization, Chania. (Fax: +30 821 37438; Web: 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) 12/98

10–13

***5th European Conference on Underwater Acoustics**, Lyon, France. (LASSO, 43 Bd. du 11 novembre 1918; Bat. 308; BP 2077, 69616 Villeurbanne cedex, France; Fax: +33 4 72 44 80 74; Web: www.ecua2000.cpe.fr)

August 2000

28–30

INTER-NOISE 2000, Nice. (Fax: +33 1 47 88 90 60; Web: www.inrets.fr/services/manif) 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) 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, Belgium. (Mrs. L. Notré, K. U. Leuven, PMA Division, Celestijnenlaan 300B, 3001 Leuven, Belgium; Fax: +32 16 32 24 82; e-mail: lieve.notre@mech.kuleuven.ac.be)

17–21

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

October 2000

3–5

WESTPRAC VII, Kumamoto. (Fax: +81 96 342 3630; Web: cogni.eecs.kumamoto-u.ac.jp/others/westprac7) 6/98

3–6

EUROMECH Colloquium on Elastic Waves in Non-destructive Testing, Prague. (Fax: +420 2 858 4695; e-mail: ok@bivoj.it.cas.cz) 10/99

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) 12/98

16–20

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

August 2001

28–30

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

10–13

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

September 2001

2–7

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

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) *10/99*

FORUM

Forum is intended for communications that raise acoustical concerns, express acoustical viewpoints, or stimulate acoustical research and applications without necessarily including new findings. Publication will occur on a selective basis when such communications have particular relevance, importance, or interest to the acoustical community or the Society. Submit such items to an appropriate associate editor or to the Editor-in-Chief, labeled FORUM. Condensation or other editorial changes may be requested of the author.

Opinions expressed are those of the individual authors and are not necessarily endorsed by the Acoustical Society of America.

Underwater sound transmission and SI units [43.05.Ky]

Clarence S. Clay

Department of Geology and Geophysics, 1215 West Dayton Street, University of Wisconsin, Madison, Wisconsin 53706-1692 and 5033 Saint Cyr Road, Middleton, Wisconsin 53562-2424

(Received 27 July 1999; accepted for publication 19 August 1999)

[S0001-4966(99)00212-X]

Very long-range acoustic transmissions are a way of measuring the temperature of water masses in the ocean, for global warming research.¹ In this field, misunderstanding between professionals in audiology–biology and underwater sound has had very serious consequences. Work has been stopped while marine biologists study the effect of the sound transmissions on whale behavior. This is a difficult and expensive task. Whale and elephant seal behavior was observed during “on” and “off” periods of transmission from a “195-dB” source 100 miles southwest of San Francisco. Converted to SI units, the acoustic power of the source is less than 260 watts. No marked changes in swimming behavior were observed by the biologists. The underwater acoustics research budget provided \$2.9 million for these animal studies and associated legal fees.²

It is the use of decibels, combined with the perspectives of people in different disciplines, which causes the misunderstanding. Acousticians who work in air use the decibel (the logarithm of the ratio of two sound pressures

squared, with the units “dB”) to report sound-pressure levels in air. In air, the thresholds of hearing and pain are approximately 0 and 130 decibels, respectively. Sound-pressure levels near the speakers of a rock band can be painful. But, air acousticians and the underwater sound community operate in different media and use grossly different reference values for their calculations of decibels. Air acousticians use a reference sound pressure of 20 micropascal. In underwater sound, reference sound pressures have been variously 20 micropascal (in the 1940’s), 0.1 pascal (1950–1960’s), and 1 micropascal (the current choice). The dB is not a linear unit like a meter, a pascal, or a watt. It is the logarithm of the ratio of a quantity to a like quantity.

The whole misunderstanding would have been avoided if the underwater sound people had used SI units (international system of metric units) of pressure and of acoustically radiated power. The acoustic power output of most underwater sound sources ranges from a few watts to kilowatts. During a meeting of ocean acousticians at the Scripps Institution of Oceanography,³ Professor Walter Munk and I proposed the following:

- (1) Sound pressures are to be reported in pascals, and
- (2) Radiated source powers are to be reported in watts.

¹Robert Iron, Meeting briefs, “Sounding Out Pacific Warming,” *Science* **279**, 1302–1303 (27 Feb. 1988).

²Echoes, Vol. 8, No. 1 (Winter 1999), p. 8.

³“Acoustic Techniques for Measuring Ocean Variables,” edited by Capt. Bob Smart and Alexander Vornovich (1995). NOAA Envr. Tech. Lab., 325 Boulder, CO, agv@etl.noaa.gov

Noise Control and SI Units [43.05.Ky]

Robert Hickling

Sonometrics Inc., 8306 Huntington Road, Huntington Woods, Michigan 48070

(Received 27 July 1999; accepted for publication 19 August 1999)

© 1999 Acoustical Society of America. [S0001-4966(99)00112-5]

The use of decibels is more deeply entrenched in noise control than in other branches of acoustics because, over the years, agencies, from international organizations down to local governments, have generated a large body of noise standards and test procedures based on decibels. Also the public appears to think the decibel is an inherent part of noise measurement. Sportscasters frequently use the word to describe the noise of a crowd. Using a complicated unit that the public does not understand may also provide noise control with a kind of aura or professional mystique.

Decibels are different from standard logarithmic representations in powers of 10. What decibel units are and how they came to be used has been well-documented.¹ However they appear to have little scientific justification.² The idea that they are needed to match the response of the human ear is a rationalization that was introduced some time after decibels were first used.² In particular the argument that decibels are needed to match the dynamic range of human hearing is untrue. Decibels have the following problems:

- (1) They create confusion, because the same word, decibel, is used for different acoustical quantities, such as sound pressure, intensity, and power. Confusion also occurs with decibels in different media, particularly when comparing the effect of sound on sea animals and man.³
- (2) The logarithmic form of decibels obscures reasoning and is a cover for inaccuracy.
- (3) Decibels are more related to the era of slide rules and log tables than to modern digital processing.
- (4) Mathematical expressions in decibel form are cumbersome and difficult to interpret.

A major problem with decibels is that they contain the implicit assumption that sound intensity is represented by the farfield approximation. Sound intensity (sound power flow per unit area) is a vector quantity.^{1,2} In the farfield approximation, sound intensity is represented by an inexact scalar quantity based on sound pressure squared. No allowance is made for the fact that practical and accurate methods of measuring sound intensity in its vector form were developed 20 years ago, making it possible, for the first time, to determine sound power accurately in a normal environment. To be able to determine sound power in a normal environment is a major engineering advance, because sound power is the best single, overall measure of a noise source. However, sound power based on sound-intensity measurement is still not used in noise standards and test procedures in industry, mainly because of the continued use of decibels and the associated assumption of the farfield approximation.

It has been suggested that confusion with decibels can be avoided by always stating the appropriate reference value.^{4,5} However this may be too onerous for most people. Similar efforts in the past to improve the use of the decibel system have not met with much success.⁶ Shoring up the decibel system does not address the bulk of the problems associated with decibels.

The obvious solution is simply to abandon decibels entirely. Acoustical measurements are first made in linear SI units and then converted to decibels. It would make sense simply to forego the final conversion and leave the measurement in its original linear SI form. All of the problems just described would then disappear and noise control could become an engineering discipline like any other. In animal acoustics and underwater sound, linear units are already being used.^{7,8} In recent noise-control papers by the writer, sound power is represented in both linear and logarithmic form.^{9,10} In architectural acoustics, it is recognized that the quality of auditoria and concert halls is determined by how well an audience can distinguish the direction of a sound source.¹¹ The best way to measure this is to use vector sound intensity, without decibels.²

However, the use of decibels cannot be abandoned easily in noise control because of the existing noise standards and test procedures based on decibels. These standards and procedures should therefore be modified so that they are expressed both in decibel and linear SI units. In this way there would not be an abrupt change with the past. Similarly comparisons on a decibel scale should be expressed as a ratio or percentage. Plots covering a wide dynamic range should be in standard logarithmic form.

Several centuries ago it took European bankers and merchants more than a hundred years to replace Roman numerals with the present system of numbers. Hopefully it will not take acousticians that long to abandon decibels. Decibels are a useless affectation, which is impeding the development of noise control as an engineering discipline.

¹A. D. Pierce, "Acoustics, An Introduction to its Physical Principles and Applications" (Acoustical Society of America, Woodbury, New York, 1989), pp. 60–65.

²R. Hickling, "Octaves, Decibels, the Farfield Approximation and Sound Power," Report No. 98-RH001, Sonometrics Inc., Huntington Woods, MI 48070 (1998).

³Anon., "Low-Frequency Sound and Marine Mammals," prepared by The Committee on Low-Frequency Sound and Marine Mammals, The Ocean Studies Board and The Committee on Geosciences, Environment and Resources, under the auspices of the National Research Council (National Academy Press, Washington, D. C., 1994).

⁴W. M. Carey, "Standard Definitions for Sound Levels in the Ocean," Editorial Article in *IEEE J. Ocean Eng.* **20**, 109–113 (1995).

⁵D. M. F. Chapman and D. D. Ellis, "The Elusive Decibel: Thoughts on Sonars and Marine Mammals," *Can. Acoust.* **26**(2), 29–31 (1998).

⁶C. W. Horton, "The Bewildering Decibel" *Electron. Eng. (U.K.)* **73**, 550–555 (1954).

⁷B. Holdobler, U. Braun, W. Gronenberg, W. R. Kirshner, and C. Peeters, "Trail Communication in the Ant *Magnaponera foetens* (Fabr.) (Formicidae, Ponerinae)," *J. Insect Physiol.* **40**, 585–593 (1994).

⁸C. S. Clay, "Underwater Sound Transmission and SI Units," *Echoes* **8**, 3 (1998).

⁹R. Hickling, L. N. Bolen, and R. F. Schumacher, "Indoor System for Efficient Measurement of the Sound Power of Light Vehicles and for Noise-Control Diagnostics," Transactions, Society of Automotive Engineers, Paper No. 891145 (1989).

¹⁰R. Hickling and P. Lee, "Determining Sound Power in Reverberation Rooms and Other Indoor Work Spaces using Vector Sound Intensity Measurement" Proceedings NOISE-CON 97, June 1997, pp. 483–488.

¹¹M. R. Schroeder, "Fractals, Chaos and Power Laws" (W. H. Freeman and Co., New York, 1991), pp. 72–79.

Improvement of sound barriers using headpieces with finite acoustic impedance

Michael Möser and Rudi Volz

Institute of Technical Acoustics, Technical University Berlin, Einsteinufer 25, D-10587 Berlin, Germany

(Received 10 June 1998; revised 16 July 1999; accepted 4 August 1999)

The paper deals with the reduction of the sound energy in the shadow region behind barriers by means of an attached body at the edge of the screen. The reflecting attributes of the barrier's headpiece are described by a locally reacting impedance. Diffraction at ideal soft and hard bodies demonstrates the basic principle of tangential power transport parallel to their surface: the impedance must be chosen so that the tangential intensity near the edge is lowered, turning the incoming power in harmless directions. The differences due to finite impedances are then discussed. The physical principles are demonstrated in frozen pictures of the sound field for the different cases. Theoretical computations show considerably improved levels in the shadow zone for larger angles of diffraction. These are compared with empirical results, and practical applications are discussed. © 1999 Acoustical Society of America. [S0001-4966(99)00712-2]

PACS numbers: 43.20.Bi, 43.20.Fn, 43.20.Hq, 43.50.Gf [DEC]

INTRODUCTION

Noise barriers are now commonly used to protect against traffic noise. The approximate equation of calculation of the shielding effect, based on the Sommerfeld considerations,¹⁻³ is well established in acoustics.

An old dream has been to achieve an improvement of insertion loss of noise shields with the minimum expenditure. Many attempts have been made to manipulate the edge of the screen to prevent disturbing noise from infiltrating towards the shadow region. Changes were made to the shape (e.g., T-shaped and arrow-shaped constructions) in combination with absorbent covering of the upper region of the screen and the use of constructions with active noise control, with differing degrees of success.⁴⁻²⁷ Recently, practical experiments *in situ* with an absorbing cylindrical headpiece on a noise screen have only resulted in small effects.^{12,16}

Instead of considering the influence of the geometry of the screen and the absorption of parts of the screen, we looked at the influence resulting from the acoustic impedance of headpieces (or parts of the screen near the edge).

In the following theoretical consideration the headpieces will be modelled as cylindrical bodies with given surface impedance, allowing a simple theoretical treatment. The subsequent discussion of the influence of the impedance uses a comparison with the (more "mathematical") limit cases of ideal soft surfaced ($Z=0$) and ideal hard surfaced ($Z\rightarrow\infty$) headpieces for noise barriers.

The next section reviews the scattering and diffraction on such idealized bodies (here cylinders).

I. THEORY

A. The sound field behind cylinders with hard and soft surface impedance

Bodies with a pressure reflection factor $R=1$ or $R=-1$ take no energy from the surrounding sound field. However, there is a fundamental difference between bodies

with a hard or soft surface impedance regarding power transportation near the surface and therefore also the formation of the shadow region.

The surface of an ideal yielding area of space $Z=0$ is free of sound pressure. For a cylinder with soft surface and radius b , the pressure p and the tangential velocity v_φ are zero:

$$Z=0: p(b)=0, v_\varphi(b)=0. \quad (1)$$

The sound energy cannot pass along a surface $Z=0$. The energy flow density which encloses the soft-surfaced body increases only gradually with the distance from the surface. Figure 1(a) shows the acoustic field in decibels of the active intensity for a cylinder with $Z=0$ (for details see Ref. 28).

Since energy transport is not possible along the "surface" $r=b$ there is also no energy supply to the region behind the cylinder. This creates an extended acoustic shadow.

The behavior of the sound field at hard-surfaced bodies is exactly opposite. The sound pressure shows maxima at the "surface" $r=b$. The progress of the sound field tangential to the surface will be impressed through the impacting sound so the field joins to the hard-surfaced body—even with an increase in sound pressure on the exposed parts of the body. Figure 1(b) shows how the energy flow density here encloses the body. In contrast to soft-surfaced bodies, for $Z\rightarrow\infty$ grazing energy transport takes place close to the surface. So with smaller reflector dimensions there is nearly no way of creating a shadow region.

B. The effect of cylindrical headpieces with soft or hard surfaces

An obvious approach is to create shadows with soft surfaces to improve the insertion loss of extended sound noise barriers. If the infiltration of sound power behind the soft body is only possible on one side after attaching an additional large screen, then the energy deflecting effect of the soft body in the shadow region will be appreciable: with $Z=0$ it has to result in darker shadow region compared to the

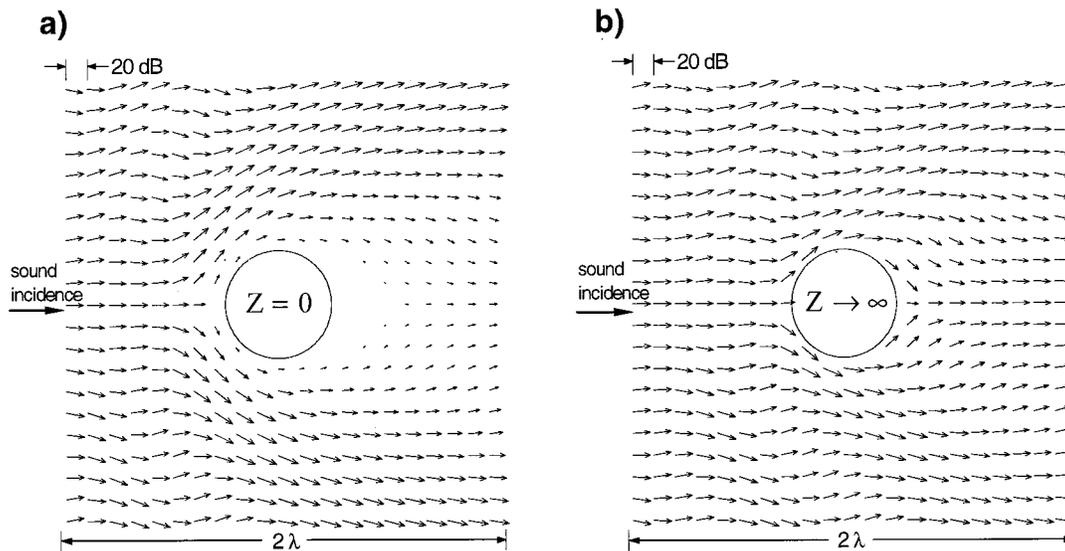


FIG. 1. Distribution of the intensity; $b/\lambda = 0.25$: (a) soft surfaced cylinder and (b) hard surfaced cylinder.

case without the headpiece. The angle of incidence of the sound wave is not crucial. The improvements will depend mainly on the diffraction angle β measured relative to the direction of incidence. The position of a chosen field point relative to the screen is of minor significance for the reduction of sound pressure level caused by the soft body.

There is reason enough to take a more detailed look at the shielding effect of a (hard surfaced) noise screen with a ‘‘sound deflecting’’ headpiece. Of course surfaces with other acoustic properties should also be included, since in air the condition $Z=0$ can only be realized over narrow bands at least passively. The problem of diffraction will be considered for a semi-infinite hard-surfaced wall with a cylinder attached at its top with known surface properties (Fig. 2). The sound field is excited by a nondirected line source a distance x_s from the center of the cylinder and enclosing an angle φ_0 with the screen. It will be assumed in all cases that the source

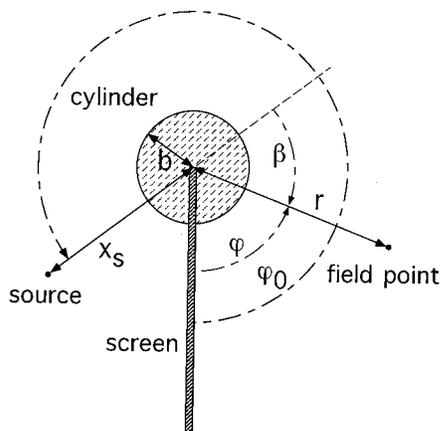


FIG. 2. Scheme of the screen with an attached cylinder where x_s is the distance between source and edge, b is the radius of the cylindrical headpiece, r is the distance between source and field point, β is the diffraction angle, φ_0 is the angle between source and screen, and φ is the angle between the field point and the screen.

creates an incident plane wave due to its large distance from the wall.

The simplest way to discuss the influence of the acoustic properties of the cylinder is to assume it is a given locally reacting impedance Z . The effect of the cylinder with the radius b on the surrounding sound field is then expressed through its impedance

$$Z = -\frac{p(b)}{v_r(b)}, \quad (2)$$

where $p(b)$ is the sound pressure on the surface of the cylinder $r=b$ and $v_r(b)$ is the outward radial component of the velocity. For a first approximation it will be sufficient to examine only locally constant impedances $\partial Z/\partial \varphi = 0$.

The principle of such a reduced model is described in text books on theoretical acoustics. The treatment of the problem was shown in Ref. 28 including some critical cases (e.g., a noise barrier without headpiece or considerations of the far field). For a sound wave with an angle of incidence φ_0 the sound pressure is composed of two parts:

$$p(r, \varphi) = p_{\text{screen}}(r, \varphi) + p_{\text{cyl}}(r, \varphi), \quad (3)$$

where p_{screen} is the pressure observed with the barrier alone and p_{cyl} is the pressure with the cylinder (without barrier).

Of these, the sound field at the screen without cylinder is

$$p_{\text{screen}}(r, \varphi) = p_Q(0) \sum_{n=0}^{\infty} \frac{2e^{j(n\pi/4)}}{\epsilon_n} J_{n/2}(kr) \cos\left(\varphi \frac{n}{2}\right) \cos\left(\varphi_0 \frac{n}{2}\right), \quad (4)$$

where k is the wave number ($k = \omega/c = 2\pi/\lambda$). Equation (4) is also represented with the Fresnel integrals (see Ref. 28). Also,

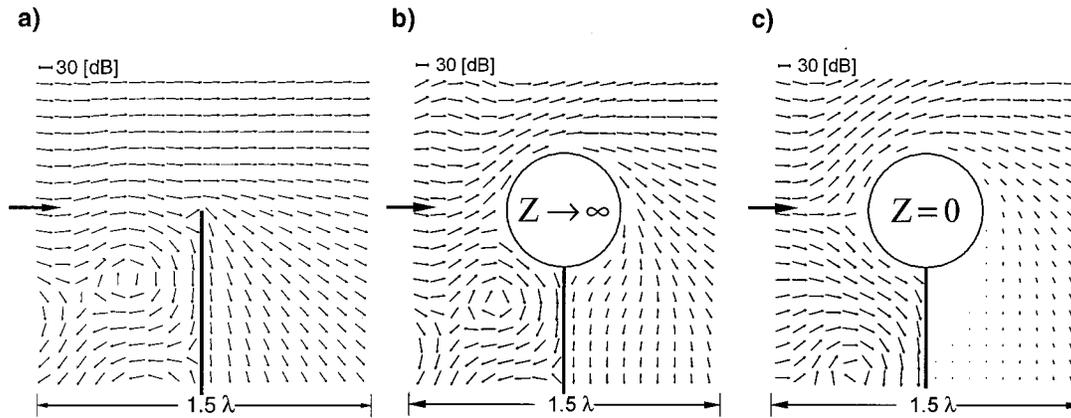


FIG. 3. Distribution of the intensity; $b/\lambda = 0.25$, $\varphi_0 = 90$ degrees: (a) screen without cylinder, (b) screen with attached hard surfaced cylinder, and (c) screen with attached soft surfaced cylinder.

$$p_{cyl}(r, \varphi) = -p_Q(0) \sum_{n=0}^{\infty} \frac{2e^{j(n\pi/4)}}{\varepsilon_n} Q_{n/2}(kb) \times H_{n/2}^{(2)}(kr) \cos\left(\frac{n}{2}\varphi\right) \cos\left(\frac{n}{2}\varphi_0\right), \quad (5)$$

with

$$Q_{n/2}(kb) = \frac{J_{n/2}(kb) + j(Z/\rho c)J'_{n/2}(kb)}{H_{n/2}^{(2)}(kb) + j(Z/\rho c)H_{n/2}^{(2)'}(kb)} \quad (6)$$

($'$ = derivative of the argument) including the variation of the sound field through the attached cylinder resulting from Eq. (2).

In the above equations, $J_{n/2}$ = spherical Bessel function, order $n/2$; $H_{n/2}^{(2)}$ = Hankel function (spherical Bessel function of the third kind), order $n/2$; and $p_Q(0)$ = sound pressure of the source in free field at $r=0$.

The expression of the field using terms with the form $\cos[\varphi(n/2)]$ makes sense: in the sense of a Fourier sum these are all φ -dependent functions whose derivatives disappear at the limits $\varphi=0$ and $\varphi=2\pi$.

The results of the numerical analysis from Eq. (3) initially validate the earlier proposals. Figure 3 shows the calculated local intensity for three different cases: at a screen and at a screen with an attached small cylinder $b/\lambda = 0.25$ for the cases $Z = \infty$ and $Z = 0$. For the soft-surfaced case the infiltration of the field in the shadow region is much less than for the hard-surfaced case. There is only a small difference between the screen alone and the screen with the hard-surfaced cylinder.

Obviously the diffraction field in the range of the geometrical shadow is also influenced by energy transport near the surface of the cylinder over larger distances as discussed in the previous paragraph: the “energy supply” of the shadow region comes mainly from the contact area near the surface $r=b$. The further away the examined field point is from the geometrical border of the shadow $\beta=0$, the more important this supply route is for the sound field. Figure 3 shows the substantial differences between the cases $Z=0$ and $Z=\infty$: even with a small headpiece with $b/\lambda = 0.25$ the differences in the intensity are more than 10 dB.

That such large differences in the insertion loss,

$$IL = 10 \times \lg \left(\frac{P_{\text{without headpiece}}^2}{P_{\text{with headpiece}}^2} \right), \quad (7)$$

are observed at large distances is already indicated by the local energy flow densities. Calculations²⁸ show that the additional decrease in sound pressure level obtained by attaching the cylindrical headpiece to the screen compared with the semi-infinite screen without headpiece in the far field $r \gg b$ and $r \gg \lambda$ will be independent of the distance r from the field point to the center of the cylinder. So the far field insertion loss $IL_{\text{far field}}$, which is interesting especially under practical conditions, is therefore particularly suitable for describing the global effect. Figure 4 shows the calculated values for variation of the radius of the cylinder in steps of $\frac{1}{3}$ octave (over four octaves) for one direction of incidence. As expected, the diffraction angle β is primarily responsible for the quality of decrease. The angle of incidence φ_0 is only of secondary importance. While hard-surfaced headpieces only have small effects, headpieces with $Z=0$ lead to considerable decreases in sound pressure level compared with the screen alone. For diffraction angles of about 22.5 degrees they account for 5–10 dB, and for 45 degrees, 8–18 dB, depending on the diameter of the headpiece. This is obvi-

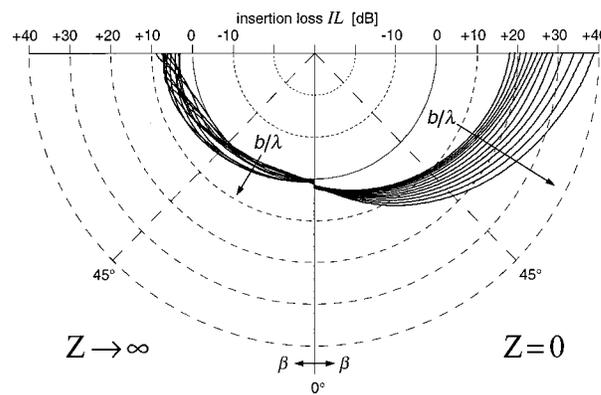


FIG. 4. Far-field insertion loss IL . Left side: screen with attached hard surfaced cylinder. Right side: screen with attached soft surfaced cylinder. $\varphi_0 = 75$ degrees; from inside to outside: $b/\lambda = 0.1; 0.125; 0.16; 0.2; \dots; 1.6$.

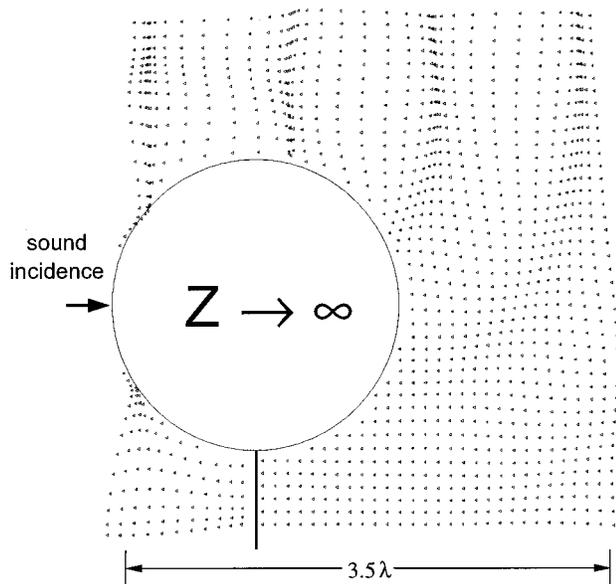


FIG. 5. Movement of the air particles; $b/\lambda=1$, $\varphi_0=90$ degrees; hard-surfaced cylinder.

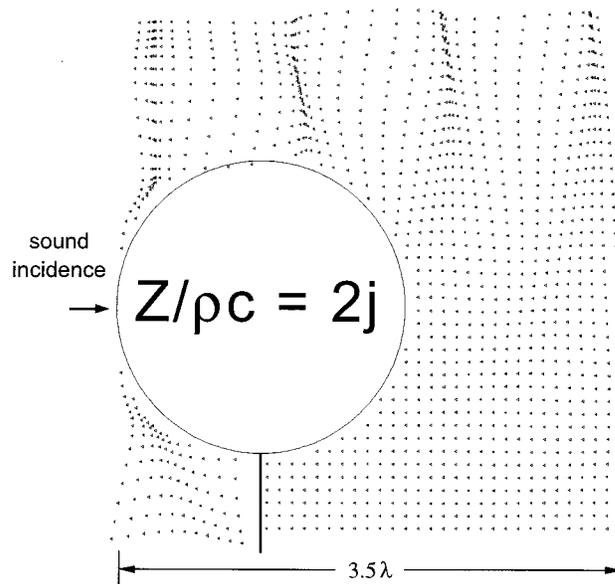


FIG. 7. Movement of the air particles; $b/\lambda=1$, $\varphi_0=90$ degrees; impedance with mass character.

ously due to the prevention of the energy flow around the edge of the screen with the attached cylinder.

Estimating the effect of impedances with finite values will require additional physical considerations which take into account the transient phenomena near the cylindrical headpiece where air particles interact with the surface of the headpiece. To make it easier to understand, the local displacements of the field points in the compressible air continuum were calculated. Figures 5–8 show equidistant point rasters (the points may be interpreted as “air particles”). This description of the sound field allows graphical reproduction of the local and global movements and compressions.

The comparison of the reproduced movements of the

“air particles” for the hard- and soft-surfaced case in Figs. 5 and 6 form the key to understanding the laws in principle, including infinite impedance of the headpiece. In the soft-surfaced case the surrounding air near the headpiece will be moved radially by the surface of the cylindrical headpiece. As always with thin gas layers without borders, the boundary layer which encloses the surface has to be interpreted as a mass (only contained gas acts as a spring). Because at $Z=0$ there is no hindrance it will be shifted to the outside if there is an existing current over-pressure at its outer side as shown in Fig. 6. The displacement of a mass is out of phase with the exciting force. In the case of the cylindrical headpiece with $Z \rightarrow \infty$ (Fig. 5), the surrounding air-mass meets with the (infinitely large) resistance on its inner side. It there-

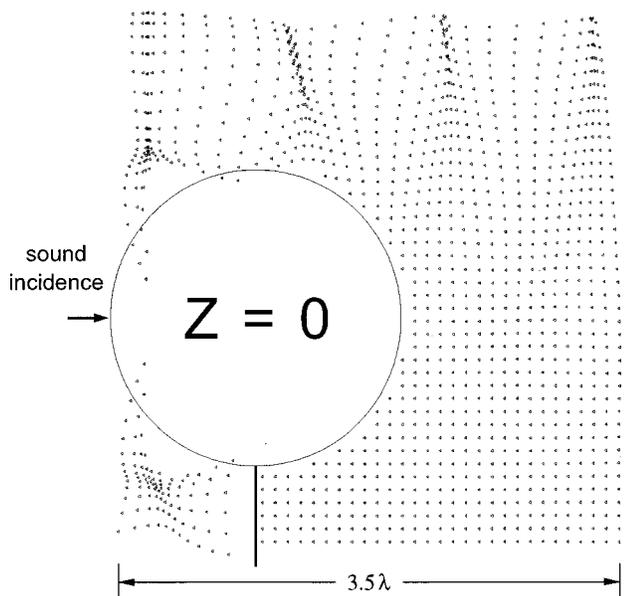


FIG. 6. Movement of the air particles; $b/\lambda=1$, $\varphi_0=90$ degrees; soft-surfaced cylinder.

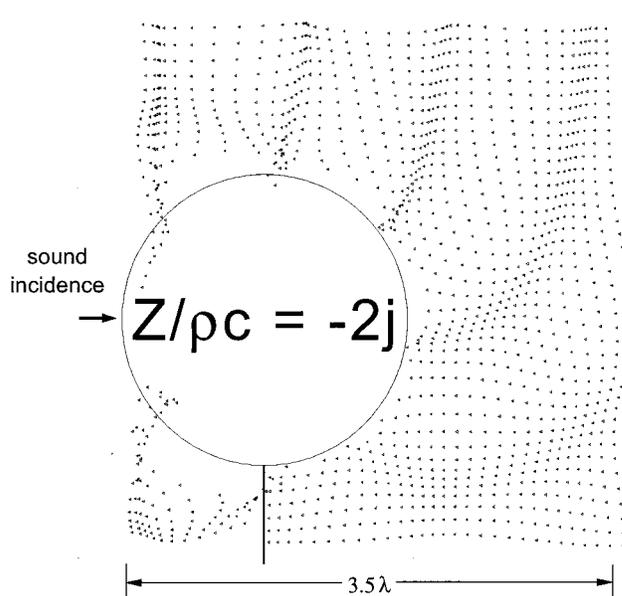


FIG. 8. Movement of the air particles; $b/\lambda=1$, $\varphi_0=90$ degrees; impedance with stiffness character.

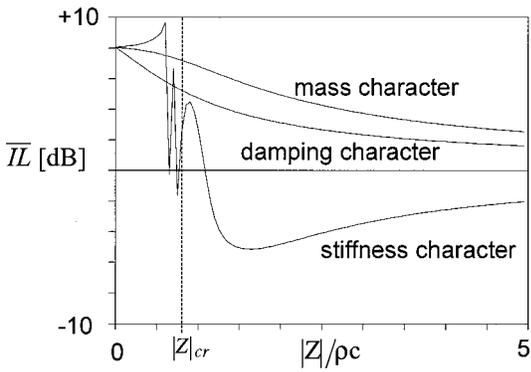


FIG. 9. Average insertion loss \overline{IL} ; $b/\lambda=0.5$; $\varphi_0=90$ degrees.

fore cannot react with the acting force from the outside with radial acceleration, but can only transmit the excited pressure distribution to the surface. This simultaneously creates the undesirable energy transport at the surface of the cylindrical headpiece, in contrast to the soft surface $Z=0$.

The interpretation of the thin “air layer” near the cylinder as a “pure mass concerning the displacement in radial direction” allows the estimation of finite surface impedances.

C. Influence of finite impedances

1. Impedance with mass character

For impedances $Z=j|Z|$ with mass behavior, the movement of the air mass will be increasingly restricted with increase of $|Z|$, so that the pressure at the surface increases with $|Z|$. Figures 6 and 7 also show that the sound field joins more closely to the cylinder than with the case $Z=0$. This reduces the improvement obtained by the headpiece, so it results in a gradual transition between the extremes $Z=0$ and $Z\rightarrow\infty$ (see Fig. 9).

2. Damping impedance

Damping impedances $Z=|Z|$ behave much like mass character impedances. The decreasing mobility of the added moving air mass with an increase of $|Z|$ results in increased sound pressure $p(b)$ near the surface, and again in gradual transition between $Z=0$ and $Z\rightarrow\infty$ with an increase in $|Z|$ (see Fig. 9).

3. Impedance with stiffness character

Impedances with spring character $Z=-j|Z|$ with the added moving air mass result in a simple resonator system. The ratio of the current frequency ω and resonance frequency $\omega_{res}=\sqrt{s''/m''}$ (s'' =stiffness, m'' =added moving mass per unit area) is critical. Below the resonance frequency the force acting along the spring is larger than the exciting force. That means that the sound field at the surface of the cylinder even is larger than for the hard surface $Z\rightarrow\infty$, resulting in a negative insertion loss. In the frequency range above the resonance frequency the force of the spring is smaller (at higher frequency much smaller) than the exciting force. Again positive insertion losses are to be expected.

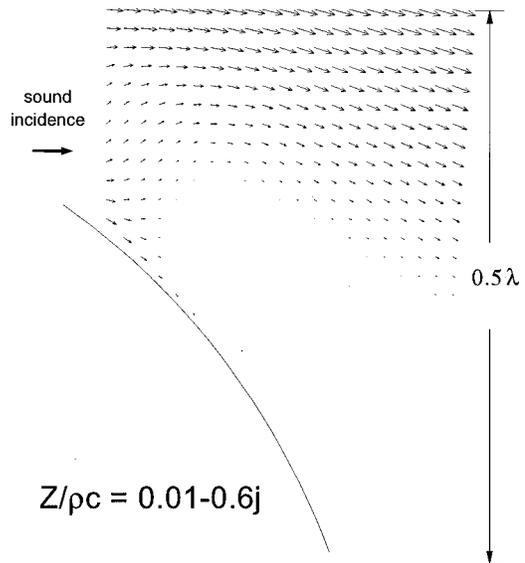


FIG. 10. Distribution of the intensity; $b/\lambda=1$, $\varphi_0=90$ degrees; impedance with stiffness character and small damping.

From these considerations it is also possible to derive the dependence of the absolute value of the impedance at constant frequencies (see Fig. 9). Adequately small $|Z|$ correspond to a low-tuned resonator and so, for this case, there will be real improvements. With rising $|Z|$ the resonance will be tuned to higher frequencies. Above a certain critical impedance the resonance frequency is higher than the observed frequency. In the range of higher absolute values of the impedance the results are worse than for the (nearly ineffective) hard-surfaced case. Figure 8 shows the “overlapping” of the sound field. Obviously the pressure at the surface is higher than for the case $Z\rightarrow\infty$ in Fig. 5.

This is also clearly demonstrated in Fig. 9. In this case the average insertion loss is used for reasons of simplicity:

$$\overline{IL} = \frac{\int_{15^\circ}^{90^\circ} p^2(\text{without headpiece}) d\beta}{\int_{15^\circ}^{90^\circ} p^2(\text{with headpiece}) d\beta}. \quad (8)$$

In contrast to impedances with mass character and damping impedances which have monotonic decreasing curves plotted against the absolute value of the impedance, the average insertion loss with spring character “jumps” at a critical limit $|Z|_{cr}$ from positive to negative values. This limit is approximated from the case of resonance with

$$|Z|_{cr} \approx m'' \omega_0. \quad (9)$$

With an estimated thickness of the added moved air layer of $\lambda/2\pi$ we obtain

$$|Z|_{cr} \approx \rho c. \quad (10)$$

Figure 9 shows an initial increase in \overline{IL} with the increase of stiffness impedance when passing $|Z|=0$. This effect cannot be explained using the principles mentioned, so there must be an extra physical mechanism. One possible interpretation is suggested by the intensity in Fig. 10. Obviously the pressure node lies in front of the surface of the cylinder for small stiffness impedance. So the headpiece is “acoustically en-

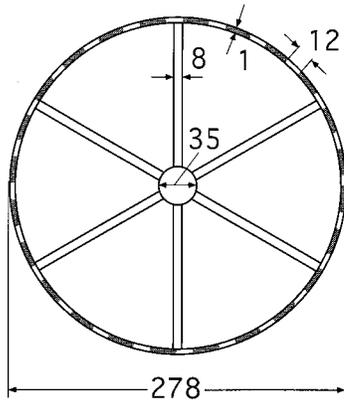


FIG. 11. Construction of the cylindrical headpiece in the form of a Helmholtz resonator.

larged'' and this explains the slight increase in insertion loss. As Fig. 9 shows, the insertion loss abruptly becomes negative if the impedance exceeds a certain critical value. Numerical calculations show that it will be approximately

$$\frac{|Z|_{\text{cr}}}{\rho c} \approx \sqrt[3]{b/\lambda}. \quad (11)$$

D. Implementation

There is a considerable range of impedance for the improvement of noise barriers through attached cylinders. Summarizing, a convenient effect is to be expected in the frequency range under consideration if

- (i) primarily small damping is considered to reduce the disadvantages in the range of larger stiffness impedances,
- (ii) a stiffness range to a limit of Z''_{cr} was considered, and
- (iii) the frequency dependence of impedance, if it has a mass behavior, should be as small as possible, with $|Z/\rho c| \leq 2$ still being acceptable.

Of course these conditions can only be fulfilled within certain frequency bands. Helmholtz resonators are suggested because impedance at their resonance frequencies is only determined by the small damping, the resulting parameters being selected over a wide frequency range with sufficiently small impedances.

1. Resonators as headpieces

Figure 11 shows a construction with a cylindrical headpiece in the form of a Helmholtz resonator with perforated surface and a segmented cavity for lateral insulation. The impedance of such a construction is

$$\frac{Z}{\rho c} = \frac{\Xi d_a}{\rho c} + j \left(\frac{\omega m''}{\rho c} - \frac{s''}{\omega \rho c} \right) \quad (12)$$

in which Ξ is the flow resistivity per unit length of the absorbing material with the thickness d_a ; m'' is the mass per unit area with $m'' = \rho(w + 1.6a)/\alpha$, where w is the thickness of the shell, a is the radius of the holes, and α is the proportion of the surface covered with holes, and s'' is the stiffness per unit area with $s''/\rho c = c/h$, where h is the effective depth

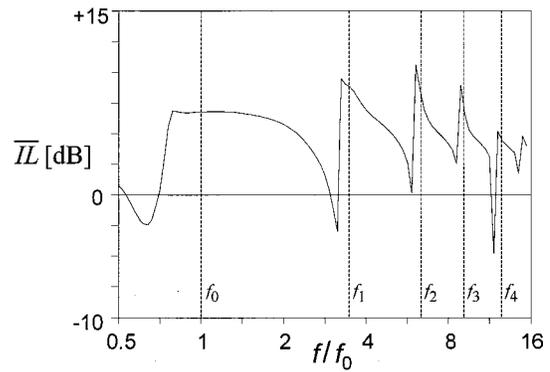


FIG. 12. Single resonator, average insertion loss \bar{IL} , $b/\lambda_0 = 1/3$; $\varphi_0 = 60$ degrees.

of the cavity, $h = b/2$, and b is the radius of the attached cylinder.

The design of this resonator follows the same principle as for absorbing resonators: the alteration $dZ/d\omega$ of the impedance in a resonant frequency is determined through the mass m'' ($dZ/d\omega = 2m''$ for $\omega = \omega_0 = \sqrt{s''/m''}$). To achieve a large frequency range with a small impedance therefore requires a sufficiently small mass coating. It is necessary to tune the desired resonance frequency

$$\omega_0 = \sqrt{\frac{\rho c^2}{m'' h}} \quad (13)$$

by keeping the required mass as small as possible and the effective depth of the cavity h as large as possible.

The theoretical calculation shows that a significant decrease in sound pressure level is feasible in the shadow region. With no damping and with consideration of resonances in depth but without lateral resonances

$$\frac{Z}{\rho c} = j \left[\frac{\omega m''}{\rho c} - \text{ctg} \left(\frac{kb}{2} \right) \right] = j \left[\frac{\omega}{\omega_0} \frac{2\pi m''}{\rho \lambda_0} - \text{ctg} \left(\pi \frac{\omega}{\omega_0} \frac{b}{\lambda_0} \right) \right]. \quad (14)$$

The multiple (depth) resonances ω_n are determined through the zeros of the impedance with

$$\frac{\omega_n b}{c} \frac{m''}{\rho b} = \text{ctg} \frac{\omega_n b}{2c}. \quad (15)$$

Figure 12 shows how the frequency dependence of impedance influences the improvement in the shadow region. For the deepest resonance frequency f_0 the radius of the cylinder is a third of the wavelength, $b/\lambda_0 = \frac{1}{3}$. The frequency response \bar{IL} can be interpreted easily from the results in the previous paragraph: in the frequency range just below a resonance there is stiffness behavior in impedance which results in a negative insertion loss. The wide regions with positive \bar{IL} correspond with a mass character in impedance through which the amount of the impedance increases to an antiresonance, describing the (nearly inefficient) case of the hard impedance.

The disadvantages in the range of stiffness of the impedance can be reduced by fixing absorbing material of a certain thickness at the hard surface end of the resonator. At lower

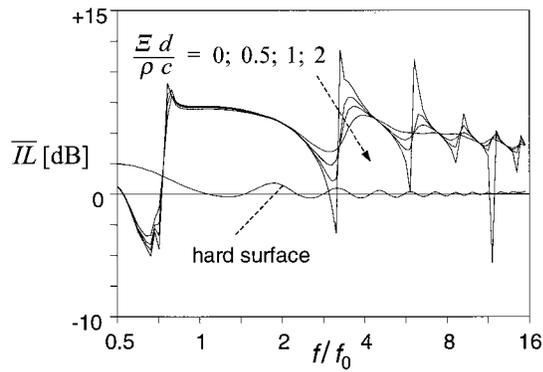


FIG. 13. Average insertion loss \overline{IL} ; $b/\lambda_0=1/3$; $\varphi_0=60$ degrees; filling factor $\gamma=0.25$.

frequencies such an absorbing layer does not yet have a negative effect on the frequency response of the impedance. Only at higher frequencies, just beginning with the first higher depth resonance, the dissipation in the absorbing layer influences the effective impedance: the unfavorable effect of the absolute impedance of stiffness will be diminished by a resistive component. The impedance Z for this headpiece is given by

$$\frac{Z}{\rho c} = \frac{j\omega m''}{\rho c} + \frac{(Z_A/\rho c)\cos[k(h-d)] + j\sin[k(h-d)]}{\cos[k(h-d)] + j(Z_A/\rho c)\sin[k(h-d)]} \quad (16)$$

$[(h-d) = \text{thickness of the remaining "free air layer"}]$. Here Z_A is the impedance of the porous layer (with the thickness d),

$$\frac{Z_A}{\rho c} = -j \frac{k_A}{k} \text{ctg}(k_A d), \quad (17)$$

with the wave number

$$k_A = k \sqrt{1 - j\Xi/\omega\rho} \quad (18)$$

in the absorbing (resistive) material ($\Xi = \text{flow resistivity per length, structure factor, and porosity are set to 1 here for simplicity}$).

Calculated average insertion losses are shown in Figs. 13 and 14. In each case the size of the cylinder b/λ_0 in the resonance frequency f_0 and the filling factor γ (= absorbing

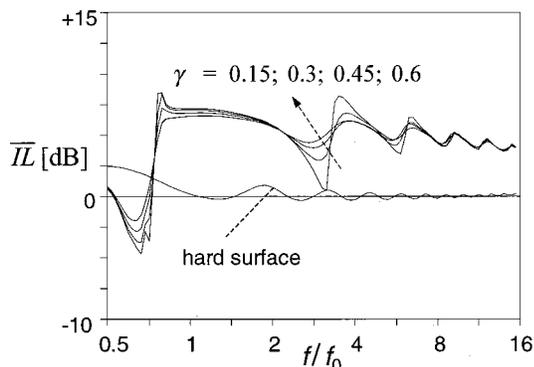


FIG. 14. Average insertion loss \overline{IL} ; $b/\lambda_0=1/3$; $\varphi_0=60$ degrees; $\Xi/\rho c=1$.

volume/total volume) is indicated. As can be seen, the frequency response of \overline{IL} is well smoothed. Particularly the periods with negative \overline{IL} are less marked, but at the price of a slight reduction in the frequency ranges with mass character of the impedance at higher frequencies. At and above the lowest resonance (the ‘Helmholtz resonance’), however, when there was a sufficiently careful application of the absorbing material, the result is hardly influenced. So there is a theoretical possibility of reducing the existing frequency selectivity of the headpiece and achieving a broadband effectiveness.

II. EXPERIMENTS

A. Construction of the cylindrical headpiece

As described, a simple Helmholtz resonator with perforated shell was applied as a headpiece for the measurements. It consists of three perforated steel tubes (length: 1 m, outside diameter: 278 mm, thickness of shell: 1 mm) with holes (diameter: 12 mm, covering 12% of surface). The cavity of the cylinder is divided into six segments by pressboard (thickness: 8 mm) with a wooden core (diameter 35 mm) (Fig. 11).

B. Equipment to measure the intensity and the insertion loss

The barrier with the attached cylindrical headpieces was tested in the anechoic room (14 m long \times 9 m wide \times 8.5 m high). The three cylindrical elements were attached on a 5-m wall of pressboard (thickness: 22 mm, mass per area 14 kg/m²). The upper edge was flushed at each side with the attached cylinders (Figs. 15 and 16). To avoid undesirable energy transport through the wall and through the attached cylindrical headpieces, a second pressboard wall (thickness: 16 mm) was fitted on the source side with rubber elements to the main wall and the floor (thickness of rubber: 1 cm). To reduce resonances between the two walls, a porous foam (thickness: 5 cm) was attached to the main wall. The segments inside the cylindrical headpiece and the space between the headpiece and the walls were completely sealed with a durable elastic material. The second wall was connected through the shell with a radial segmenting plate inside the headpiece, so the lower cavity on the source side was closed to the outside. With this arrangement energy transport through the lower cavity should be obstructed, since the energy flow would have to cross two segmenting plates to reach the receiver side.

The noise screen butts at both ends onto a highly absorbing wall of the anechoic room, creating a triangular area. To

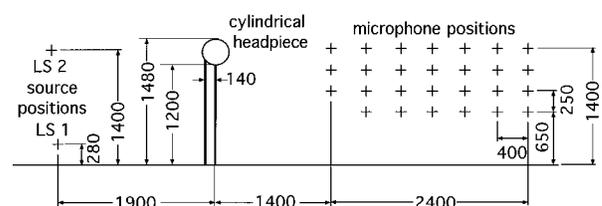


FIG. 15. Side view of the measurement equipment.

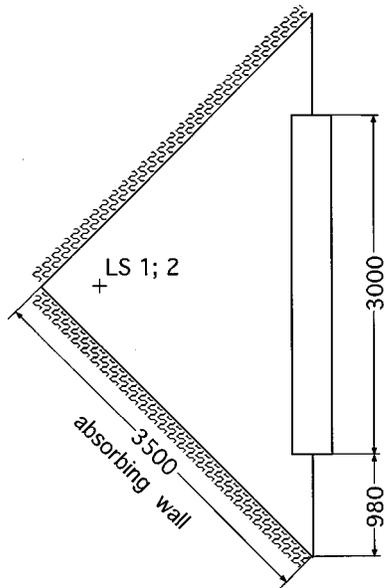


FIG. 16. Top view of the measurement equipment.

avoid diffraction of sound energy at the vertical edges of the wall, an extension was created for 35 cm on both sides with attached pressboard plates (thickness: 16 mm).

To avoid energy transport at the lower edge of the wall, the supporting grid of the anechoic room on the source side was completely covered with pressboard (thickness: 19 mm) and all slits were sealed. To reduce reflections from the ground, porous foam (thickness: 50 mm) was laid on the source side. A multiple loudspeaker ("Dodekaeder") was used as source.

Two series of measurements were taken: in one case the headpiece was covered with a heavy foil (mass per area: 5 kg/m²) and in the other case the headpiece construction was uncovered.

1. Measurement of the intensity

The intensity was measured in *x* and *y* directions at 42 measurement points near the cylindrical headpiece on a plane perpendicular to the longitudinal axis of the cylindrical elements in the center of the wall using a B&K intensity probe with 1/2-in. microphones. The cross power spectrum was

calculated. The upper frequency was 3.2 kHz and band-limited random white noise was used as source signal.

The source was positioned 190 cm from the wall and 140 cm above the ground.

The intensity was measured twice. In one case the headpiece was covered with a heavy foil and in the other case the headpiece was uncovered.

2. Local intensity in 1/3-octave bands

Figure 17(a) shows the intensity with covering, Fig. 17(b) without covering of the headpiece. In the 1/3-octave bands up to $f_m = 630$ Hz there are deteriorations due to the effect of the headpiece. From $f_m = 800$ Hz up to $f_m = 2500$ Hz obvious improvements are recognizable in the complete shadow region. Only in the 1/3-octave band with $f_m = 2000$ Hz are there small deteriorations for some microphone positions. The improvements are up to 10 dB in the 1/3-octave band with $f_m = 800$ Hz (Fig. 17).

3. Measurement of the insertion loss

The sound-induced phenomena are characterized by strong local variations of sound pressure level. Since this three-dimensional distribution will be changed by covering the headpiece, the critical examination of the acoustical effect of such headpieces for noise barriers requires numerous measurement positions for the observation of the acoustic field. For this reason the sound pressure levels were measured at 32 microphone positions behind the barrier in the frequency range up to 3.2 kHz (for details see Ref. 29).

The following measurements represent the insertion loss *ILm* in dB resulting from the following level difference:

$$ILm = L_{(with\ covering)} - L_{(without\ covering)} \quad (19)$$

$L_{(with\ covering)}$ characterizes the level measured when the cylindrical headpiece was covered with the heavy foil (mass per area: 5 kg/m²) and $L_{(without\ covering)}$ characterizes the level when the cylindrical headpiece was not covered.

Insertion losses *ILm* with a positive sign indicate the real benefit obtained from the effect of the cylindrical headpiece compared with a construction of the same geometrical dimensions.

To control the stability of the source equipment, the sound pressure level was measured with a reference micro-

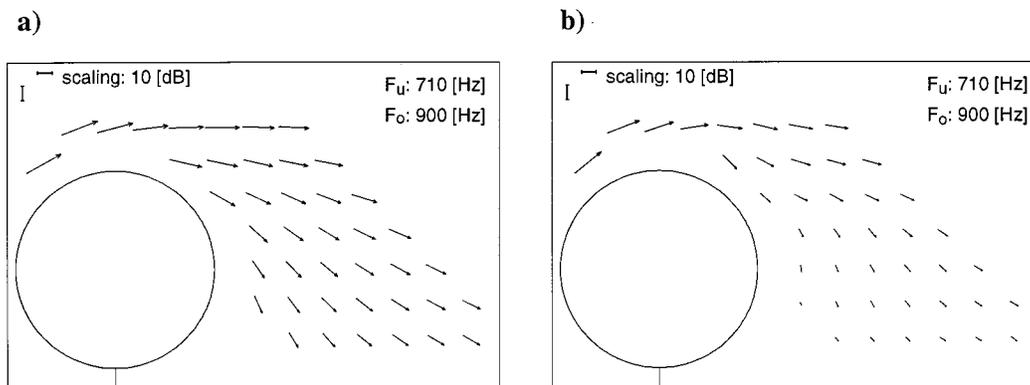


FIG. 17. Distribution of the intensity; 1/3-octave band, $f_m = 800$ Hz; (a) with covering and (b) without covering.

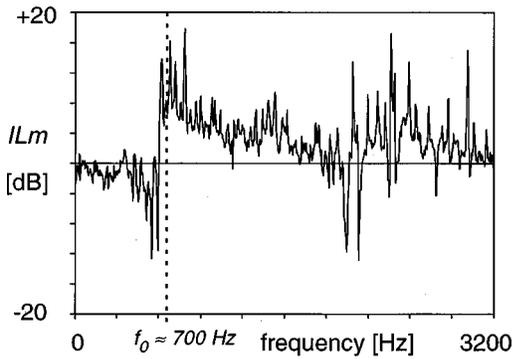


FIG. 18. Insertion loss IL_m (with and without covering): height of source: 0.28 m; distance to the wall: 2.60 m; height above ground: 0.65 m; and diffraction angle: 50 degrees.

phone at the beginning and the end of a measurement series. The differences between these reference levels were less than 1 dB.

Band-limited random noise was used as source signal and the source was positioned 190 cm from the wall and 28 or 140 cm above the floor.

In Fig. 18 the insertion loss was shown for one microphone position depending on frequency.

In Figs. 19–22 the insertion loss in the frequency range from 355 to 2800 Hz is shown in one $\frac{1}{3}$ -octave and three octave bands as a function of the so-called diffraction angle β defined in Fig. 2. The squares and the “plus” signs in the figures are for source position 1 (height: 0.28 m) and the rhombs and the crosses are for source position 2 (height: 1.40 m). The “plus” signs and the crosses correspond to five positions at a greater distance (7.83 m) from the wall.

The results can be summarized as followed:

- (1) Over the frequency range from 280 to 560 Hz the cylindrical headpiece on average has a negative effect of 0 to 5 dB in the $\frac{1}{3}$ -octave bands, which is worst in the band with the center frequency of 500 Hz. In the octave band from 355 to 710 Hz the negative effects are from 0 to 3 dB (see also Fig. 20).
- (2) Over the two octaves from 710 Hz (center frequency $f_m = 800$ Hz) up to the highest measured $\frac{1}{3}$ -octave band ($f_m = 2500$ Hz) there are only improvements. In the $\frac{1}{3}$ -octave band with the center frequency of 800 Hz the

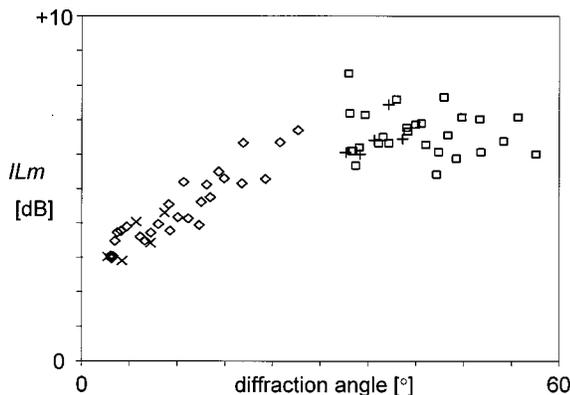


FIG. 19. Insertion loss IL_m (with and without covering); $\frac{1}{3}$ -octave band, $f_m = 800$ Hz.

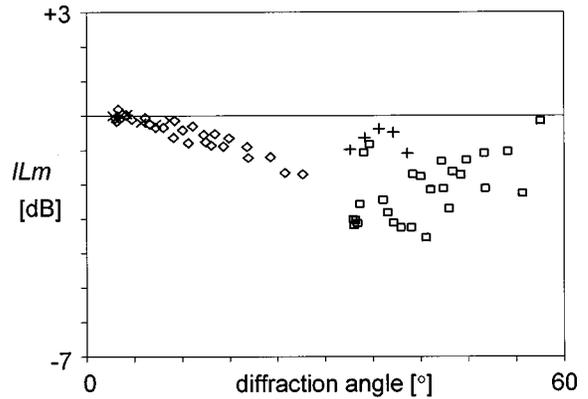


FIG. 20. Insertion loss IL_m (with and without covering); octave band, $f_m = 500$ Hz.

improvements are from 3 to 7 dB, increasing with an increase in the diffraction angle up to about 40 degrees (Fig. 19). In the range of the diffraction angle from 35 up to 60 degrees the improvements are to more than 5 dB with an average of 7 dB. For the following three $\frac{1}{3}$ -octave bands ($f_m = 1000$ Hz up to $f_m = 1600$ Hz) an improvement with the diffraction angle was also obtained of up to 5 dB. In the highest $\frac{1}{3}$ -octave band ($f_m = 2500$ Hz) the improvements amount to 1 to 3 dB, increasing slightly with the diffraction angle. When combining the frequency band from 710 to 1400 Hz in one octave band (Fig. 21) it can be shown that improvements from 2 to 5 dB can be observed depending on the effect of the headpiece, increasing with the diffraction angle. For the next octave band (1400–2800 Hz) (Fig. 22) improvements from 1 to 3 dB can be observed. Again in this case an increase of the improvements with the diffraction angle can be recognized.

C. Conclusion of the results of the measurements

The measurements of the intensity and the insertion loss show the obvious potential of the cylindrical headpiece.

Summarizing, at the “best” $\frac{1}{3}$ -octave band with $f_m = 800$ Hz an additional shielding up to 7 dB could be obtained at high diffraction angles due to the headpiece. For the

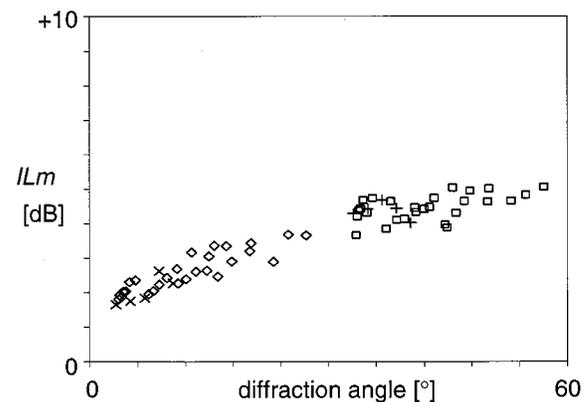


FIG. 21. Insertion loss IL_m (with and without covering); octave band, $f_m = 1000$ Hz.

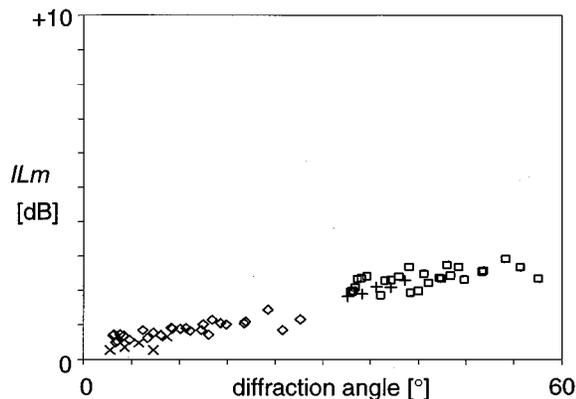


FIG. 22. Insertion loss IL_m (with and without covering); octave band, $f_m = 2000$ Hz.

“best” octave band with $f_m = 1000$ Hz and high diffraction angles the improvements are up to 5 dB. The results for the headpiece do not include the effects from the longer geometrical path of the sound over the headpiece, as is common in the literature. However, the results remain below the theoretical expectations, especially at larger angles of diffraction. The next section tries to explain this and to find a strategy for future improvements.

III. DISCUSSION

A. Influence of transverse coupling in the cylindrical segments

The essential aspect of diffraction at noise barriers is that the surrounding volume and energy flow near the edge is the determining source for the shadow region. Therefore, it is important to reduce the energy flow transported tangentially past the headpiece, simultaneously realizing a “turn away” of the incident sound to other regions.

The (nearly) ideal headpiece for noise barriers therefore has a constant impedance along its circumference on the sur-

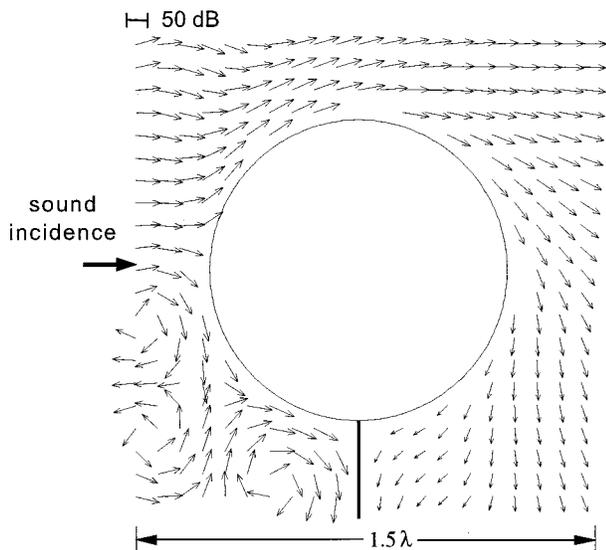


FIG. 23. Distribution of the intensity; hard-surfaced cylinder ($Z \rightarrow \infty$).

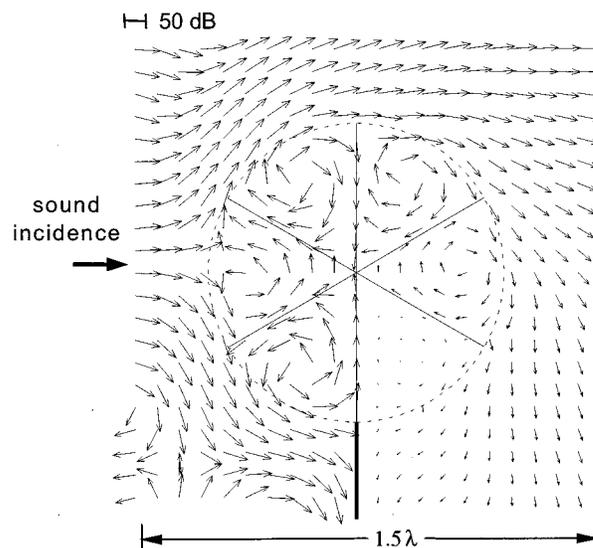


FIG. 24. Distribution of the intensity; six segments (calculated in the fundamental resonance frequency f_0).

face of $Z=0$. Because there is also no developing velocity tangential to the surface, there is no volume flow or transport from active intensity near the surface, so that the desired effect will be achieved.

Tangential transient effects near the surface therefore always result in a worsening of the shielding effect. They could be due to a divergence from the “ideal value” impedance $Z=0$ as described in Sec. IC. The energy transport along the headpiece could also occur through locally varying transverse distribution in the segments (with finite width) of the cylindrical headpiece. To examine the influence of transverse coupling, the sound field was calculated inside and in front of a covered headpiece in which the local field distribution in the independent segments would be considered using wave equation solutions and with the inner part of the

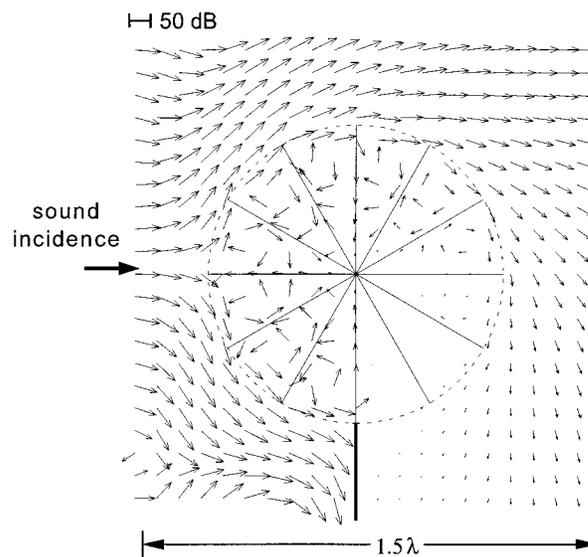


FIG. 25. Distribution of the intensity; 12 segments (calculated in the fundamental resonance frequency f_0).

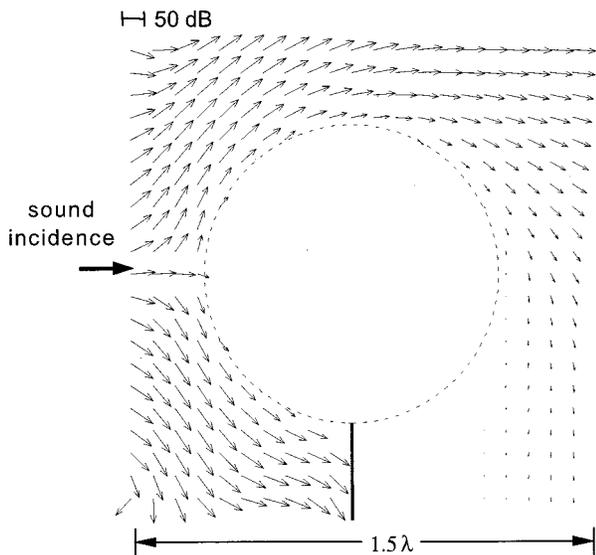


FIG. 26. Distribution of the intensity; soft-surfaced cylinder ($Z=0$).

cylindrical headpiece separated from the outer space by covering with a mass per area m'' .

A graphical description of the results is shown in Figs. 23–26 with the local distribution of the active intensity ($b/\lambda=0.5$ in all cases). Figure 23 shows the case of a hard surface, Fig. 26 the case of a soft surface ($Z=0$), and Figs. 24 and 25 demonstrate the realistic cases of a segmented cylindrical headpiece with 6 or 12 equal segments. It can be seen that compared with the “ideal case” $Z=0$, six segments cause a distinctly higher energy transport. The improvement of a “real headpiece” is therefore smaller than expected for the “ideal case” $Z=0$. Doubling the number of segments (Fig. 25) produces a case close to the locally independent impedance $Z=0$. It is clear that the pressure buildup at the surface of the cylindrical headpiece, which is not possible for a constant impedance $Z=0$, creates an infiltration of the field into the segments of the real construction. The sound field is joining more closely to the surface $r=b$, resulting in a higher energy transport to the shadow region.

The results can be seen on the calculated far-field insertion loss (compared with the screens without headpiece, $b=0$). Figures 27 and 28 show the curve depending on the diffraction angle. The frequency was varied in $\frac{1}{3}$ -octave bands. On the left side is the “real case” and on the right side the “ideal case” $\partial Z/\partial\varphi=0$. The parameters are approximately in accordance with the test assembly when the frequency f is in the range $700\text{ Hz} \leq f \leq 1400\text{ Hz}$ ($b/\lambda = \frac{1}{3}$ corresponds to the resonance frequency at $f=700\text{ Hz}$). Figure 27 shows the comparison with the case with six segments which was also used for the measurements. The real effect is distinctly less than that expected from the idealization, and at the same time the negative insertion loss at the highest frequency approaches zero ($\beta < 60$ degrees). These tendencies are reflected well in the earlier results mentioned. Only if the number of segments is increased, insertion loss does approximate the optimum possible value. Figure 28 shows the case of 12 segments at which nearly no differences to the constant impedance were observed. Further evaluations with eight

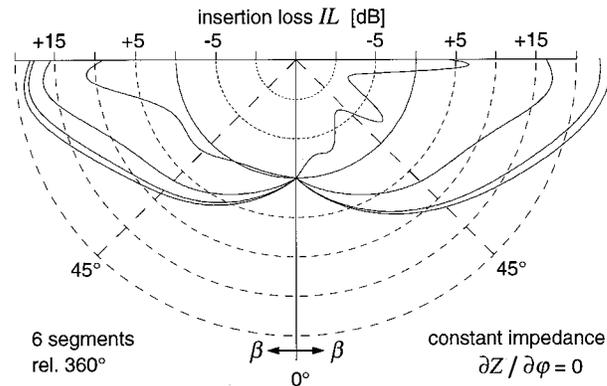


FIG. 27. Far-field insertion loss IL ; $b/\lambda = \frac{1}{3}$ specify the lowest resonance ($Z=0$). Left side: segmented cylinder with six caverns. Right side: cylinder with constant impedance. $\varphi_0=75$ degrees; from outside to inside: $b/\lambda = \frac{1}{3}; \frac{1}{3} \times 2^{1/3}; \frac{1}{3} \times 2^{2/3}; \frac{2}{3}$.

and ten segments show a gradual transition between Figs. 27 and 28. For 10 segments the insertion losses are only slightly lower than for 12 segments. It is concluded that the influence of the limited width of the segments has been underestimated in the past. In the next phase of the project this effect will receive more attention.

Although the simple structure with six segments remained below the optimum, relevant results were nevertheless obtained.

IV. CONCLUSION

A principle for the improvement of noise barriers with (nearly) no increase in height through a “sound deflecting surface design” of headpieces was shown—expressed by the acoustical surface impedance “near zero.” A deflection of the incident sound is possible in principle away from the shadow region. The acoustical shadow region is then more pronounced. In this case the absorption at the headpieces is only of secondary importance. The predicted effects could largely be verified for a cylinder with Helmholtz resonators. The design of the cylinder (in segments) was found to be less than optimum.

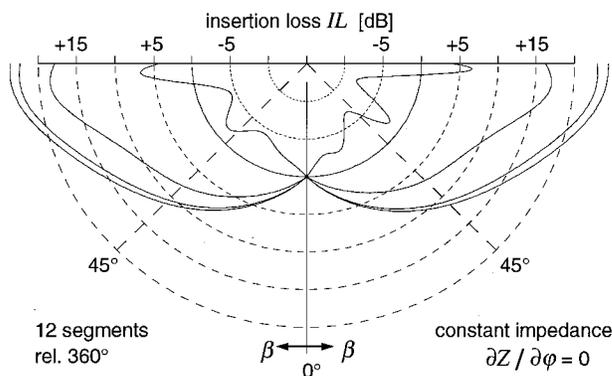


FIG. 28. Far-field insertion loss IL . Left side: segmented cylinder with 12 segments. Right side: cylinder with constant impedance. $b/\lambda = \frac{1}{3}$ specify the lowest resonance ($Z=0$); $\varphi_0=75$ degrees; from outside to inside: $b/\lambda = \frac{1}{3}; \frac{1}{3} \times 2^{1/3}; \frac{1}{3} \times 2^{2/3}; \frac{2}{3}$.

Future work will be aimed at improving results and increasing the available bandwidth by the selection of proper frequency responses of the impedance.

ACKNOWLEDGMENTS

This work was supported by the DFG (Deutsche Forschungsgemeinschaft).

- ¹A. Sommerfeld, *Vorlesungen über Theoretische Physik* (Akademische Verlagsgesellschaft, Leipzig, 1964), Vol. IV, pp. 215–238.
- ²E. Skudrzyk, *The Foundations of Acoustics* (Springer Verlag, Wien, 1971), Chap. XXV, pp. 557–591.
- ³L. Beranek, *Noise and Vibration Control* (Institute of Noise Control Engineering, Washington, DC, 1988), Chap. 7, pp. 174–178.
- ⁴G. F. Butler, “A Note on Improving the Attenuation Given by a Noise Barrier,” *J. Sound Vib.* **32**(3), 367–369 (1974).
- ⁵A. D. Rawlins, “Diffraction of Sound by a Rigid Screen with a Soft or Perfectly Absorbing Edge,” *J. Sound Vib.* **45**(1), 53–67 (1976).
- ⁶L. S. Wirt, “The Control of Diffracted Sound by Means of Thndners (Shaped Noise Barriers),” *Acustica* **42**(2), 73–88 (1979).
- ⁷D. N. May and M. M. Osman, “The Performance of Sound Absorptive, Reflective, and T-Profile Noise Barriers in Toronto,” *J. Sound Vib.* **71**(1), 65–71 (1980).
- ⁸D. N. May and M. M. Osman, “Highway Noise Barriers: New Shapes,” *J. Sound Vib.* **71**(1), 73–101 (1980).
- ⁹S. I. Hayek, “Mathematical Modeling of Complex Noise Barriers,” *Proc. Inter-Noise*, 803–806 (1982).
- ¹⁰K. Iida, Y. Kondoh, and Y. Okada, “Research on a Device for Reducing Noise,” *Transp. Res. Rec.* **983**, 51–54 (1983).
- ¹¹D. A. Hutchins, H. W. Jones, and L. T. Russell, “Model Studies of Barrier Performance in the Presence of Ground Surfaces: II. Different Shapes,” *J. Acoust. Soc. Am.* **75**, 1817–1826 (1984).
- ¹²K. Yamamoto, K. Taya, M. Yamashita, and K. Tanaka, “Reduction of Road Traffic Noise by Absorptive Cylinder Adapted at the Top of a Barrier,” *Proc. Inter-Noise*, 349–352 (1989).
- ¹³K. Fujiwara, “Influences of Covering Materials on the Noise Reducing Efficiency of a Barrier with Absorbing Edge,” *Proc. Inter-Noise*, 353–356 (1989).
- ¹⁴S. I. Hayek, “Mathematical Modeling of Absorbent Highway Noise Barriers,” *Appl. Acoust.* **31**, 77–100 (1990).
- ¹⁵D. C. Hothersall, D. H. Crombie, and S. N. Chandler-Wilde, “The Performance of T-Profile and Associated Noise Barriers,” *Appl. Acoust.* **32**, 269–287 (1991).
- ¹⁶K. Fujiwara and N. Furuta, “Sound Shielding Efficiency of a Barrier with a Cylinder at the Edge,” *Noise Control Eng. J.* **37**(1), 5–11 (1991).
- ¹⁷S. Ise, H. Yano, and H. Tachibana, “Basic Study on Active Noise Barriers,” *J. Acoust. Soc. Jpn.* **12**(6), 299–306 (1991).
- ¹⁸A. I. Papadopoulos and C. G. Don, “A Study of Barrier Attenuation by Using Acoustic Impulses,” *J. Acoust. Soc. Am.* **90**, 1011–1018 (1991).
- ¹⁹D. C. Hothersall, S. N. Chandler-Wilde, and M. N. Hajmirzae, “Efficiency of Single Noise Barriers,” *J. Sound Vib.* **146**(2), 303–322 (1991).
- ²⁰A. Omoto, F. Tsutsumi, and K. Fujiwara, “An Effect of the Active Canceling of the Edge Potential on the Noise Barriers with Different Acoustic Surface Conditions,” *J. Acoust. Soc. Jpn.* **48**(12), 854–862 (1992).
- ²¹M. Hasebe, “A Study on the Sound Reduction by T-Profile Noise Barrier,” *J. Acoust. Soc. Jpn. (E)* **14**(2), 113–115 (1993).
- ²²A. Omoto and K. Fujiwara, “A Study of an Actively Controlled Noise Barrier,” *J. Acoust. Soc. Am.* **94**, 2173–2180 (1993).
- ²³D. H. Crombie and D. C. Hothersall, “The Performance of Multiple Noise Barriers,” *J. Sound Vib.* **176**(4), 459–473 (1994).
- ²⁴G. R. Watts, “Acoustic Performance of new Designs of Traffic Noise Barriers: Full Scale Tests,” *J. Sound Vib.* **177**(3), 289–305 (1994).
- ²⁵M. Hasebe, “Sound Reduction by a T-Profile Noise Barrier,” *J. Acoust. Soc. Jpn. (E)* **16**(3), 173–179 (1995).
- ²⁶P. McIver and A. D. Rawlins, “Diffraction by a Rigid Barrier With a Soft or Perfectly Absorbent End,” *Wave Motion* **22**(4), 387–402 (1995).
- ²⁷G. R. Watts, “Acoustic Performance of a Multiple Edge Noise Barrier Profile at Motorway Sites,” *Appl. Acoust.* **47**, 47–66 (1996).
- ²⁸M. Möser, “Die Wirkung von zylindrischen Aufsätzen an Schallschirmen,” *Acustica* **81**, 565–586 (1995).
- ²⁹M. Möser and R. Volz, “Schallschirme mit zylinderförmigen Aufsätzen—erste Ergebnisse,” *Proc. DAGA*, 136–137 (1997).

Wideband quantitative ultrasonic imaging by time-domain diffraction tomography

T. Douglas Mast

Applied Research Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 3 April 1999; revised 27 August 1999; accepted 30 August 1999)

A quantitative ultrasonic imaging method employing time-domain scattering data is presented. This method provides tomographic images of medium properties such as the sound speed contrast; these images are equivalent to multiple-frequency filtered-backpropagation reconstructions using all frequencies within the bandwidth of the incident pulse employed. However, image synthesis is performed directly in the time domain using coherent combination of far-field scattered pressure waveforms, delayed and summed to numerically focus on the unknown medium. The time-domain method is more efficient than multiple-frequency diffraction tomography methods, and can, in some cases, be more efficient than single-frequency diffraction tomography. Example reconstructions, obtained using synthetic data for two- and three-dimensional scattering of wideband pulses, show that the time-domain reconstruction method provides image quality superior to single-frequency reconstructions for objects of size and contrast relevant to medical imaging problems such as ultrasonic mammography. The present method is closely related to existing synthetic-aperture imaging methods such as those employed in clinical ultrasound scanners. Thus, the new method can be extended to incorporate available image-enhancement techniques such as time-gain compensation to correct for medium absorption and aberration correction methods to reduce error associated with weak scattering approximations. © 1999 Acoustical Society of America.

[S0001-4966(99)04612-3]

PACS numbers: 43.20.Fn, 43.60.Rw, 43.80.Vj, 43.20.Px [ANN]

INTRODUCTION

Quantitative imaging of tissue properties is a potentially useful technique for diagnosis of cancer and other pathological conditions. Inverse scattering methods such as diffraction tomography can provide quantitative reconstruction of tissue properties including sound speed, density, and absorption. However, although previous inverse scattering methods have achieved high resolution and quantitative accuracy, such methods have not yet been incorporated into commercially successful medical ultrasound imaging systems.

Current inverse scattering methods are lacking in several respects with respect to conventional B-scan and synthetic aperture imaging techniques. Previous methods of diffraction tomography, including methods based on the Born and Rytov approximations,^{1,2} and higher-order nonlinear approaches,^{3,4} have usually been based on single-frequency scattering, while current diagnostic ultrasound scanners employ wideband time-domain signals. The use of wideband information in image reconstruction is known to provide increased point and contrast resolution,^{5,6} both of which are important for medical diagnosis.^{5,7,8}

Several approaches have been used to incorporate wideband scattering information into quantitative ultrasonic imaging. One group of methods employs time-domain tomography based on Radon-transform relationships that hold (under the assumption of weak scattering) between scattered acoustic fields and the reflectivity or scattering strength of the medium. Pioneering work in this area^{9,10} employed measurements of reflectivity in pulse-echo mode, while later studies have incorporated aberration correction^{11,12} and multiple-angle scattering measurements.^{13,14} A limitation of

these methods, however, is that the Radon transform relationship strictly holds only when the medium is insonified by an impulsive (infinite bandwidth) wave. When pulses of finite bandwidth are employed, image quality can degrade significantly.¹⁵

A number of linear and nonlinear diffraction tomography methods have been implemented using scattering data for a number of discrete frequencies (e.g., Refs. 16–19). Although use of multiple-frequency data provides improvements in image quality, computational requirements for multiple-frequency imaging are typically large because the computational cost is proportional to the number of frequencies employed. To achieve image quality competitive with present diagnostic scanners, together with quantitative imaging of tissue properties, present frequency-domain methods may require solution of the inverse scattering problem for many frequencies within the bandwidth of the transducer employed. This approach thus demands a high computational cost, so that high-quality real-time imaging may not be presently feasible using current frequency-domain inverse scattering methods.

Very few previous workers have investigated direct use of time-domain waveform data for inverse scattering methods analogous to frequency-domain diffraction tomography. Several methods^{20,21} have used frequency decomposition of scattered pulses to construct a wideband estimate of the spatial Fourier transform of an unknown medium; after appropriate averaging and interpolation, this transform can be inverted to obtain a wideband Born reconstruction of the medium. A study reported in Ref. 22 has showed that broadband synthetic aperture imaging using linear arrays is closely

related to inverse scattering using filtered backpropagation. A related method, suggested in Ref. 23, provides a time-domain reconstruction algorithm that employs filtered backpropagation of scattered waveforms measured on a circular boundary. However, the time domain reconstruction formula of Ref. 23 yields reconstructions that are less general than multiple-frequency reconstructions obtained using the same signal bandwidth.

Another approach, related both to multiple-frequency methods and direct time-domain methods, has recently been presented.²⁴ This work extends the eigenfunction method of Ref. 19 to use the full bandwidth of the incident pulse waveform. In the extended method, eigenfunctions and eigenvalues of a scattering operator are computed to obtain a frequency-dependent representation of the scattering medium. Fourier synthesis is then applied to obtain a time-dependent estimate of the medium. A cross-correlation operation removes the time dependence of the estimate as well as its dependence on the waveform employed.

The present paper offers a new approach to wideband quantitative imaging: a time-domain inverse scattering method that overcomes some of the limitations of previous frequency-domain and time-domain quantitative imaging methods. The new method provides tomographic reconstructions of unknown scattering media using the entire available bandwidth of the signals employed. Reconstructions are performed using scattering data measured on a surface surrounding the region of interest, so that the method is well suited to ultrasonic mammography. The reconstruction algorithm is derived as a simple delay-and-sum formula similar to synthetic-aperture algorithms employed in conventional clinical scanners. However, unlike current clinical scanners, the present method can provide quantitative images of tissue properties such as the spatially dependent sound speed. Reconstructions obtained in this manner are equivalent to reconstructions obtained by combining conventional frequency-domain diffraction tomography reconstructions for all frequencies within the signal bandwidth of interest. The current method, however, can be even more efficient than single-frequency diffraction tomography. The method is applicable both to two-dimensional and three-dimensional image reconstruction. The direct time-domain nature of the reconstruction algorithm allows straightforward incorporation of depth- and frequency-dependent amplitude correction to compensate for medium absorption as well as aberration correction methods to overcome limitations of the Born approximation.

I. THEORY

A. The time-domain reconstruction algorithm

An inverse scattering algorithm, applicable to quantitative imaging of tissue and other inhomogeneous media, is derived below. For simplicity of derivation, the medium is modeled as a fluid medium defined by the sound speed contrast function

$$\gamma(\mathbf{r}) = \frac{c_0^2}{c(\mathbf{r})^2} - 1, \quad (1)$$

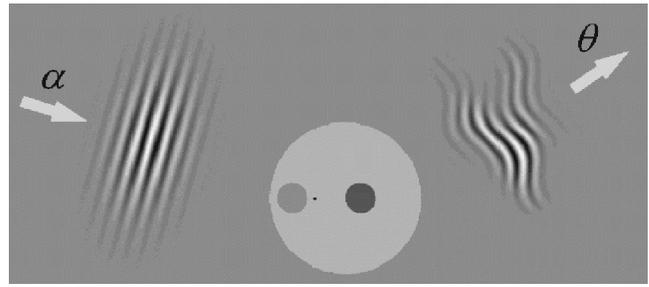


FIG. 1. Scattering configuration. An incident pressure pulse $f(t - \boldsymbol{\alpha} \cdot \mathbf{r}/c)$ is scattered by an inhomogeneous medium and the time-domain scattered pressure $p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t)$ is measured at a radius R in the far field.

where c_0 is a background sound speed and $c(\mathbf{r})$ is the spatially dependent sound speed defined at all points \mathbf{r} . For the scope of the initial derivation, the medium is assumed to have constant density, no absorption, and weak scattering characteristics; extensions to the reconstruction algorithm that overcome these limiting assumptions are discussed in the following section.

For the model of the scattering medium represented by Eq. (1), the time-domain scattered acoustic pressure $p_s(\mathbf{r}, t)$ obeys the wave equation²⁵

$$\nabla^2 p_s(\mathbf{r}, t) - \frac{1}{c_0^2} \frac{\partial^2 p_s(\mathbf{r}, t)}{\partial t^2} = \frac{\gamma(\mathbf{r})}{c_0^2} \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2}, \quad (2)$$

where $p(\mathbf{r}, t)$ is the total acoustic pressure in the medium.

The scattering configuration considered here is sketched in Fig. 1. The medium is subjected to a pulsatile plane wave propagating in the direction of the unit vector $\boldsymbol{\alpha}$,

$$p_{\text{inc}}(\mathbf{r}, \boldsymbol{\alpha}, t) = f(t - \mathbf{r} \cdot \boldsymbol{\alpha}/c_0), \quad (3)$$

where f is the time-domain waveform and c_0 is the background sound speed. The scattered wavefield $p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t)$ is measured at a fixed radius R in the far field, where $\boldsymbol{\theta}$ corresponds to the direction unit vector of a receiving transducer element. (Alternatively, if scattering measurements are made in the near field, the far-field acoustic pressure can be computed using exact transforms that represent propagation through a homogeneous medium.¹⁶)

A general time-domain solution for the wave equation (2), valid for two-dimensional (2D) or three-dimensional (3D) scattering, is then

$$p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t) = \int_{-\infty}^{\infty} \hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega) e^{-i\omega t} d\omega, \quad (4)$$

where $\hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega)$ is a single frequency component of the scattered wavefield,

$$\hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t) e^{i\omega t} dt, \quad (5)$$

given exactly by²⁵

$$\begin{aligned} \hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega) &= k^2 \hat{f}(\omega) \int \mathbf{G}_0(R\boldsymbol{\theta} - \mathbf{r}_0, \omega) \\ &\quad \times \gamma(\mathbf{r}_0) \hat{p}(\mathbf{r}_0, \boldsymbol{\alpha}, \omega) dV_0. \end{aligned} \quad (6)$$

In Eq. (6), k is the wave number ω/c_0 and $\hat{p}(\mathbf{r}_0, \boldsymbol{\alpha}, \omega)$ is the total acoustic pressure associated with the unit-amplitude incident plane wave $e^{ik\boldsymbol{\alpha} \cdot \mathbf{r}_0}$. The integral in Eq. (6) is taken over the entire support of γ in \mathbb{R}^2 for 2D scattering or in \mathbb{R}^3

for 3D scattering. The free-space Green's function, represented by G_0 in Eq. (6), is²⁶

$$G_0(\mathbf{r}, \omega) = \frac{i}{4} H_0^{(1)}(kr) \quad \text{for 2D scattering}$$

and (7)

$$G_0(\mathbf{r}, \omega) = \frac{e^{ikr}}{4\pi r} \quad \text{for 3D scattering,}$$

where $H_0^{(1)}$ is the zeroth-order Hankel function of the first kind and r is the magnitude of the vector \mathbf{r} .

The far-field scattered pressure, when specified for all incident-wave directions $\boldsymbol{\alpha}$, measurement directions $\boldsymbol{\theta}$, and times t , comprises the data set to be used for reconstruction of the unknown medium. The inverse scattering problem, specified by Eq. (6) for a single frequency component, is to reconstruct the unknown medium contrast $\gamma(\mathbf{r})$ using the measured data $\hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega)$.

The starting point for the present time-domain inverse scattering method is conventional single-frequency diffraction tomography. Under the assumption of weak scattering, one can make the Born approximation, in which the total pressure $\hat{p}(\boldsymbol{\alpha}, \omega)$ in Eq. (6) is replaced by the plane wave $e^{ik\mathbf{r}\cdot\boldsymbol{\alpha}}$. For scattering measurements made at a radius R in the far field, the linearized inverse problem of Eq. (6) can be then solved for any frequency component using filtered backpropagation,^{2,16,27} i.e.,

$$\gamma_B(\mathbf{r}, \omega) = \frac{\hat{\mu}(\omega)e^{-ikR}}{\hat{f}(\omega)} \iint \Phi(\boldsymbol{\theta}, \boldsymbol{\alpha}) \times \hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega)e^{ik(\boldsymbol{\theta}-\boldsymbol{\alpha})\cdot\mathbf{r}} dS_\alpha dS_\theta, \quad (8)$$

where

$$\hat{\mu}(\omega) = -\sqrt{\frac{kR}{8i\pi^3}},$$

$$\Phi(\boldsymbol{\theta}, \boldsymbol{\alpha}) = |\sin(\theta - \alpha)| \quad \text{in 2D,}$$

and (9)

$$\hat{\mu}(\omega) = \frac{kR}{4\pi^3}, \quad \Phi(\boldsymbol{\theta}, \boldsymbol{\alpha}) = |\boldsymbol{\theta} - \boldsymbol{\alpha}| \quad \text{in 3D.}$$

Each surface integral in Eq. (8) is performed over the entire measurement circle for the 2D case and over the entire measurement sphere for the 3D case. Equation (8) provides an exact solution to the linearized inverse scattering problem for a single frequency component of the scattered wavefield $p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t)$. The resulting reconstruction, $\gamma_B(\mathbf{r}, \omega)$, has spatial frequency content limited by the "Ewald sphere" of radius $2k$ in wavespace.¹

To improve upon the single-frequency formulas specified by Eq. (8), one can extend the spatial-frequency content of reconstructions by exploiting wideband scattering information. The method outlined here synthesizes a "multiple-frequency" reconstruction $\gamma_M(\mathbf{r})$ by formally integrating single-frequency reconstructions $\gamma_B(\mathbf{r}, \omega)$ over a range of frequencies ω . A generalized formula for this approach can be written

$$\gamma_M(\mathbf{r}) = \frac{\int_0^\infty \hat{g}(\omega) \gamma_B(\mathbf{r}, \omega) d\omega}{\int_0^\infty \hat{g}(\omega) d\omega}, \quad (10)$$

where $\hat{g}(\omega)$ is an appropriate frequency-dependent weighting function. In practice, the weighting function $\hat{g}(\omega)$ is chosen to be bandlimited because (for a given set of physical scattering measurements) the frequency-dependent contrast $\gamma_B(\mathbf{r}, \omega)$ can only be reliably reconstructed for a finite range of frequencies ω associated with the spectra of the incident waves employed. Thus, the integrands in Eq. (10) are non-zero only over the support of $\hat{g}(\omega)$ and the corresponding integrals are finite.

Using Eq. (8), and making the definition

$$N \equiv 2 \int_0^\infty \hat{g}(\omega) d\omega, \quad (11)$$

Eq. (10) can be written in the form

$$\gamma_M(\mathbf{r}) = \frac{2}{N} \int_0^\infty \hat{g}(\omega) \frac{\hat{\mu}(\omega)e^{-ikR}}{\hat{f}(\omega)} \iint \Phi(\boldsymbol{\theta}, \boldsymbol{\alpha}) \times \hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega)e^{ik(\boldsymbol{\theta}-\boldsymbol{\alpha})\cdot\mathbf{r}} dS_\alpha dS_\theta d\omega. \quad (12)$$

If the frequency weight $\hat{g}(\omega)$ is now specified to incorporate the incident-pulse spectrum $\hat{f}(\omega)$ and to compensate for the frequency- and dimension-dependent coefficient $\hat{\mu}(\omega)$,

$$\hat{g}(\omega) = \frac{\hat{f}(\omega)}{\hat{\mu}(\omega)}, \quad (13)$$

Eq. (12) reduces to the form

$$\gamma_M(\mathbf{r}) = \frac{2}{N} \iint \Phi(\boldsymbol{\theta}, \boldsymbol{\alpha}) \int_0^\infty \hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega) \times e^{-ik[R+(\boldsymbol{\alpha}-\boldsymbol{\theta})\cdot\mathbf{r}]} d\omega dS_\alpha dS_\theta. \quad (14)$$

The choice of frequency weight from Eq. (13) allows the multiple-frequency reconstruction formula of Eq. (12) to be greatly simplified. Specifically, the inner integral of Eq. (14) resembles a weighted inverse Fourier transform of the frequency-domain scattered field $\hat{p}(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega)$. To obtain an explicit time-domain expression for $\gamma_M(\mathbf{r})$, Eq. (14) can be rewritten using the definition of $\hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega)$ from Eq. (5) to yield

$$\gamma_M(\mathbf{r}) = \frac{1}{N} \iint \Phi(\boldsymbol{\theta}, \boldsymbol{\alpha}) \times \mathbf{L} \left[p_s \left(\boldsymbol{\theta}, \boldsymbol{\alpha}, R/c_0 + \frac{(\boldsymbol{\alpha}-\boldsymbol{\theta})\cdot\mathbf{r}}{c_0} \right) \right] dS_\alpha dS_\theta, \quad (15)$$

where \mathbf{L} denotes the linear operator

$$\mathbf{L}[\psi(t)] = 2 \int_0^\infty \hat{\psi}(\omega) e^{-i\omega t} d\omega \quad (16)$$

and $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$ using the definition from Eq. (5).

Using the conjugate symmetry of $\hat{\psi}(\omega)$ [i.e., $\hat{\psi}(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega) = \hat{\psi}^*(\boldsymbol{\theta}, \boldsymbol{\alpha}, -\omega)$ for any real $\psi(t)$], the real part of

$\mathbf{L}[\psi(t)]$ is shown to be simply $\psi(t)$. Similarly, using the convolution theorem as well as the conjugate symmetry of $\psi(t)$, the imaginary part of $\mathbf{L}[\psi(t)]$ is seen to be an inverse Hilbert transform²⁸ of $\psi(t)$,

$$\text{Im}[\mathbf{L}[\psi(t)]] = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\psi(\tau)}{t-\tau} d\tau = \mathbf{H}^{-1}[\psi(t)]. \quad (17)$$

This transform, also known as a quadrature filter, applies a phase shift of $\pi/2$ to each frequency component of the input signal.

Thus, the time-domain reconstruction formula can finally be written

$$\gamma_M(\mathbf{r}) = \frac{1}{N} \iint \Phi(\boldsymbol{\theta}, \boldsymbol{\alpha}) \left(p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \tau) + i\mathbf{H}^{-1}[p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \tau)] \right) dS_\alpha dS_\theta, \quad (18)$$

where

$$\tau = R/c_0 + \frac{(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{r}}{c_0}. \quad (19)$$

The direction-dependent weight $\Phi(\boldsymbol{\theta}, \boldsymbol{\alpha})$, which is the same as the ‘‘filter’’ employed in single-frequency filtered back-propagation, is given for the 2D and 3D cases by Eq. (9).

Equation (18) is notable in several respects. First, it provides a linearized reconstruction that employs scattering information from the entire signal bandwidth without any frequency decomposition of the scattered wavefield. Second, the delay term τ corresponds exactly to the delay required to construct a focus at the point \mathbf{r} by delaying and summing the scattered wavefield $p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t)$ for all measurement directions $\boldsymbol{\theta}$ and incident-wave directions $\boldsymbol{\alpha}$. Thus, the time-domain reconstruction formula given by Eq. (18) can be regarded as a quantitative generalization of confocal time-domain synthetic aperture imaging, in which signals are synthetically delayed and summed for each transmit/receive pair to focus at the image point of interest.^{22,29,30}

A reconstruction formula similar to, although less general than, Eq. (18) was independently derived in Ref. 23 for the two-dimensional inverse scattering problem. In view of the present derivation, the method of ‘‘probing by plane pulses’’ in Ref. 23 can be regarded to yield a multiple-frequency reconstruction of $\text{Re}[\gamma_M(\mathbf{r})]$, while the present method yields the complex function $\gamma_M(\mathbf{r})$. In Ref. 23, this method was proposed as a more convenient way to implement narrow-band diffraction tomography. However, the numerical results given below show that the reconstruction formula of Eq. (18), when directly implemented using wideband signals, provides considerable improvement in image quality over narrow-band reconstructions.

Reconstructions using Eq. (18) can be performed using any pulse waveform. However, the frequency compounding defined by Eq. (10) is most straightforwardly interpreted if the frequency weight $\hat{g}(\omega)$ has a phase that is independent of frequency. This criterion can be met, for instance, if the incident pulse waveform $f(t)$ is even in time,

$$f(t) = f(-t), \quad (20)$$

so that $\hat{f}(\omega)$ is purely real. [Similarly, if the incident pulse waveform is odd in time, $\hat{f}(\omega)$ is purely imaginary and Eq. (18) can still be employed.]

However, supposition of a frequency-independent phase for $\hat{f}(\omega)$ does not result in any loss of generality. For any linear-phase signal, such that the Fourier transform has the form

$$\hat{f}(\omega) = |\hat{f}(\omega)| e^{i\omega\zeta}, \quad \omega > 0, \quad (21)$$

an additional delay term of magnitude ζ can be applied to all scattered signals to obtain the signals associated with the purely-real spectrum $|\hat{f}(\omega)|$. In general, the scattered field associated with a desired waveform $f(t)$ can be determined for an arbitrary waveform $u(t)$ from the deconvolution operation

$$[p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t)]_{f(t)} = \mathbf{F}^{-1} \left[\frac{\hat{f}(\omega)}{\hat{u}(\omega)} [p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t)]_{u(t)} \right]. \quad (22)$$

For stable deconvolution using Eq. (22), the desired $\hat{f}(\omega)$ should not have significant frequency components outside the bandwidth of $\hat{u}(\omega)$.

B. Extensions to the reconstruction algorithm

For large tissue structures at high ultrasonic frequencies, weak scattering approximations such as the Born approximation are of limited validity. Thus, for problems of interest to medical ultrasonic imaging, reconstructed image quality can be improved by aberration correction methods that incorporate higher-order scattering and propagation effects. The present time-domain reconstruction formula (18) provides a natural framework for quantitative imaging with aberration correction. In general, if the background medium is known or can be estimated, the received scattered signals can be processed to provide an estimate of the scattered field that would be measured for the same scatterer within a homogeneous background medium. This approach essentially removes higher-order scattering effects from the measured far field scattering, so that a Born inversion can be performed on the modified data; similar processes occur implicitly in many nonlinear inverse scattering methods.³¹

For example, a simple implementation of aberration correction can be derived if one makes the assumption that background inhomogeneities result only in cumulative delays (or advances) of the incident and scattered wavefronts. This crude model does not include many propagation and scattering effects important to ultrasonic aberration, but has been shown to provide a reasonable first approximation of local delays in wavefronts propagating through large-scale tissue models.^{32,33} Given this approximation, the total delay for an angle ϕ and a point position \mathbf{r} is given by

$$\delta\tau(\boldsymbol{\phi}, \mathbf{r}) = \int_{\xi} c(\boldsymbol{\xi})^{-1} d\xi - \frac{R}{c_0}, \quad (23)$$

where the integral is performed along the line that joins the spatial points \mathbf{r} and $R\boldsymbol{\phi}$. Aberration-corrected reconstructions can then be performed using Eq. (18) with τ replaced by the corrected delay term

$$\tau \rightarrow R/c_0 + \frac{(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{r}}{c_0} + \delta\tau(\boldsymbol{\alpha}, \mathbf{r}) + \delta\tau(\boldsymbol{\theta}, \mathbf{r}). \quad (24)$$

Improved approximations could be obtained by application of the delay function $\delta\tau(\boldsymbol{\phi}, \mathbf{r})$ after numerical backpropagation of the far-field scattered wavefronts through a homogeneous medium^{34,35} or by compensation for both delay and amplitude variations.^{36,37} More general, although much more computationally expensive, aberration correction could also be performed by synthetic focusing using full-wave numerical computation of acoustic fields within an estimated realization of the unknown medium. A method of this kind has been implemented, within the context of a frequency-domain diffraction tomography method, in Ref. 19.

The present imaging method has been derived using simplifying assumptions including zero absorption and constant density for the scattering medium. However, these assumptions do not substantially restrict the validity of the method. For example, the effect of absorption can be reduced using time-gain compensation, with or without frequency-dependent corrections,³⁸ of received scattered signals for each transmit/receive pair. Such time-gain compensation could be performed either using an estimated bulk attenuation for the medium (as with current clinical ultrasound scanners), or by implementation of an adaptive attenuation model in a manner similar to the time-shift compensation scheme discussed above.

Inclusion of density variations as well as sound speed variations adds additional complication to the time-domain diffraction tomography algorithm derived here. For single-frequency diffraction tomography in the presence of sound speed and density variations, the quantity $\gamma_B(\mathbf{r}, \omega)$ reconstructed by Eq. (8) can be shown³⁹ to provide an estimate of a physical quantity that depends both on sound speed variations and density variations. In the notation used here, this quantity can be written

$$\gamma'(\mathbf{r}) = \gamma(\mathbf{r}) - \gamma(\mathbf{r})\gamma_\rho(\mathbf{r}) + \frac{1}{2k^2} \nabla^2 \gamma_\rho(\mathbf{r}), \quad (25)$$

where the density variation is defined $\gamma_\rho = 1 - \rho_0/\rho(\mathbf{r})$. Thus, for time-domain reconstructions of media with density variations, the reconstruction formula of Eq. (18) will provide the estimate

$$\gamma_M(\mathbf{r}) \approx \gamma(\mathbf{r}) - \gamma(\mathbf{r})\gamma_\rho(\mathbf{r}) + \frac{1}{2k_0^2} \nabla^2 \gamma_\rho(\mathbf{r}), \quad (26)$$

where k_0 is the wave number corresponding to the center frequency of the pulse employed. For media such as human tissue, where density variations are fairly small and abrupt density transitions are rare, the last two terms of Eq. (26) are small compared to $\gamma(\mathbf{r})$, so that the reconstruction algorithm derived above can still be regarded to provide an image of the sound-speed variation function $\gamma(\mathbf{r})$. However, if desired, a reconstruction employing pulses with two distinct center frequencies could allow separation of sound speed and density variations by techniques similar to those described in Ref. 16 or 39.

II. COMPUTATIONAL METHODS

The time-domain inverse scattering method described above has been tested with 2D and 3D synthetic data prepared using three numerical methods: a Born approximation method for point scatterers and 3D slabs, an exact series solution for cylindrical inhomogeneities, and a k -space method for arbitrary 2D inhomogeneous media.

The time-domain waveform employed for all the computations reported here was

$$f(t) = \cos(\omega_0 t) e^{-t^2/(2\sigma^2)}, \quad (27)$$

where $\omega_0 = 2\pi f_0$ for a center frequency of f_0 and σ is the temporal Gaussian parameter. This waveform has the real, even Fourier transform

$$\hat{f}(\omega) = \sqrt{\frac{\sigma^2}{8\pi}} (e^{-\sigma^2(\omega - \omega_0)^2/2} + e^{-\sigma^2(\omega + \omega_0)^2/2}). \quad (28)$$

Values used for the computations reported here were $f_0 = 2.5$ MHz and $\sigma = 0.25$ μ s, so that the -6 dB bandwidth of the signal was 1.5 MHz. These parameters correspond closely to those of an existing 2048-element ring transducer.⁴⁰

For the case of point scatterers, the contrast function γ was assumed to take the form

$$\gamma(\mathbf{r}) = \sum_1^M \mu_j \delta(\mathbf{r} - \mathbf{r}_j). \quad (29)$$

Using the far-field form of the 2D Green's function and neglecting multiple scattering, Eq. (6) for the scattered far field can be rewritten as

$$\hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega) = -k^2 \sqrt{\frac{i}{8\pi k R}} \hat{f}(\omega) \sum_j \mu_j e^{ik(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{r}_j} \quad (30)$$

for each frequency component of interest. Time-domain waveforms were synthesized by using Eq. (30) for each frequency with $\hat{f}(\omega) > 10^{-3}$ and inverting the frequency-domain scattered wavefield by a fast Fourier transform (FFT) implementation of Eq. (4). The temporal sampling rate employed was 10 MHz. An analogous formula, with a different multiplicative constant, was also employed for the 3D case.

The Born approximation was also used to compute three-dimensional scattering for slab-shaped objects defined by the equation

$$\gamma(\mathbf{r}) = \gamma_0 H(a_x - |x|) H(a_y - |y|) H(a_z - |z|). \quad (31)$$

For this object, the linearized forward problem can be solved analytically. Under the Born approximation, the frequency-domain scattered far field has the form

$$\begin{aligned} \hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega) &= 2\hat{f}(\omega) \gamma_0 a_x a_y a_z e^{ikR}/(\pi R) \\ &\times \frac{\sin[kL_x(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{e}_x]}{kL_x(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{e}_x} \frac{\sin[kL_y(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{e}_y]}{kL_y(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{e}_y} \\ &\times \frac{\sin[kL_z(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{e}_z]}{kL_z(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{e}_z}, \end{aligned} \quad (32)$$

where \mathbf{e}_x , \mathbf{e}_y , and \mathbf{e}_z represent unit vectors in the x , y , and z directions. The time domain scattered pressure $p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, t)$ is

obtained, as for the point scatterer case described above, by inverse transformation of the frequency-domain wavefield for all frequencies within the bandwidth of interest.

For 2D cylindrical inhomogeneities, an analogous procedure was followed, except that the frequency-domain scattered wavefield $\hat{p}_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \omega)$ was computed using an exact series solution²⁵ for each frequency component of interest. In implementation of the series solution, summations were truncated when the magnitude of a single coefficient dropped below 10^{-12} times the sum of all coefficients.

Solutions were also obtained for arbitrary 2D inhomogeneous media using a time-domain k -space method.⁴¹ Grid sizes of 256×256 points, a spatial step of 0.0833 mm, and a time step of $0.02734 \mu\text{s}$ were employed. Scattered acoustic pressure signals on a circle of virtual receivers were recorded at a sampling rate of 9.144 MHz. The receiver circle, which had a radius of 3.0 mm in these computations, completely contained the inhomogeneities used. Far-field waveforms were computed by Fourier transforming the time-domain waveforms on the near-field measurement circle, transforming these to far-field waveforms for each frequency using a numerically exact transformation method,¹⁶ and performing inverse Fourier transformation to yield time-domain far-field waveforms. All forward and inverse temporal Fourier transforms, as well as angular transforms occurring in the near-field-far-field transformation,¹⁶ were performed by FFT.

The time-domain imaging method was directly implemented using Eq. (18), evaluated using straightforward numerical integration over all incident-wave and measurement directions employed. The reconstruction formula employed can be explicitly written as

$$\begin{aligned} \gamma_M(\mathbf{r}) = & \frac{1}{N_{2D}} \int_0^{2\pi} \int_0^{2\pi} |\sin(\alpha - \theta)| \left(p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \tau) \right. \\ & \left. + i\mathbf{H}^{-1}[p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \tau)] \right) d\alpha d\theta, \\ \tau = R/c_0 + & \frac{(\cos \alpha - \cos \theta) \cdot x + (\sin \alpha - \sin \theta) \cdot y}{c_0} \end{aligned} \quad (33)$$

for the 2D case, where α and θ are the angles corresponding to the direction vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$, and as

$$\begin{aligned} \gamma_M(\mathbf{r}) = & \frac{1}{N_{3D}} \int_0^{2\pi} \int_0^\pi \int_0^{2\pi} \int_0^\pi |\boldsymbol{\alpha} - \boldsymbol{\theta}| \left(p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \tau) \right. \\ & \left. + i\mathbf{H}^{-1}[p_s(\boldsymbol{\theta}, \boldsymbol{\alpha}, \tau)] \right) \sin(\Phi_\alpha) \sin(\Phi_\theta) d\Phi_\alpha \\ & \times d\Theta_\alpha d\Phi_\theta d\Theta_\theta, \\ \tau = R/c_0 + & \frac{(\boldsymbol{\alpha} - \boldsymbol{\theta}) \cdot \mathbf{r}}{c_0}, \\ \boldsymbol{\alpha} - \boldsymbol{\theta} = & (\cos \Theta_\alpha \sin \Phi_\alpha - \cos \Theta_\theta \sin \Phi_\theta) \cdot \mathbf{e}_x \\ & + (\sin \Theta_\alpha \sin \Phi_\alpha - \sin \Theta_\theta \sin \Phi_\theta) \cdot \mathbf{e}_y \\ & + (\cos \Phi_\alpha - \cos \Phi_\theta) \cdot \mathbf{e}_z \end{aligned} \quad (34)$$

for the 3D case, where Θ_α and Φ_α are direction angles for the incident-wave direction $\boldsymbol{\alpha}$ and Θ_θ and Φ_θ are direction

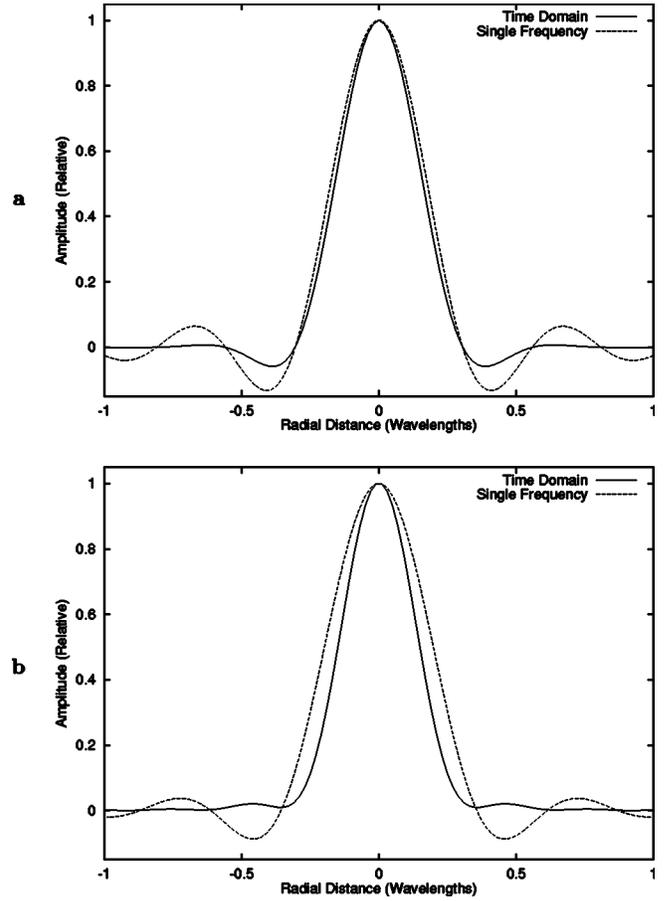


FIG. 2. Point-spread functions for time-domain and single-frequency diffraction tomography methods. In each panel, the vertical scale corresponds to the relative amplitude of the reconstructed contrast $\gamma(\mathbf{r})$, while the horizontal scale corresponds to number of wavelengths at the center frequency. (a) Two-dimensional case. (b) Three-dimensional case.

angles for the measurement direction $\boldsymbol{\theta}$. For each case, the normalization factor N was determined from Eq. (11) with $\hat{g}(\omega) = \hat{f}(\omega) / \hat{\mu}(\omega)$ and $\mu(\omega)$ given by Eq. (9). Before evaluation of the argument τ for each signal, the time-domain waveforms were resampled at a sampling rate of 16 times the original rate. This resampling was performed using FFT-based Fourier interpolation. The inverse Hilbert transform was performed for each signal using an FFT implementation of Eq. (16). Values of the pressure signals at the time τ were then determined using linear interpolation between samples of the resampled waveforms. The integrals of Eqs. (33) and (34) were implemented using discrete summation over all transmission and measurement directions employed.

Computations were also performed using the time-domain diffraction tomography algorithm for limited-aperture data. For these reconstructions, the integrals of Eq. (33) were evaluated only for angles corresponding to transmitters and receivers within a specified aperture of angular width ϕ_{ap} , i.e.,

$$|\alpha| \leq \phi_{\text{ap}}/2, \quad |\theta - \pi| \leq \phi_{\text{ap}}/2. \quad (35)$$

Use of a small value for ϕ_{ap} corresponds to use of a small aperture in pulse-echo mode.

III. NUMERICAL RESULTS

Two-dimensional and three-dimensional point-spread functions (PSF) for the present time-domain diffraction tomography method are illustrated in Fig. 2. The time-domain reconstructions shown here, like the other time-domain reconstructions shown in this paper, were obtained using an incident pulse of center frequency 2.5 MHz and a Gaussian envelope corresponding to a -6 dB bandwidth of 1.5 MHz. Point-spread functions were determined by reconstructing a point scatterer located at the origin. For the 2D case, in which the point scatterer can be regarded as a thin wire, synthetic scattering data was obtained using the Born approximation method outlined above for 16 incident-wave directions and 64 measurement directions. The 3D time-domain reconstruction was obtained using Born data for 72 incident-wave directions and 288 measurement directions, each evenly spaced on a rectangular grid defined by the angles Θ and Φ . For comparison, analogous point-spread functions are also shown for standard frequency-domain diffraction tomography reconstructions using single-frequency (2.5 MHz) data.

For the 2D case illustrated in Fig. 2, the time-domain PSF has a slightly narrower peak, indicating that point resolution has been slightly improved by the increased bandwidth employed in the time domain method. More significantly, sidelobes of the time-domain PSF are significantly smaller than those for the single-frequency PSF (the first sidelobe is reduced by 7 dB, while the second is reduced by 19 dB), so that contrast resolution for time-domain diffraction tomography is seen to be much higher than for single-frequency diffraction tomography. For the 3D case, the time-domain reconstruction shows a much more dramatic improvement over the single-frequency reconstruction. In this case, the time-domain solution shows significant increases in both the point resolution (PSF width at half-maximum reduced by 27%) and contrast resolution (first sidelobe reduced by 13 dB and second sidelobe reduced by 18 dB). Furthermore, a comparison of the PSFs for 2D and 3D time-domain reconstruction indicates that much higher image quality is achievable for 3D time-domain imaging than for the 2D case. This increase in image quality suggests that the time-domain diffraction tomography method proposed here may benefit from the overdetermined nature of the general wideband 3D inverse scattering problem.^{42,43}

The effect of transmit and receive aperture characteristics on image quality is illustrated in Fig. 3. Panels (a) and (b) of Fig. 3 show the point-spread function for a number of aperture configurations, each employing 64 measurement directions. Figure 3(a) shows the point-spread function for reconstructions obtained using 1, 4, 8, and 16 incident-wave directions. The point scatterer is clearly imaged even for the reconstruction using one incident-wave direction. Optimal image quality (indistinguishable from reconstructions with 64 incident-wave directions) is obtained for 16 incident-wave directions, so that scattering data obtained using one incident-wave direction for each group of four measurement directions appears to be sufficient for the present reconstruction method.

The effect of limited view range on the point spread

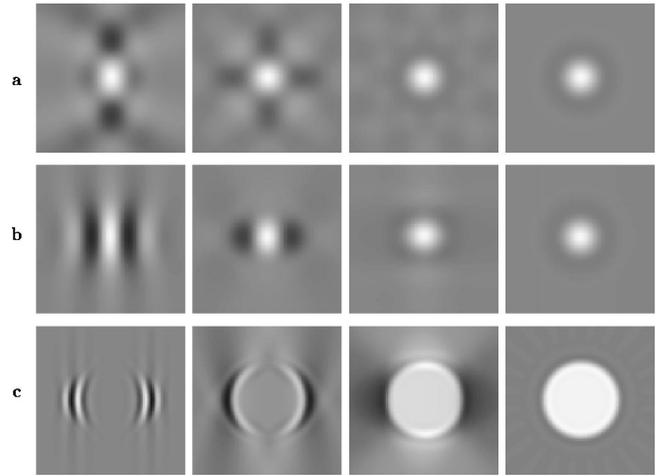


FIG. 3. Effect of aperture characteristics on image quality. Each panel shows the real part of a time-domain reconstruction, $\text{Re}[\gamma_M]$, on a linear grayscale with white representing the maximum amplitude of $|\gamma_M(\mathbf{r})|$ and black represents -1 times the maximum amplitude. (a) Point-spread functions for the same waveform parameters as Fig. 2. Each panel shows an area of 0.6×0.6 mm², corresponding to one square wavelength at the center frequency. Left to right: 1, 4, 8, and 16 incident-wave directions. (b) Point-spread functions for aperture sizes of $\pi/2$, π , $3\pi/2$, and 2π radians, format as in previous panel. (c) Real parts of reconstructions for a homogeneous cylinder ($a = 1.0$ mm, $\gamma = 0.02$). The area shown in each panel is 2.0×2.0 mm². Left to right: aperture sizes of $\pi/2$, π , $3\pi/2$, and 2π radians.

function is also illustrated in Fig. 3. Panel (b) shows the point-spread function for four differently limited apertures, while panel (c) shows reconstructions of a homogeneous cylinder ($a = 1.0$ mm, $\gamma = 0.02$) for the same apertures. In each case, limitation of the transmit and receive apertures to angles near the backscatter direction (aperture size $\pi/2$) results in images that resemble a conventional B-scans. Use of apertures corresponding to pulse-echo mode in the large-aperture limit (aperture size π) yield higher resolution in all directions. Using three-fourths of a circular aperture (size $3\pi/2$) yields image quality close to that for the full aperture (2π) case. The characteristics of all these images result from the set of spatial-frequency vectors interrogated by each group of scattering measurements.¹ Apertures with only a limited range of transmit and receive directions [e.g., the “b-scan” apertures shown in the first column of panels (b) and (c)] provide only information corresponding to large spatial frequency vectors oriented nearly on-axis, so that such images mainly show those edges that are nearly perpendicular to the axis of the aperture.

Reconstructions performed using exact solutions for scattering from cylindrical inhomogeneities provide a straightforward means to assess the accuracy of the time-domain scattering method for a range of object sizes and contrasts. A number of example reconstructions are shown in Figs. 4 and 5. The number of measurement directions for all cylinder reconstructions was chosen based on an empirical test of the number required for a satisfactory image of a homogeneous cylinder; for a cylinder of radius 1 mm, the required number of measurement directions was determined to be approximately 96. Based on spatial-frequency sampling considerations, the number of measurement directions was increased in proportion to the size of the inhomogeneous

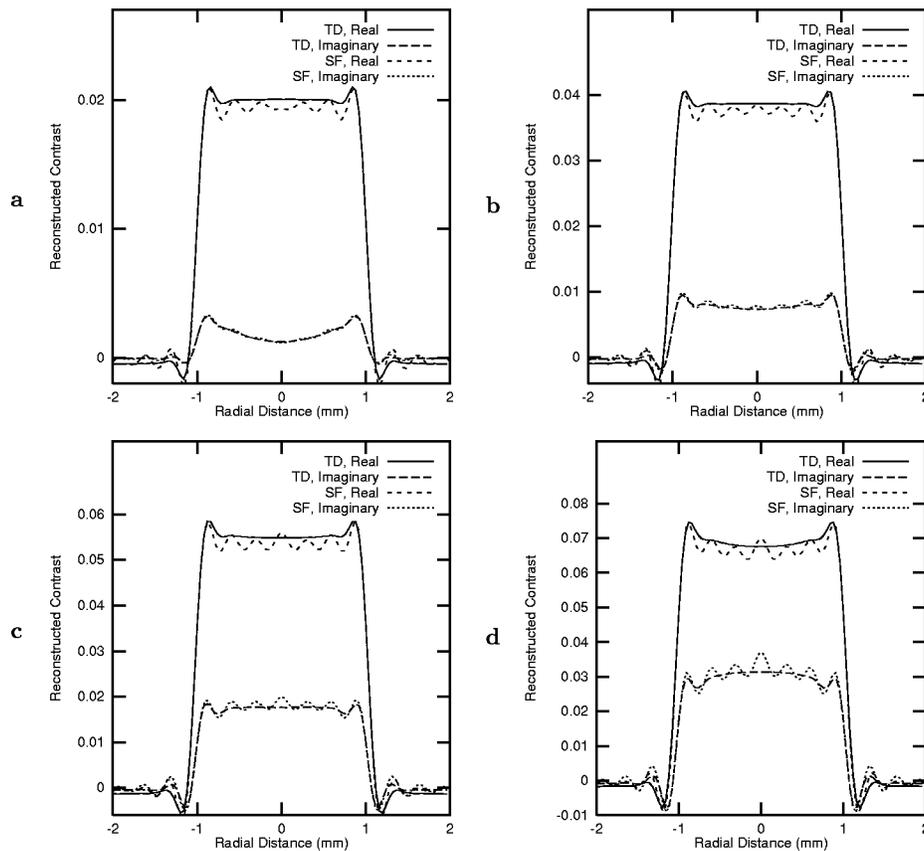


FIG. 4. Cross sections of reconstructed contrast functions $\gamma(\mathbf{r})$ for a cylinder of radius 1 mm, using time-domain (TD) and single-frequency (SF) diffraction tomography. Waveform parameters are as in Fig. 1. (a) $\gamma=0.02$. (b) $\gamma=0.04$. (c) $\gamma=0.06$. (d) $\gamma=0.08$.

region to be reconstructed. Since the results shown in Fig. 3 indicate that considerably fewer incident-wave directions than measurement directions are needed, the number of incident directions was chosen to be one-quarter the number of measurement directions in each case.

Cross sections of time-domain and single-frequency reconstructions, plotted in Fig. 4, show the relative accuracy of each reconstruction method for a cylinder of 1-mm radius and purely real contrast ranging from $\gamma=0.02$ to $\gamma=0.08$. For the synthetic scattering data in each case, 96 measurement directions and 24 incident-wave directions were employed. The time-domain reconstructions show improvement over the single-frequency reconstructions both in improved contrast resolution (smaller sidelobes outside the support of the cylinder) and in decreased ringing (Gibbs phenomenon) artifacts within the support of the cylinder. However, for increasing contrast values, both methods show similar increases in phase error, as indicated by increased imaginary parts of the reconstructed contrast. This error results from the Born approximation, which is based on the assumption that the incident wave propagates through the inhomogeneous medium without distortion. Perturbations in the local arrival time of the incident wavefront, which are more severe for higher contrasts and larger inhomogeneities, can result in a scattered field that is phase shifted relative to the ideal case assumed in the Born approximation; linear inversion of this phase-distorted data naturally results in a phase-distorted reconstruction of the scattering medium. (A complementary

explanation of this phase error, based on the unitarity of the scattering operator, is given in Ref. 19.)

A test of image fidelity for the time-domain reconstruction method is shown in Fig. 5. The real parts of time-domain reconstructions are shown as grayscale images for homogeneous cylinders with radii between 1 and 4 mm and contrasts between $\gamma=0.02$ and $\gamma=0.08$. The number of measurement directions employed for the synthetic scattering data was 96 for the 1-mm radius cylinders, 192 for the 2-mm cylinders, 288 for the 3-mm cylinders, and 384 for the 4-mm cylinders. In each case, four incident-wave directions per measurement direction were used. The first row of this figure corresponds to the time-domain reconstructions shown in Fig. 4.

The images shown in Fig. 5 provide a basis for evaluating the ability of the present time-domain diffraction tomography method to image homogeneous objects of various sizes and contrasts. In this figure, images of $\text{Re}[\gamma_M]$ show uniform quality for small cylinder sizes and contrasts, but poorer image quality for larger sizes and contrasts. For the largest size and contrast employed ($a=4.0$ mm, $\gamma=0.08$), the reconstruction primarily shows the edges of the cylinder and fails to image the interior. Particularly notable is that the “matrix” of images in Fig. 5 is nearly diagonal; that is, a linear increase in object contrast causes image degradation comparable to a corresponding linear increase in object size. Thus, a nondimensional parameter directly relevant to image quality for homogeneous objects is $ka\gamma$, where k is a domi-

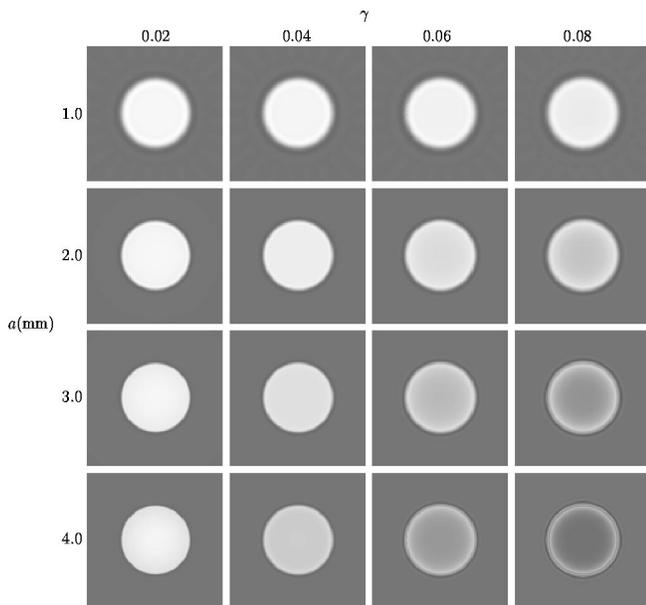


FIG. 5. Images of time-domain reconstructions for cylinders of varying radius a and contrast γ . Each panel shows the real part of the reconstructed contrast, $\text{Re}[\gamma_M(\mathbf{r})]$, for a pulse of center frequency 2.5 MHz and -6 dB bandwidth 1.5 MHz. The area shown in each panel is $2a \times 2a$. All images are shown on a linear, bipolar gray scale where white represents the maximum amplitude of $|\gamma_M(\mathbf{r})|$ and black represents -1 times the maximum amplitude.

nant wave number, a is the object radius, and γ is the object contrast. Using the wave number $k_0 = 10.472$ rad/mm corresponding to the center frequency of 2.5 MHz and a sound speed of 1.5 mm/ μ s, the reconstructions shown in Fig. 5 indicate that the interior of the cylinder is imaged satisfactorily for the approximate range $ka \gamma < 2.5$. This result is consistent with a previous study of single-frequency diffraction tomography, in which adequate Born reconstructions of cylinders were obtained for the parameter range $ka \gamma \leq 2.2$.⁴⁴

Reconstructions for several scattering objects without special symmetry are shown in Fig. 6. All of these reconstructions were performed using synthetic data produced by the k -space method described in Ref. 41. Synthetic scattering data were computed for 64 incident-wave directions and 256 measurement directions in each case. The first panel shows a reconstruction of a cylinder of radius 2.5 mm and contrast $\gamma = -0.0295$ with an internal cylinder of radius 0.2 mm and contrast $\gamma = 0.0632$. These contrast values correspond, based on tissue parameters given in Ref. 32, to the sound-speed contrasts of human skeletal muscle for the outer cylinder and of human fat for the inner cylinder. The second panel shows a reconstruction of a 2.5-mm-radius cylinder with random internal structure. The third reconstruction shown employed a portion of a chest wall tissue map from Ref. 45. In this case, the synthetic data was obtained using a tissue model⁴⁵ that incorporates both sound speed and density variations, so that the reconstructed quantity is given by Eq. (26). In Fig. 6(c), black denotes connective tissue ($\gamma = -0.1073$, $\gamma_\rho = 0.1134$), dark gray denotes muscle ($\gamma = -0.0295$, $\gamma_\rho = 0.0543$), and light gray denotes fat ($\gamma = 0.0632$, $\gamma_\rho = -0.0453$).

The real part of each reconstruction in Fig. 6 shows

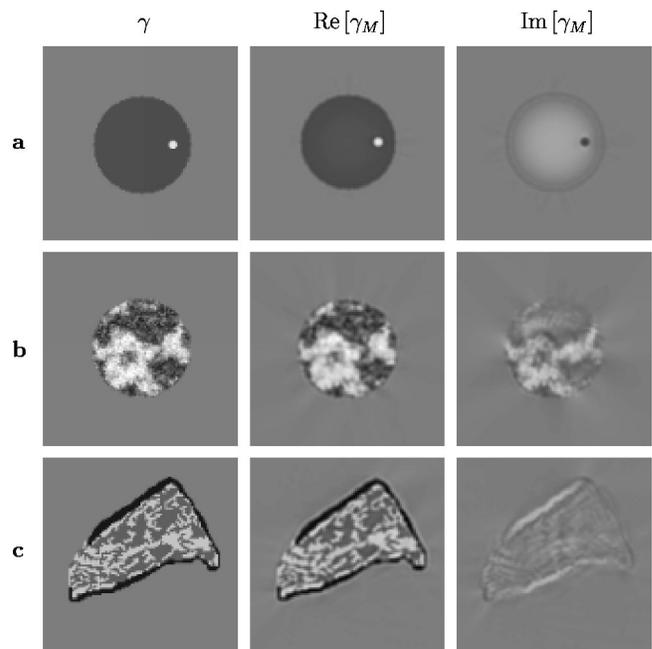


FIG. 6. Time-domain reconstructions from full-wave synthetic data for three arbitrary scattering objects. Each row shows the actual (purely real) contrast function γ together with the real and imaginary parts of the reconstructed contrast function γ_M , using the same linear bipolar gray scale for each panel. Each panel shows a reconstruction area of 5×5 mm². (a) Cylinder, radius 2.5 mm, with an internal cylinder of radius 0.2 mm. (b) Cylinder, radius 2.5 mm, with random internal structure. (c) Tissue structure, with variable sound speed and density, from chest wall cross section 5L in Ref. 45.

good image quality, with high resolution and very little evidence of artifacts. Particularly notable is the accurately detailed imaging of internal structure for the random cylinder and the chest wall cross section. As expected, the density variations present in the chest wall cross section have not greatly affected the image appearance; there is, however, a slight edge enhancement, associated with the Laplacian term in Eq. (26), at boundaries between tissue regions. Also notable is the nearly complete absence of any artifacts outside the scatterer in each case; this result indicates that high contrast resolution has been achieved. However, in each case, the imaginary part of the reconstruction is nonzero, indicating that the Born approximation is not fully applicable. The imaginary parts of each reconstruction are, however, small compared to the real parts. Thus, simple aberration correction methods [of which one example is given by Eq. (24)] could substantially reduce this phase error, as for multiple-frequency diffraction tomography in Ref. 19.

Three-dimensional reconstructions of a homogeneous slab are shown in Fig. 7. The scatterer is characterized by Eq. (31) with $\gamma_0 = 0.01$, $a_x = 0.5$ mm, $a_y = 1.0$ mm, and $a_z = 1.5$ mm. Synthetic data was computed using Eq. (34) for 288 incident-wave directions and 1152 measurement directions, each evenly spaced in the angles Φ and Θ . Signal parameters were as for the examples above, except that the initial sampling rate for the time-domain signals was 9.0 MHz. Iso-surface renderings of the real part of γ_M are shown for the surfaces $\gamma_M = 0.0025$. Since the scattering data were obtained using a Born approximation for the 3D case, the

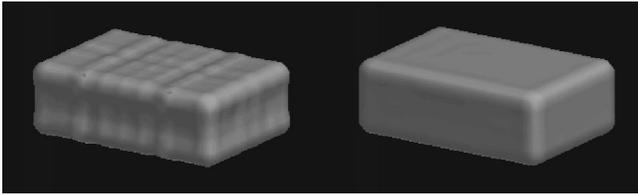


FIG. 7. Three-dimensional reconstructions of a uniform slab with contrast $\gamma=0.01$. Each reconstruction shows an isosurface rendering of the surface $\gamma_M=0.0025$. Left: single-frequency reconstruction. Right: time-domain reconstruction.

imaginary part of each reconstruction is identically zero for both reconstructions. Consistent with the point-spread functions shown in Fig. 2, the time-domain reconstruction is much more accurate than the single-frequency reconstruction. While the single-frequency reconstruction shows an erroneously rippled surface, the time-domain reconstruction is smooth. The time-domain reconstruction is nearly identical to the original object except for some rounding of the sharp edges due to the limited high-frequency content of the signal employed. The length scale of the rounded edges is on the order of one-half the wavelength of the highest frequency in the pulse, i.e., about 0.2 mm for the -6 -dB cutoff of 3.25 MHz.

Since three-dimensional inverse scattering is a computationally demanding problem, comparison of computational efficiency for single-frequency and time-domain methods is of interest. For both reconstructions shown in Fig. 7, identical discretizations of the reconstructed medium were employed. Both computations included solution of the applicable linearized forward problem as well as the inverse problem. Nonetheless, the time-domain method was more efficient than the single-frequency method; the total CPU time required on a 200-MHz AMD K6 processor was 133.3 CPU min for the time-domain method and 287.4 CPU min for the single-frequency method. This gain in efficiency was possible because the greatest computational expense occurred in the “backpropagation” of the signals for each reconstruction point. For the single-frequency method, this step required evaluation of complex exponentials for each incident-wave direction, measurement direction, and spatial point. For the time-domain method, however, the computationally intensive steps (including the forward problem solution and Fourier interpolation of the scattered signals) needed only to be performed once for each transmit/receive pair. For the backpropagation step, performed at each point in the 3D spatial grid, the time-domain reconstruction method required only linear interpolation of the oversampled farfield pressure waveforms.

IV. CONCLUSIONS

A new method for time-domain ultrasound diffraction tomography has been presented. The method provides quantitative images of sound speed variations in unknown media; when two pulse center frequencies are employed, the method is also capable of imaging density variations. Reconstructions performed using this method are equivalent to multiple-

frequency reconstructions using filtered backpropagation, but can be obtained with much greater efficiency.

The time-domain reconstruction algorithm has been derived as a simple filtered delay-and-sum operation applied to far-field scattered signals. This algorithm is closely related to time-domain confocal synthetic aperture imaging, so that it can be considered a generalization of imaging algorithms employed in current clinical instruments. The simplicity of the imaging algorithm allows straightforward addition of features such as time-gain compensation and aberration correction.

Numerical results obtained using synthetic data for 2D and 3D scattering objects show that the time-domain method can yield significantly higher image quality (and, in some cases, also greater efficiency) than single-frequency diffraction tomography. Quantitative reconstructions, obtained using signal parameters comparable to those for present-day clinical instruments, show accurate imaging of objects with simple deterministic structure, random internal structure, and structure based on a cross-sectional tissue model. The method is hoped to be useful for diagnostic imaging problems such as the detection and characterization of lesions in ultrasonic mammography.

ACKNOWLEDGMENTS

This research was funded by the Breast Cancer Research Program of the U.S. Army Medical Research and Materiel Command, under Grant No. DAMD17-98-1-8141. The author is grateful for helpful discussions with Adrian I. Nachman, Feng Lin, and Robert C. Waag.

- ¹E. Wolf, “Three-dimensional structure determination of semi-transparent objects from holographic data,” *Opt. Commun.* **1**, 153–156 (1969).
- ²A. J. Devaney, “A filtered backpropagation algorithm for diffraction tomography,” *Ultrason. Imaging* **4**, 336–350 (1982).
- ³D. T. Borup, S. A. Johnson, W. W. Kim, and M. J. Berggren, “Nonperturbative diffraction tomography via Gauss-Newton iteration applied to the scattering integral equation,” *Ultrason. Imaging* **14**, 69–85 (1992).
- ⁴S. Gutman and M. Klibanov, “Two versions of quasi-Newton method for multidimensional inverse scattering problem,” *J. Comput. Acoust.* **1**, 197–228 (1993).
- ⁵F. W. Kremkau, *Diagnostic Ultrasound: Physical Principles and Exercises* (Grune and Stratton, New York, 1980).
- ⁶J. L. Schwartz and B. D. Steinberg, “Ultrasparse, ultrawideband arrays,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 376–393 (1998).
- ⁷S. H. Maslak, “Computed sonography,” in *Ultrasound Annual 1985* (Raven, New York, 1985), pp. 1–16.
- ⁸E. L. Madsen, B. S. Garra, J. A. Parks, A. C. Skelly, and J. A. Zagzebski, *AIUM Quality Control Manual for Gray-Scale Ultrasound Scanners* (American Institute of Ultrasound in Medicine, Laurel, MD, 1995).
- ⁹S. J. Norton, “Reconstruction of a two-dimensional reflecting medium over a circular domain: exact solution,” *J. Acoust. Soc. Am.* **67**, 1266–1273 (1980).
- ¹⁰S. J. Norton and M. Linzer, “Ultrasonic reflectivity imaging in three dimensions: exact inverse scattering solutions for plane, cylindrical, and circular apertures,” *IEEE Trans. Biomed. Eng.* **28**, 202–220 (1980).
- ¹¹J. H. Kim, S. B. Park, and S. A. Johnson, “Tomographic imaging of ultrasonic reflectivity with correction for acoustic speed variations,” *Ultrason. Imaging* **6**, 304–312 (1984).
- ¹²C. Q. Lan and W. Xiong, “An iterative method of ultrasonic reflection mode tomography,” *IEEE Trans. Med. Imaging* **13**, 419–425 (1994).
- ¹³D. Miller, M. Oristaglio, and G. Beylkin, “A new slant on seismic imaging: migration and internal geometry,” *Geophysics* **52**, 943–964 (1987).

- ¹⁴T. Melamed, Y. Ehrlich, and E. Heymann, "Short-pulse inversion of inhomogeneous media: a time-domain diffraction tomography," *Inverse Probl.* **12**, 977–993 (1996).
- ¹⁵S. Pourjavid and O. Tretiak, "Ultrasound imaging through time-domain diffraction tomography," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **38**, 74–85 (1991).
- ¹⁶A. Witten, J. Tuggle, and R. C. Waag, "A practical approach to ultrasonic imaging using diffraction tomography," *J. Acoust. Soc. Am.* **83**, 1645–1652 (1988).
- ¹⁷T. J. Cavicchi and W. D. O'Brien, "Numerical study of higher-order diffraction tomography via the sinc basis moment method," *Ultrason. Imaging* **11**, 42–74 (1989).
- ¹⁸H. Gan, R. Ludwig, and P. L. Levin, "Nonlinear diffractive inverse scattering for multiple scattering in inhomogeneous acoustic background media," *J. Acoust. Soc. Am.* **97**, 764–776 (1995).
- ¹⁹T. D. Mast, A. I. Nachman, and R. C. Waag, "Focusing and imaging using eigenfunctions of the scattering operator," *J. Acoust. Soc. Am.* **102**, 715–725 (1997).
- ²⁰J. M. Blackledge, R. E. Burge, K. I. Hopcraft, and R. J. Wombell, "Quantitative diffraction tomography: I. Pulsed acoustic fields," *J. Phys. D* **20**, 1–10 (1987).
- ²¹S. Mensah and J.-P. Lefebvre, "Enhanced compressibility tomography," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 1245–1252 (1997).
- ²²G. Prokoph and H. Ermert, "A comparison of broadband holographic and tomographic imaging concepts," *Acoust. Imaging* **18**, 381–390 (1991).
- ²³V. A. Burov and O. D. Romyantseva, "Linearized inverse problem of scattering in monochromatic and pulse modes," *Acoust. Phys.* **40**, 34–42 (1996).
- ²⁴F. Lin, A. I. Nachman, and R. C. Waag, "Quantitative imaging using a time-domain eigenfunction method," submitted to *J. Acoust. Soc. Am.* (1999).
- ²⁵P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (New York, McGraw-Hill, 1968), Chap. 8.
- ²⁶P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. I.
- ²⁷G. Beylkin, "The fundamental identity for iterated spherical means and the inversion formula for diffraction tomography and inverse scattering," *J. Math. Phys.* **24**, 1399–1400 (1982).
- ²⁸A. Papoulis, *The Fourier Integral and Its Applications* (McGraw-Hill, New York, 1962), Chap. 10.
- ²⁹J. Ylitalo, E. Alasaarela, and J. Koivukangas, "Ultrasound holographic B-scan imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **36**, 376–383 (1989).
- ³⁰K. E. Thomenius, "Evolution of ultrasound beamformers," *Proc. IEEE Ultrason. Symp.* **2**, 1615–1622 (1996).
- ³¹R. Snieder, "A perturbative analysis of non-linear inversion," *Geophys. J. Int.* **101**, 545–556 (1990).
- ³²T. D. Mast, L. M. Hinkelman, M. J. Orr, V. W. Sparrow, and R. C. Waag, "Simulation of ultrasonic pulse propagation through the abdominal wall," *J. Acoust. Soc. Am.* **102**, 1177–1190 (1997) [Erratum: **104**, 1124–1125 (1998)].
- ³³T. D. Mast, L. M. Hinkelman, M. J. Orr, and R. C. Waag, "The effect of abdominal wall morphology on ultrasonic pulse distortion. Part II. Simulations," *J. Acoust. Soc. Am.* **104**, 3651–3664 (1998).
- ³⁴D.-L. Liu and R. C. Waag, "Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation," *J. Acoust. Soc. Am.* **96**, 649–660 (1994).
- ³⁵C. Dorme and M. Fink, "Ultrasonic beam steering through inhomogeneous layers with a time reversal mirror," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**(1), 167–175 (1996).
- ³⁶S. Kirshnan, P.-C. Li, and M. O'Donnell, "Adaptive compensation for phase and magnitude aberrations," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**(1), 44–55 (1996).
- ³⁷Q. Zhu and B. D. Steinberg, "Deabberation of incoherent wavefront distortion: an approach toward inverse filtering," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 575–589 (1997).
- ³⁸I. Claesson and G. Salomonsson, "Frequency- and depth-dependent compensation of ultrasonic signals," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **35**, 582–592 (1988).
- ³⁹A. J. Devaney, "Variable density acoustics tomography," *J. Acoust. Soc. Am.* **78**, 120–130 (1985).
- ⁴⁰T. T. Jansson, T. D. Mast, and R. C. Waag, "Measurements of differential scattering cross section using a ring transducer," *J. Acoust. Soc. Am.* **103**, 3169–3179 (1998).
- ⁴¹L. P. Souriau, T. D. Mast, D.-L. D. Liu, M. Tabei, A. I. Nachman, and R. C. Waag, "A k -space method for large-scale models of wave propagation in tissue," submitted to *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* (1999).
- ⁴²A. Nachman, "Reconstructions from boundary measurements," *Ann. Math.* **128**, 531–576 (1988).
- ⁴³D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd ed. (Springer-Verlag, Berlin, 1998), Chap. 10.
- ⁴⁴M. Slaney, A. C. Kak, and L. E. Larsen, "Limitations of imaging with first-order diffraction tomography," *IEEE Trans. Microwave Theory Tech.* **32**, 860–874 (1984).
- ⁴⁵T. D. Mast, L. M. Hinkelman, M. J. Orr, and R. C. Waag, "Simulation of ultrasonic pulse propagation, distortion, and attenuation in the human chest wall," submitted to *J. Acoust. Soc. Am.* **106**, 3665–3677 (1999).

A Galerkin method for the numerical analysis of diffraction by a rectangular screen

Nathaniel J. de Lautour

Physics Department, University of Auckland, Private Bag 92019, Auckland, New Zealand

(Received 19 February 1999; revised 27 July 1999; accepted 13 August 1999)

A novel numerical analysis of acoustic diffraction by a rectangular screen is described. The boundary integral model of the system is solved by a Galerkin method using as a basis the scaling functions of discrete wavelet theory. The use of scaling functions enables the quadruple integrals in the Galerkin method to be analytically reduced to double integrals, and the singular and hypersingular integrals can be found from a recurrence formula. Numerical tests show that the new method is more efficient than a boundary element method based on collocation, particularly when the screen is irradiated near grazing incidence. © 1999 Acoustical Society of America.

[S0001-4966(99)01912-8]

PACS numbers: 43.20.Fn, 02.70.Pt, 02.60.Nm [ANN]

INTRODUCTION

The formulation of boundary value problems as boundary integral equations (BIEs) is commonly used for the solution of radiation and scattering problems. There are many advantages of the boundary integral formulation: the dimension of the problem is reduced by one; exterior problems with infinite domains can be treated without artificial boundaries; and the solutions automatically satisfy the radiation condition. In addition, the stiffness matrices derived from BIEs are generally well conditioned. Of interest in this paper is the BIE obtained from the normal derivative of the Helmholtz integral formula,

$$\frac{\partial u^{\text{inc}}(\mathbf{x}')}{\partial n(\mathbf{x}')} = \lim_{\mathbf{x}'' \rightarrow \mathbf{x}'} \int_{\Gamma} \Delta u(\mathbf{x}) \frac{\partial^2 G(\mathbf{x}, \mathbf{x}'')}{\partial n(\mathbf{x}) \partial n(\mathbf{x}'')} dS(\mathbf{x}), \quad (1)$$

which is used to model acoustic diffraction by thin, rigid, objects.¹⁻¹¹ BIEs such as (1) are typically solved by the boundary element method (BEM). In the BEM, Δu is discretized by finite-element functions, and the data, $\partial u^{\text{inc}}(\mathbf{x}')/\partial n(\mathbf{x}')$, are discretized by collocation. That is, instead of requiring that (1) hold for all \mathbf{x}' on Γ , we restrict \mathbf{x}' to a finite number of representative points on this surface, and require that (1) holds only at those points. Since the data enter into the approximation only at the collocation points, a large number of such points may be required for accurate modeling at high frequencies. On the other hand, in the Galerkin method the data are discretized by projection onto a set of test functions. With an appropriate choice of test functions, one expects the Galerkin method to yield a more efficient discretization of the data, and a more compact representation of Eq. (1).

In this paper, it is shown that it is possible to solve Eq. (1) efficiently by means of a Galerkin method in the case where Γ represents a rectangular screen, and that this method has significant advantages over the collocation approach. The method will initially be discussed with reference to the equation

$$g(\mathbf{x}') = \int_{\Omega} K(\mathbf{x}' - \mathbf{x}) f(\mathbf{x}) dV(\mathbf{x}), \quad (2)$$

since Eq. (1) reduces to this form when Γ is flat. The appropriate interpretation of the integral in (2) when K is singular is postponed until Secs. II and III. The Galerkin method discretizes Eq. (2) by using a collection of basis functions ϕ_j and test functions $\tilde{\phi}_i$, which results in the linear system

$$g_i = \sum_j K_{ij} f_j, \quad (3)$$

where

$$g_i = \int_{\Omega} \tilde{\phi}_i(\mathbf{x}') g(\mathbf{x}') dV', \quad f(\mathbf{x}) = \sum_j f_j \phi_j(\mathbf{x}), \quad (4)$$

and

$$K_{ij} = \int_{\Omega} \int_{\Omega} \tilde{\phi}_i(\mathbf{x}') K(\mathbf{x}' - \mathbf{x}) \phi_j(\mathbf{x}) dV dV'. \quad (5)$$

It will be shown that by using the scaling functions of discrete wavelet theory for both basis and test functions, the four-dimensional integrals in Eq. (5) can be analytically reduced to two-dimensional integrals. This feature will be discussed in Sec. I. In addition, the use of scaling functions enables integrals containing singularities of the form $|\mathbf{x}' - \mathbf{x}|^{-\alpha}$ to be computed by means of a recurrence formula, which is initiated by integrals over nonsingular regions of the kernel. The evaluation of such integrals in the cases $\alpha = 1$ and $\alpha = 3$ is discussed in Sec. II.

Once the matrix K_{ij} has been found, it is still necessary to solve (3) for f_j . This may be done using an iterative equation solver if K is sufficiently well conditioned. Such solvers require the efficient computation of the matrix-vector product $\sum_j K_{ij} u_j$. When the kernel is convolutional and the basis functions are suitably chosen, this product can be rapidly evaluated using the fast Fourier transform (FFT). This is discussed in Secs. I and IV.

Results of a comparison between the Galerkin method and a collocation-based boundary element method (CBEM)

are given in Sec. IV. The test problem used for the comparison was the acoustic diffraction of an incident plane wave by a unit square screen. In a calculation of the field diffracted by the screen, the rate of convergence of each method was compared for three angles of incidence over a range of wave number. Finally, a summary of the paper and conclusions are given in Sec. V.

I. THE GALERKIN METHOD

The solution of Eq. (2), when Ω is the unit square in \mathbb{R}^2 , will be discussed here. The more general case, where Ω is rectangular, can be solved using the same method, by transforming the region of integration back to the unit square. Although the singular nature of K is not discussed here, the method is applicable to the case where K is singular, as long as the integral in Eq. (2) is interpreted appropriately. The evaluation of integrals containing singularities requires specialized techniques, and will be deferred to Sec. II.

When Ω is the unit square, Eq. (2) can be written as

$$g(x', y') = \int_0^1 \int_0^1 K(x' - x, y' - y) f(x, y) dx dy. \quad (6)$$

In the discretization of (6) it is desirable that the resulting linear system inherits the convolutional structure of the equation so that a fast solution employing the FFT is possible. This can be achieved if both basis and test functions are constructed from the regular translations of a single compactly supported function. The integrals in (4) and (5) can then be interpreted as local averages of the data and the kernel. Families of one-dimensional functions of this kind have recently been constructed and put into a mathematically rigorous framework.^{12,13} In each family, a basis is constructed from integer translations of a single function ϕ which satisfies the equation

$$\phi(x) = \sqrt{2} \sum_{n=N_1}^{N_2} h_n \phi(2x - n), \quad (7)$$

which is called the refinement or dilation equation. ϕ is known in the wavelet literature as a scaling function, and the numbers h_n which define it are called filter coefficients. The properties of these functions have been thoroughly investigated, and methods for generating evenly spaced samples of ϕ exist.¹²⁻¹⁴

Here, basis and test functions for the solution of (6) are constructed from the tensor products of scaled and translated versions of ϕ , which are defined by

$$\phi_m^j(x) = 2^{j/2} \phi(2^j x - m), \quad (8)$$

where both j and m are integers. It is straightforward to show from (7) that $\phi_m^j(x)$ is nonzero only on the interval $2^{-j}[N_1 + m, N_2 + m]$, that is, it is compactly supported.

Rather than solving (6) as it stands, it was found to be better to extend the range of integration to all of \mathbb{R}^2 , giving

$$g(x', y') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x' - x, y' - y) f(x, y) dx dy, \quad (9)$$

and to express $f(x, y)$ as a linear combination of functions which have negligible energy outside the unit square. In the

TABLE I. Filter coefficients for the first eight B-spline scaling functions.

N	n	$\sqrt{2}_N h_n$	N	n	$\sqrt{2}_N h_n$		
1	0,1	1	2	-1,1	$\frac{1}{2}$		
				0	1		
3	-1,2	$\frac{1}{4}$	4	-2,2	$\frac{1}{8}$		
				0,1	$\frac{1}{2}$		
					$\frac{3}{4}$		
5	-2,3	$\frac{1}{16}$	6	-3,3	$\frac{1}{32}$		
				-1,2	$\frac{3}{16}$		
					0,1	$\frac{15}{32}$	
						$\frac{5}{8}$	
7	-3,4	$\frac{1}{64}$	8	-4,4	$\frac{1}{128}$		
				-2,3	$\frac{1}{16}$		
					-1,2	$\frac{7}{32}$	
						0,1	$\frac{7}{32}$
							$\frac{35}{64}$

problem of acoustic diffraction by thin objects it is known that the potential difference vanishes at the edges of the object, and that there is a singularity in the slope of this function at the edges.¹ It is therefore desirable to choose scaling functions which also respect these edge properties as much as possible.

The family of orthonormal scaling functions constructed by Daubechies¹² was found to be unsatisfactory due to its asymmetry. It was found that the use of asymmetric scaling functions as a basis tended to result in asymmetric reconstructions of f even when this was known to be symmetric. The cardinal B-spline functions are much better candidates for basis and test functions. These functions are compactly supported piecewise polynomials which satisfy Eq. (7). Filter coefficients for the first eight of these functions are listed in Table I. Throughout this paper, the B-spline with coefficients ${}_N h_n$ will be denoted as ${}_N \phi(x)$. All the functions in this family are symmetric about their centers, and it can be shown that ${}_N \phi$ is C^{N-2} . Each B-spline has a biorthogonal dual function, the translates of which could be used for the test functions, but they were not sufficiently smooth for this purpose. However, it is not essential that the basis and test functions be orthogonal, and so the B-splines were used for both. The first B-spline suitable as a basis function is ${}_2 \phi$, the triangular hat function. The basis, for $N \geq 2$, is defined by

$$\{\Phi_{mn}^j(x, y) = {}_N \phi_m^j(x) {}_N \phi_n^j(y), \quad m, n \in M\},$$

$$M = \{1, \dots, 2^j - 1\}. \quad (10)$$

In the case $N=2$, this basis ensures that the unknown is zero on the edges of the unit square. The other edge property, that there is a singularity in the slope at the edges, is not satisfied. In fact, when $N > 2$ the basis in (10) fails to satisfy either of the edge properties. However, as j increases, the basis functions become increasingly localized and are better able to represent the rapid variation at the edges. This behavior is illustrated, in one dimension, in Fig. 1. Moreover, numerical experiments showed that the solutions obtained for $N=2, 3$, and 4 agreed very closely both between themselves and with

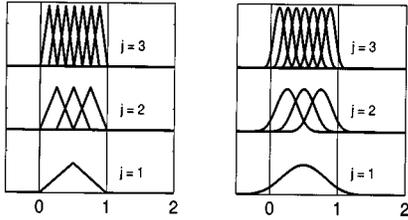


FIG. 1. The $k=1$ through $k=2^j-1$ translates of ${}_2\phi_k^j$ (left) and ${}_4\phi_k^j$ (right) are shown, for increasing j . The basis for $[0,1]^2$ is constructed from tensor products of these functions.

the collocation method described in Sec. IV. This agreement was felt to justify the use of the basis.

In the comparisons between the Galerkin and collocation methods presented in Sec. IV, the basis in (10) for $N = \{2,3,4\}$ was used for $[0,1]^2$, and the test functions were also chosen from this set. The Galerkin discretization of (9) is then

$$g_{m'n'} = \sum_{m,n \in M} K_{m'n'mn} f_{mn}, \quad m', n' \in M, \quad (11)$$

where, letting $\int dS$ denote integration over the xy plane,

$$g_{m'n'} = \int \Phi_{m'n'}^j(x,y) g(x,y) dS, \quad (12)$$

$$f(x,y) = \sum_{m,n \in M} f_{mn} \Phi_{mn}^j(x,y),$$

and

$$K_{m'n'mn} = \int \int \Phi_{m'n'}^j(x',y') \times K(x'-x, y'-y) \Phi_{mn}^j(x,y) dS dS'. \quad (13)$$

With a change of variable, the double integral in (13) may be converted to

$$\int K(x',y') \int_{-\infty}^{\infty} \phi_{m'}^j(x+x') \phi_m^j(x) dx \times \int_{-\infty}^{\infty} \phi_{n'}^j(y+y') \phi_n^j(y) dy dS'. \quad (14)$$

It turns out that there is a useful property of scaling functions, the correlation property, which enables the integrals over x and y in (14) to be evaluated analytically. It can be shown that the correlation of two scaling functions is also a scaling function, since the correlation function satisfies Eq. (7). To evaluate the integrals in (14) by way of this property, first let $\hat{\phi}$ denote the autocorrelation of ϕ , that is

$$\hat{\phi}(x') = \int_{-\infty}^{\infty} \phi(x+x') \phi(x) dx. \quad (15)$$

By substituting for ϕ in this expression, using Eq. (7) it can be shown that $\hat{\phi}$ satisfies the dilation equation

$$\hat{\phi}(x') = \sqrt{2} \sum_{k=-(N_2-N_1)}^{N_2-N_1} \hat{h}_k \hat{\phi}(2x'-k), \quad (16)$$

where

$$\hat{h}_k = \frac{1}{\sqrt{2}} \sum_l h_l h_{l+k}. \quad (17)$$

Then, by using (8), the integral over x in (14) is given by

$$\int_{-\infty}^{\infty} \phi_{m'}^j(x+x') \phi_m^j(x) dx = \hat{\phi}(2^j x' + m - m') = 2^{-(j/2)} \hat{\phi}_{m'-m}^j(x'), \quad (18)$$

and of course a similar relationship holds for the integral over y . Consequently, (13) becomes

$$K_{m'n'mn} = 2^{-j} \int K(x,y) \hat{\Phi}_{m'-m, n'-n}^j(x,y) dS, \quad (19)$$

where $\hat{\Phi}_{kl}^j(x,y) = {}_N\hat{\phi}_k^j(x) {}_N\hat{\phi}_l^j(y)$. Values of $\hat{\phi}$ at the dyadic rationals can be calculated by way of the algorithms given by Strang,¹⁴ and so quadrature rules with evenly spaced points can be used for the numerical evaluation of the nonsingular parts of (19). It can be shown that the B -spline scaling functions satisfy the rule ${}_N\hat{\phi} = {}_{2N}\phi$, that is, the autocorrelation of a B -spline is a B -spline of twice the degree. Thus the autocorrelations of the set of functions $\{{}_N\phi, N=2,3,4\}$, which were used in the numerical tests of Sec. IV are $\{{}_N\phi, N=4,6,8\}$, respectively. The filter coefficients for all of these functions are listed in Table I.

Because the basis has been constructed from the regular translations of a single function, the integrals in (19) are functions of $m'-m$ and $n'-n$ and not of these indices individually. Consequently, one may put $K_{m'n'mn} = R_{m'-m, n'-n}$ and (11) may be written as the two-dimensional Toeplitz system,

$$g_{m'n'} = \sum_{m,n} R_{m'-m, n'-n} f_{mn}, \quad (20)$$

which can be solved efficiently with the FFT in conjunction with an iterative solver. In addition, since the size of R is the square root of the size of K , the cost of construction of the stiffness matrix is greatly reduced, and the use of much larger numbers of variables in the basis is feasible.

II. EVALUATION OF SINGULAR INTEGRALS

When Γ is rectangular, Eq. (1) may be solved by the Galerkin method presented in the previous section. In the calculation of the matrix elements for this equation it is necessary to evaluate certain singular integrals of the form

$$(\text{CPV or HFP}) \int \hat{\Phi}_{kl}^j(x,y) (x^2+y^2)^{-\alpha/2} dS. \quad (21)$$

As indicated, the existence of these integrals is dependent upon their interpretation as either Cauchy principal values or Hadamard finite parts, according to the value of α . These integrals arise from a consideration of the limit in Eq. (1), which is discussed in Sec. III. Here, a method for the accurate numerical computation of the integrals in (21) is given. The method is based upon the recursive definition of scaling functions, which enables a recursive formula for these integrals to be constructed. To derive this formula, one begins with the definition

$$R_{kl}^\alpha = \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^2 \setminus C_\epsilon} \Phi_{kl}(x,y) (x^2 + y^2)^{-\alpha/2} dS, \quad (22)$$

where $\mathbb{R}^2 \setminus C_\epsilon$ denotes the xy plane with a disc of radius ϵ centered on the origin removed, and $\hat{\Phi}_{kl}(x,y) = \hat{\phi}_k(x)\hat{\phi}_l(y)$. (The hat on Φ has been removed since the following applies to all scaling functions, not only those which are the autocorrelations of other scaling functions.) By using the dilation equation for ϕ it can be seen that

$$R_{kl}^\alpha = 2 \sum_{m,n=N_1}^{N_2} h_m h_n \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^2 \setminus C_\epsilon} \Phi_{2k+m,2l+n}(2x,2y) \times (x^2 + y^2)^{-\alpha/2} dS, \quad (23)$$

and then, with a change of variable,

$$R_{kl}^\alpha = 2^{\alpha-1} \sum_{m,n=N_1}^{N_2} h_m h_n R_{2k+m,2l+n}^\alpha. \quad (24)$$

When both k and l are in the set $\{-N_2, \dots, -N_1\}$, the integrand of R_{kl}^α contains a singularity. Therefore, given that $N_1 < N_2$, the integrals on the right-hand side of this system include all of the singular integrals on the left-hand side as well as integrals which are nonsingular, that is, cases in which the support of $\Phi_{2k+m,2l+n}$ does not include the origin. By evaluating these nonsingular integrals using ordinary numerical quadrature, Eq. (24) can be made into a square system of equations and solved for the unknown singular integrals. If Φ_{kl} in (22) is replaced by its scaled relative Φ_{kl}^j then the integral, denoted $R_{kl}^{\alpha,j}$, is given by

$$R_{kl}^{\alpha,j} = 2^{(\alpha-1)j} R_{kl}^\alpha, \quad (25)$$

which can be shown by a change of variable.

The solution of Eq. (24) can be demonstrated in the case $\alpha = 1$ with the function

$${}_1\phi(x) = \begin{cases} 1 & 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases},$$

which is the first of the B -spline scaling functions. It satisfies the dilation equation

$${}_1\phi(x) = {}_1\phi(2x) + {}_1\phi(2x-1). \quad (26)$$

Equation (24), written in matrix-vector form for this function, is

$$\begin{bmatrix} R_{-1-1}^1 \\ R_{0-1}^1 \\ R_{-10}^1 \\ R_{00}^1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} R_{-1-1}^1 \\ R_{0-1}^1 \\ R_{-10}^1 \\ R_{00}^1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} R_{-2-2}^1 + R_{-1-2}^1 + R_{-2-1}^1 \\ R_{0-2}^1 + R_{1-2}^1 + R_{1-1}^1 \\ R_{-20}^1 + R_{-21}^1 + R_{-11}^1 \\ R_{10}^1 + R_{01}^1 + R_{11}^1 \end{bmatrix}. \quad (27)$$

The second term on the right of (27) contains only integrals which are nonsingular. By evaluating these numerically, the system becomes

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} R_{-1-1}^1 \\ R_{0-1}^1 \\ R_{-10}^1 \\ R_{00}^1 \end{bmatrix} = 1.762747 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad (28)$$

which yields the trivial solution 1.762 747 for each integral. This value may be confirmed by analytic integration. Evaluating R_{kl}^α in this fashion is too cumbersome for the higher-order B -splines, consequently a MATLAB code was written to both construct and solve the system in (24). To check the validity of the method, the integrals R_{kl}^1 for tensor products of the B -spline ${}_2\phi$ were derived by analytic means. These values are given in the Appendix, and were found to be in good agreement with those determined from (24).

When $\alpha \geq 2$, the integrals defined in (22) are divergent. However, in the case where $\alpha = 3$, if R_{kl}^3 is redefined by

$$R_{kl}^3 = \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^2 \setminus C_\epsilon} \Phi_{kl}(x,y) (x^2 + y^2)^{-3/2} dS - \frac{2\pi\Phi_{kl}(0,0)}{\epsilon}, \quad (29)$$

then it turns out that these quantities also satisfy Eq. (24) with $\alpha = 3$. This definition is known as a Hadamard finite part.¹⁵ It is necessary to calculate these quantities for the solution of (1) by the Galerkin method. Problematically, the system of equations in (24) was found to be singular when $\alpha = 3$. However, experimentation revealed that solutions to this system can be obtained when ϕ is symmetric about the origin. In this case, the integrals obey the eightfold symmetry relation $R_{kl}^\alpha = R_{\pm k, \pm l}^\alpha = R_{\pm l, \pm k}^\alpha$, and so a considerable redundancy exists in (24). This redundancy can be eliminated by adding together columns in the matrix representation of the system which correspond to variables that are known to be equal by symmetry. In general, the resulting overdetermined system of equations must be solved in a least-squares sense. Whether this solution is exact to within round-off error can be checked by substitution into (24). It was found by experiment that such numerically exact solutions of (24) could be obtained from the least-squares solution of the overdetermined system. The process of constructing this modified system can be easily automated, and in this manner it was possible to find the integrals R_{kl}^1 and R_{kl}^3 for the functions

$$\{\Phi_{kl}(x,y) = {}_N\phi_k(x){}_N\phi_l(y), \quad N = 4, 6, 8\}, \quad (30)$$

which were needed for the numerical tests described in Sec. IV. The values of these integrals are given in Table II.

III. IMPLEMENTATION OF THE METHOD

Problems of radiation and scattering by thin objects present numerical difficulties when the standard BIEs are applied. It is known that discretizations of the Helmholtz integral formula are poorly conditioned for small scatterer thickness.¹¹ To avoid this problem, a finitely thin scatterer can be modeled as having zero thickness. In the limit of zero thickness, the governing equation is the normal derivative of the Helmholtz integral formula, which reduces to Eq. (1) when the scatterer is rigid. The numerical analysis of this

TABLE II. The singular integrals R_{kl}^1 and R_{kl}^3 for the functions $\{\hat{\Phi}_{kl}(x,y) = {}_N\phi_k(x){}_N\phi_l(y), N=4,6,8\}$. Values not listed can be found from the relation $R_{kl}^\alpha = R_{\pm k, \pm l}^\alpha = R_{\pm l, \pm k}^\alpha$.

N	k	l	R_{kl}^1	R_{kl}^3	
4	0	0	2.110 023 118 35	-5.811 966 397 90	
	0	1	1.182 094 513 59	-0.177 591 204 97	
	0	2	0.524 610 134 20	0.223 647 211 32	
	1	1	0.808 074 159 68	0.471 931 146 68	
	1	2	0.465 185 338 39	0.140 678 505 51	
	2	2	0.361 933 328 56	0.055 945 658 24	
6	0	0	1.738 742 548 93	-3.225 704 020 82	
	0	1	1.145 452 607 62	-0.609 227 350 20	
	0	2	0.545 006 005 66	0.256 242 290 78	
	0	3	0.343 728 257 44	0.051 229 994 61	
	1	1	0.834 031 047 38	0.109 727 158 38	
	1	2	0.478 799 403 34	0.177 427 568 14	
	1	3	0.325 133 548 32	0.042 201 886 52	
	2	2	0.367 149 136 57	0.068 108 796 99	
	2	3	0.283 280 692 85	0.026 346 821 15	
	3	3	0.239 240 103 72	0.015 109 649 32	
	8	0	0	1.513 143 677 59	-2.143 781 683 53
		0	1	1.091 723 228 85	-0.655 540 905 36
0		2	0.563 781 152 73	0.193 216 649 16	
0		3	0.348 268 897 04	0.062 461 195 83	
0		4	0.255 690 261 63	0.019 583 279 85	
1		1	0.836 217 689 02	-0.078 865 772 90	
1		2	0.493 448 352 08	0.163 582 031 67	
1		3	0.328 882 335 23	0.050 019 623 76	
1		4	0.247 732 256 39	0.017 622 345 02	
2		2	0.373 471 805 41	0.080 120 043 14	
2		3	0.285 601 285 67	0.029 313 865 50	
2		4	0.227 658 850 62	0.013 329 036 69	
3		3	0.240 547 099 37	0.016 044 682 72	
3		4	0.202 861 910 66	0.009 167 655 36	
4		4	0.178 722 718 43	0.006 125 012 16	

equation by the Galerkin method in the case of a flat scatterer is discussed in this section. The method is implemented by separating the kernel into its hypersingular, weakly singular, and nonsingular parts. Integrals involving singular parts of the kernel may be evaluated using the methods described in Sec. II, while the nonsingular integrals may be evaluated using numerical quadrature.

The derivation of the thin scatterer equation begins with the Helmholtz integral formula

$$u(\mathbf{x}') = \int_{\Gamma} u(\mathbf{x}) \frac{\partial G(\mathbf{x}, \mathbf{x}')}{\partial n(\mathbf{x})} - G(\mathbf{x}, \mathbf{x}') \frac{\partial u(\mathbf{x})}{\partial n(\mathbf{x})} dS(\mathbf{x}) + u^{\text{inc}}(\mathbf{x}'). \quad (31)$$

In this equation, $n(\mathbf{x})$ is the coordinate of the normal to the surface Γ at \mathbf{x} , u^{inc} is the incident acoustic potential, and G is the free space Green's function which in three dimensions is

$$G(\mathbf{x}, \mathbf{x}') = \frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x}-\mathbf{x}'|}. \quad (32)$$

In the case where Γ is vanishingly thin, it is necessary to distinguish the two sides of Γ as Γ^+ and Γ^- . Allowance must be made for a discontinuity in both the potential and its normal derivative on Γ , and so it is useful to make the definitions

$$\Delta u = u^+ - u^- \quad \text{and} \quad \Delta u_n = \frac{\partial u^+}{\partial n} - \frac{\partial u^-}{\partial n}, \quad (33)$$

where $u^+ = u(\Gamma^+)$ and $u^- = u(\Gamma^-)$. Then, the normal derivative of the acoustic potential $u(\mathbf{x})$ satisfies^{2,7}

$$\frac{\partial u(\mathbf{x}')}{\partial n(\mathbf{x}')} = - \int_{\Gamma} \Delta u(\mathbf{x}) \frac{\partial^2 G(\mathbf{x}, \mathbf{x}')}{\partial n(\mathbf{x}) \partial n(\mathbf{x}')} - \Delta u_n(\mathbf{x}) \frac{\partial G(\mathbf{x}, \mathbf{x}')}{\partial n(\mathbf{x}')} dS(\mathbf{x}) + \frac{\partial u^{\text{inc}}(\mathbf{x}')}{\partial n(\mathbf{x}')}, \quad (34)$$

where \mathbf{x}' is any point that does not lie on Γ , and $n(\mathbf{x})$ is now the coordinate along the normal to Γ^+ , pointing away from Γ^- . The field at a point $\mathbf{x}' \in \Gamma$ can be interpreted as the limiting value of the integral at a point which approaches \mathbf{x}' from the interior. Consequently, when the scatterer is rigid and both \mathbf{x}' and \mathbf{x} lie on Γ , one has

$$\frac{\partial u^{\text{inc}}(\mathbf{x}')}{\partial n(\mathbf{x}')} = \lim_{\mathbf{x}'' \rightarrow \mathbf{x}'} \int_{\Gamma} \Delta u(\mathbf{x}) \frac{\partial^2 G(\mathbf{x}, \mathbf{x}'')}{\partial n(\mathbf{x}) \partial n(\mathbf{x}'')} dS(\mathbf{x}). \quad (35)$$

It is important to note that the operations of taking the limit and integration do not commute here. The kernel inside the integral is hypersingular, and if the limit in (35) is taken before integration, the integral will diverge.

The special case when Γ is a region in the xy plane will now be considered. Let \mathbf{x} and \mathbf{x}' be points in the xy plane, and let \mathbf{x}'' be the point lying a distance D vertically above \mathbf{x}' , that is, $\mathbf{x}'' = \mathbf{x}' + (0,0,D)$. It can be shown that

$$\frac{\partial^2 G(\mathbf{x}, \mathbf{x}'')}{\partial n(\mathbf{x}) \partial n(\mathbf{x}'')} = \left[\left(\frac{i3k}{r} + k^2 - \frac{3}{r^2} \right) D^2 - ikr + 1 \right] \frac{e^{ikr}}{4\pi r^3} = K(x' - x, y' - y, D), \quad (36)$$

where $r = |\mathbf{x}'' - \mathbf{x}|$. In this case, Eq. (35) is a convolutional integral equation. The matrix elements of the Galerkin method applied to this equation with the kernel given by (36) are defined by

$$K_{m'n'mn} = \lim_{D \rightarrow 0} \int_{\Gamma} \int_{\Gamma} \Phi_{m'n'}^j(x', y') \times K(x' - x, y' - y, D) \Phi_{mn}^j(x, y) dS dS'. \quad (37)$$

The first step in the calculation of these integrals is the reduction of the quadruple integral to a double integral by means of the correlation property of scaling functions discussed in Sec. I. Using the correlation property gives

$$K_{m'n'mn} = \lim_{D \rightarrow 0} 2^{-j} \int_{\Gamma} \Phi_{m'-m, n'-n}^j(x, y) K(x, y, D) dS. \quad (38)$$

Next, it is necessary to decompose the kernel into its singular and nonsingular parts, so that the singular integrals may be evaluated using the method given in Sec. II. Carrying out this decomposition gives

$$K(x,y,D) = -\frac{3D^2}{4\pi r^5} + \frac{1}{4\pi r^3} + \frac{k^2}{8\pi r} + \frac{(1-ik\rho)(e^{ik\rho}-1-ik\rho+\frac{1}{2}k^2\rho^2)}{4\pi\rho^3} + \frac{ik^3}{8\pi}. \quad (39)$$

Here, $\rho = (x^2 + y^2)^{1/2}$, $r = (\rho^2 + D^2)^{1/2}$, and the limit $D \rightarrow 0$ has been taken in the nonsingular part of this expression. The first two terms on the right of (39) are responsible for the divergent behavior of the integral in (35) when $\lim_{x',y' \rightarrow x}$ is taken inside the integral sign. It can be shown, given certain restrictions on f , that

$$\lim_{D \rightarrow 0} \int_{\Gamma} f(x,y) \left[-\frac{3D^2}{r^5} + \frac{1}{r^3} \right] dS = \lim_{\epsilon \rightarrow 0} \int_{\Gamma \setminus C_{\epsilon}} f(x,y) \rho^{-3} dS - \frac{2\pi f(0,0)}{\epsilon}, \quad (40)$$

where C_{ϵ} denotes a disc of radius ϵ centered on the origin. The right-hand side is a Hadamard finite part which was discussed in Sec. II in the context of scaling functions. There it was shown how such integrals may be evaluated when f is the tensor product of scaling functions. The issue of existence of these integrals was not discussed. In fact, it can be shown that if f is $C^{1,\alpha}$ Hölder continuous, that (40) exists.⁹ In the numerical tests of the Galerkin method the performance of basis functions constructed from tensor products of the B -splines $\{N\phi, N=2,3,4\}$ was investigated. While the tensor products of 3ϕ and 4ϕ are $C^{1,\alpha}$, those of 2ϕ are not. However, this presents no practical impediment to its use as a basis function. If Δu is expanded in a basis using tensor products of 2ϕ then the points at which the basis fails to be $C^{1,\alpha}$ form a set of measure zero, and consequently the integrals in (37) exist.

The third term on the right of (39) is weakly singular in the limit $D \rightarrow 0$. It can be shown that

$$\lim_{D \rightarrow 0} \int_{\Gamma} f(x,y) r^{-1} dS = \lim_{\epsilon \rightarrow 0} \int_{\Gamma \setminus C_{\epsilon}} f(x,y) \rho^{-1} dS, \quad (41)$$

that is, the limit of the integral on the left reduces to a Cauchy principal value integral. The evaluation of the principal value when f is a tensor product of scaling functions was also discussed in Sec. II. [It should be pointed out that there is no real need to reinterpret the integrals involving the singular terms in (39) as finite parts and principal values. It can be shown that the limit as $D \rightarrow 0$ of the integrals in (40) and (41) satisfies Eq. (24) with $\alpha = 3$ and $\alpha = 1$, respectively, when f is the tensor product of scaling functions.]

With the aid of the decomposition in (39), and the rescaling relation in (25), the matrix elements may be expressed as

$$K_{m'n'mn} = \frac{2^j}{4\pi} R_{m'-m,n'-n}^3 + \frac{2^{-j}k^2}{8\pi} R_{m'-m,n'-n}^1 + \frac{2^{-j}}{4\pi} R_{m'-m,n'-n}^{\text{reg}}, \quad (42)$$

where

$$R_{kl}^{\text{reg}} = \int \Phi_{kl}^j \left\{ \rho^{-3} (1-ik\rho) \left(e^{ik\rho} - 1 - ik\rho + \frac{1}{2}k^2\rho^2 \right) + \frac{ik^3}{2} \right\} dS. \quad (43)$$

Those of the integrals R_{kl}^3 and R_{kl}^1 which are singular may be calculated with the aid of Eq. (24), and these values, for the functions listed in (30), are given in Table II. The regular integrals denoted by R_{kl}^{reg} , and those of the integrals R_{kl}^3 and R_{kl}^1 which are nonsingular, may be calculated with ordinary numerical quadrature. For the work presented in Sec. IV Simpson's rule was used.

IV. NUMERICAL RESULTS

Previous reported solutions of (35) for flat screens have been obtained using methods based on collocation.^{2,5,6,9,10} Here, the efficiency of a collocation-based boundary element method (CBEM) is compared to the Galerkin method by calculating the field diffracted by a unit square screen irradiated by a plane wave. The convergence of each method was tested for three different angles of incidence. The elements used in the CBEM were square, with a single collocation point at their centers. On each element the potential difference was approximated by a constant, and the matrix elements were calculated by the method of Terai.² Because of the spatial uniformity of the BEM mesh, the stiffness matrix was two-dimensionally Toeplitz, as in (20).

Although higher-order finite-element functions could in principle be used to represent Δu , the main concern in this paper is the accuracy of representation of the data, $(\partial u^{\text{inc}}/\partial n)$. This is not improved with the use of higher-order finite-element functions, since the data enter the approximation only at the collocation points. The data are more effi-

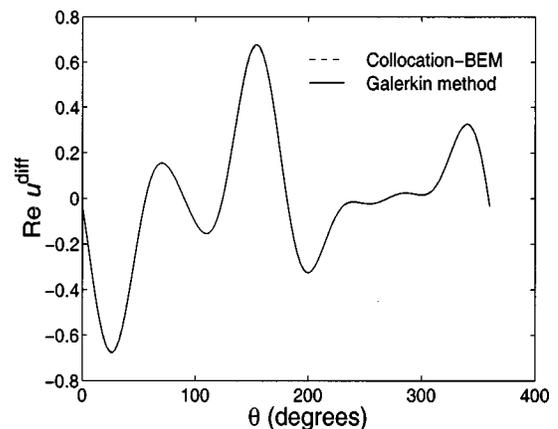


FIG. 2. The real part of the diffracted field, $k = 10$, $\theta = 30^\circ$.

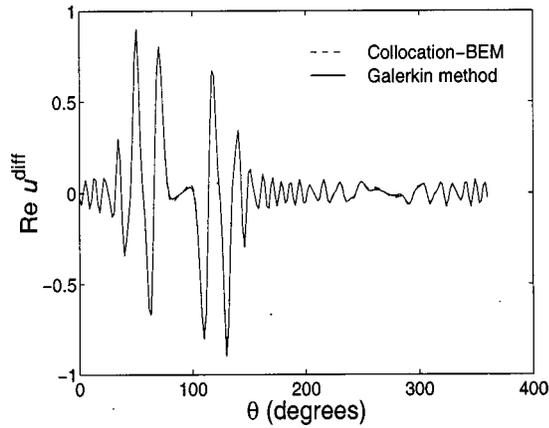


FIG. 3. The real part of the diffracted field, $k=100$, $\theta=60^\circ$.

ciently represented in the Galerkin method, by projection onto the set of test functions, which yields a comparatively smaller system of equations.

For both methods it was necessary to solve the two-dimensional Toeplitz system

$$g_{m'n'} = \sum_{m,n} R_{m'-m,n'-n} f_{mn}. \quad (44)$$

To obtain accurate solutions at high frequencies, from 10^3 to 10^4 basis functions were required. Because of the singularity in the kernel, the system matrix K is generally well conditioned, and so it is possible to obtain solutions by an iterative method. The stabilized biconjugate gradient method¹⁶ was found to be effective for the solution of this system (it is available as the routine BICGSTAB in MATLAB 5). All iterative methods work toward a solution by repeated multiplication of vectors by K , and so to be faster than direct methods an efficient means of calculating $K\mathbf{x}$ is essential. Because of the two-dimensional Toeplitz structure of K this product may be calculated efficiently through the use of the FFT. To use the FFT, Eq. (44) must be expressed as a circular convolution, which can be achieved by appropriate zero padding.

The rate of convergence of each method was tested by calculating the field diffracted by a plane wave incident on a screen occupying the unit square in the xy plane. The wave

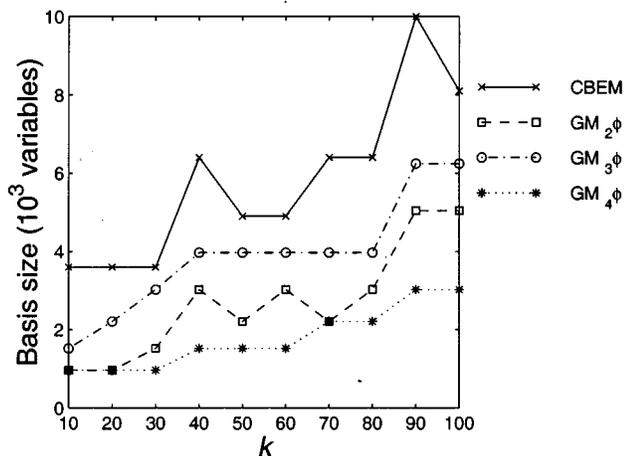


FIG. 4. The basis size required for convergence of the collocation-BEM (CBEM) and the Galerkin method (GM), for $\theta=0^\circ$ (normal incidence).

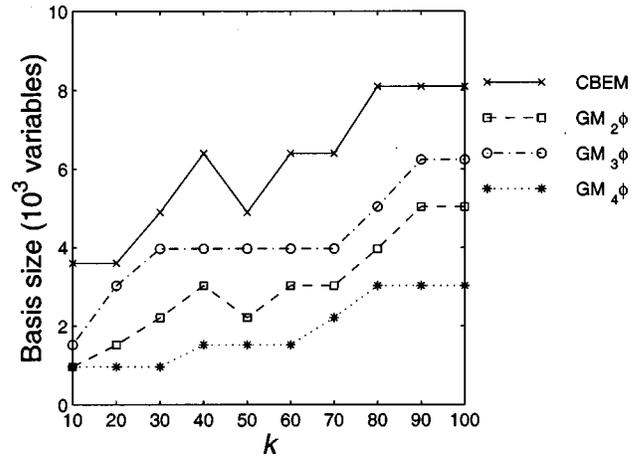


FIG. 5. The basis size required for convergence of the collocation-BEM (CBEM) and the Galerkin method (GM), for $\theta=30^\circ$.

vector of the incident plane wave was chosen to be parallel to the xz plane, and the diffracted field was determined at 201 equispaced points on the circle satisfying $(x-1/2)^2 + z^2 = 1$, in the plane $y=1/2$. The diffracted field at these points due to the potential difference Δu was calculated numerically from

$$u^{\text{diff}}(\mathbf{x}') = - \int_{\Gamma} \Delta u(\mathbf{x}) \frac{\partial G(\mathbf{x}, \mathbf{x}')}{\partial n(\mathbf{x})} dS(\mathbf{x}). \quad (45)$$

There was good, but not perfect, agreement between values of u^{diff} calculated by the two methods. The typical best and worst case comparisons are demonstrated in Figs. 2 and 3, which show, respectively, the real part of the diffracted field for $k=10$ and $\theta=30^\circ$, and for $k=100$ and $\theta=60^\circ$.

Convergence was measured by solving the CBEM and Galerkin systems with an increasing number of variables until

$$\sum_i |u_i^{\text{diff},n} - u_i^{\text{diff},n-1}|^2 < 10^{-4} \sum_i |u_i^{\text{diff},n}|^2, \quad (46)$$

where $u_i^{\text{diff},n}$ is the diffracted field at the i th point on the circle at iteration n . In the Galerkin method, increases in the

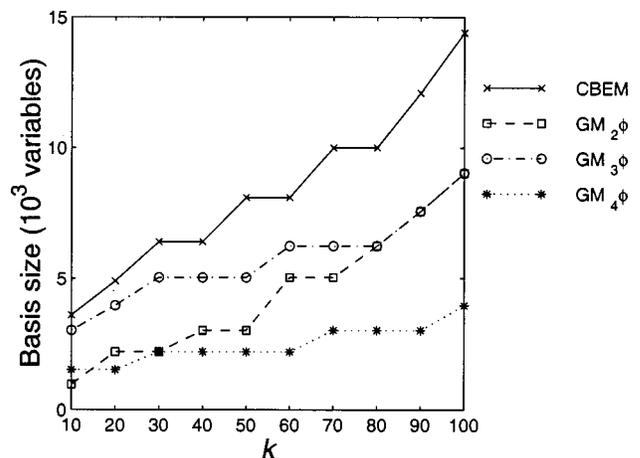


FIG. 6. The basis size required for convergence of the collocation-BEM (CBEM) and the Galerkin method (GM), for $\theta=60^\circ$.

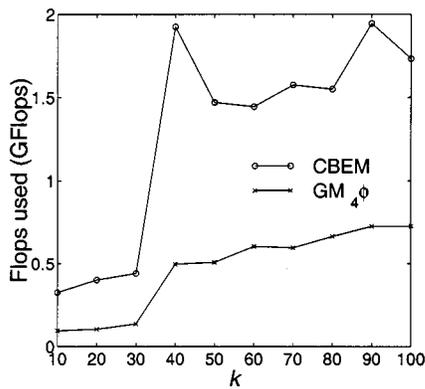


FIG. 7. Floating point operations (flops) used by the collocation-BEM (CBEM) and the Galerkin method (GM) for $\theta=0^\circ$ (normal incidence).

size of the basis were not made by increasing the scale j , but rather by fixing the scale and rescaling the geometry of the problem when more variables were needed. This approach provided more flexibility in setting the size of the basis.

The number of variables required for convergence, as defined by (46), was determined for the angles of incidence 0 (normal incidence), 30, and 60°, for wave numbers from $k=10$ to $k=100$ in steps of 10. The results, shown in Figs. 4–6, show that the Galerkin method outperformed the collocation method for each of the B -spline functions used as a basis. The basis derived from the cubic B -spline (4ϕ) significantly outperformed those derived from the linear (2ϕ) and quadratic (3ϕ) B -splines. At the largest tested wave number and with an incident angle of 60°, the Galerkin method, using the cubic B -spline, converged with less than one-third of the number of variables required for convergence of the CBEM. Moreover, the advantage of the Galerkin method clearly increases with frequency.

The more compact representation of the system by the Galerkin method results in reduced computation time. Figures 7–9 compare the computation time, measured by the number of floating point operations used, for the cases presented in Figs. 4–6. The best performance of the Galerkin method was obtained using the cubic B -spline, and only these results are shown. As described previously, the iterative solver employs the two-dimensional fast Fourier transform (FFT2) to compute the discrete convolution required at each iteration. To enable the use of a fast radix-2 FFT, the

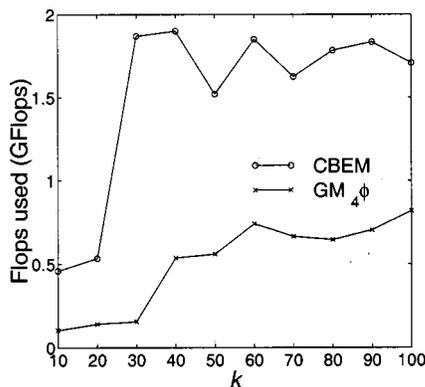


FIG. 8. Floating point operations (flops) used by the collocation-BEM (CBEM) and the Galerkin method (GM) for $\theta=30^\circ$.

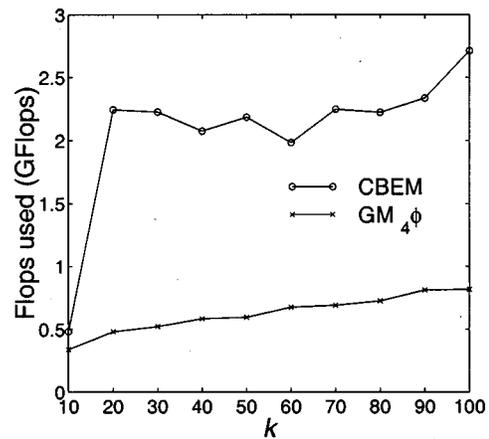


FIG. 9. Floating point operations (flops) used by the collocation-BEM (CBEM) and the Galerkin method (GM) for $\theta=60^\circ$.

matrix R in (44) was padded out to the nearest power of two. The jump in computation time for the collocation method evident in each figure is caused by going from a 128×128 to a 256×256 point FFT2. By contrast, in the Galerkin method the more compact basis enabled a 128×128 point FFT2 to be used in each case.

As far as memory requirements are concerned, because of the FFT-based iterative solution method, the entire system matrix is never explicitly constructed. Memory is mainly expended on the storage of the matrix R in (44) (suitably padded with zeros) which in the numerical tests had dimensions 128×128 for the Galerkin method and in most cases 256×256 for the collocation method. This is roughly 0.25 and 1 MB of memory, respectively, and so memory requirements for both methods are not very significant.

V. SUMMARY

A hypersingular boundary integral equation, which models acoustic diffraction by a rectangular screen, has been solved by means of a Galerkin method using scaling functions as basis and test functions. The principal obstacle to the implementation of such a technique, the evaluation of four-dimensional singular and regular integrals, was overcome by the recursive nature of scaling functions. The use of these functions permits the accurate evaluation of the singular integrals and the reduction of the number of dimensions for which numerical integration is required from four to two. The numerical performance of the Galerkin method was compared with a constant element collocation-based boundary element method. It was shown that considerable reductions in basis size could be achieved, with consequent improvements in computation time over the collocation method. This method seems at present to be limited to flat surfaces, since the quadruple integral in Eq. (13) cannot otherwise be reduced to a double integral.

APPENDIX: ANALYTIC VALUES OF SOME SINGULAR INTEGRALS

To check that the computation of the singular integrals by means of Eq. (24) was accurate, comparison with some exact analytic values was made. A completely analytic

evaluation is possible for certain Cauchy principal value integrals defined by Eq. (22). The integrals R_{00}^1 , $R_{\pm 1 \pm 1}^1$, and $R_{\pm 1,0}^1$ were determined for $\Phi_{kl}(x,y) = {}_2\phi_k(x) {}_2\phi_l(y)$ analytically. The results were

$$\begin{aligned}
 R_{00}^1 &= \frac{-4\sqrt{2}+4}{3} + 4 \ln(1+\sqrt{2}), \\
 R_{\pm 1 \pm 1}^1 &= \frac{2}{3}(\sqrt{2}-1) + 2 \ln\left(\frac{2+\sqrt{5}}{1+\sqrt{2}}\right) + \frac{2}{3}(\sqrt{5}-\sqrt{2}-2) \\
 &\quad + \frac{8}{3}\sqrt{5}-4\sqrt{2} + 8 \ln\left(\frac{2(\sqrt{2}+1)}{\sqrt{5}+1}\right) \\
 &\quad + 6 \ln\left(\frac{\sqrt{5}-2}{\sqrt{2}-1}\right), \tag{A1} \\
 R_{\pm 10}^1 = R_{0\pm 1}^1 &= -\frac{5}{3}\sqrt{5} + 4 \ln(\sqrt{5}+1) - 4 \ln(2) + \frac{5}{3} \\
 &\quad + \frac{4}{3}\sqrt{2} + 2 \ln(\sqrt{2}-1) - 2 \ln(\sqrt{2}+1) \\
 &\quad - 2 \ln(\sqrt{5}-2).
 \end{aligned}$$

In general, a completely analytic evaluation of R_{kl}^3 , defined in Eq. (29), does not appear possible even for the simplest B -splines. However, by a combination of analytic and numeric integration, the integrals R_{00}^3 and $R_{\pm 2 \pm 2}^3$ for $\Phi_{kl}(x,y) = {}_4\phi_k(x) {}_4\phi_l(y)$ were determined to be

$$R_{00}^3 = -5.811\,966\,397\,897\,0, \quad R_{\pm 2 \pm 2}^3 = 0.055\,945\,658\,26. \tag{A2}$$

These values compare well with those found by solving Eq. (24), which are listed in Table II.

- ¹P. A. Martin and F. J. Rizzo, "On boundary integral equations for crack problems," Proc. R. Soc. London, Ser. A **421**, 341–355 (1989).
- ²T. Terai, "On calculation of sound fields around three-dimensional objects by integral equation methods," J. Sound Vib. **69**, 71–100 (1980).
- ³X. F. Wu, A. D. Pierce, and J. H. Ginsberg, "Variational method for computing surface acoustic pressure on vibrating bodies, applied to transversely oscillating disks," IEEE J. Oceanic Eng. **OE-12**, 412–418 (1987).
- ⁴R. Martinez, "A boundary integral formulation for thin-walled shapes of revolution," J. Acoust. Soc. Am. **87**, 523–531 (1990).
- ⁵F. Cassot, "Contribution à l'étude de la diffraction par un écran mince," Acustica **34**, 64–71 (1975).
- ⁶P. Filippi and G. Dumery, "Etude theorique et numerique de la diffraction par un écran mince," Acustica **21**, 343–350 (1969).
- ⁷P. Filippi, "Layer potentials and acoustic diffraction," J. Sound Vib. **54**, 473–500 (1977).
- ⁸M. Guiggiani, G. Krishnasamy, T. J. Rudolph, and F. J. Rizzo, "A general algorithm for the numerical solution of hypersingular boundary integral equations," J. Appl. Mech.-T. ASME **59**, 604–614 (1992).
- ⁹G. Krishnasamy, L. W. Scherrer, T. J. Rudolph, and F. J. Rizzo, "Hypersingular boundary integral equations: Some applications in acoustic and elastic wave scattering," J. Appl. Mech.-T. ASME **57**, 404–414 (1990).
- ¹⁰T. W. Wu and G. C. Wan, "Numerical modeling of acoustic radiation and scattering from thin bodies using a Cauchy principal integral equation," J. Acoust. Soc. Am. **92**, 2900–2906 (1992).
- ¹¹R. Martinez, "The thin-shape breakdown (TSB) of the Helmholtz integral equation," J. Acoust. Soc. Am. **90**, 2728–2738 (1991).
- ¹²I. Daubechies, "Orthonormal bases of compactly supported wavelets," Commun. Pure Appl. Math. **41**, 909–996 (1988).
- ¹³I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics (SIAM, Philadelphia, 1992), Vol. 61.
- ¹⁴G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Wellesley-Cambridge, Wellesley, Massachusetts, 1996).
- ¹⁵J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Yale University Press, New Haven, Connecticut, 1923).
- ¹⁶H. A. van der Vorst, "BI-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems," SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **13**, 631–644 (1992).

Acoustic response of a periodic layer of nearly rigid spherical inclusions in an elastic solid

Konstantin Maslov^{a)} and Vikram K. Kinra^{b)}

Department of Aerospace Engineering, Texas A&M University, College Station, Texas 77843-3141

(Received 18 September 1998; revised 1 June 1999; accepted 9 August 1999)

Reflection and transmission spectra of a plane longitudinal wave normally incident on a periodic (square) array of identical spherical particles in a polyester matrix are measured at wavelengths which are comparable to the particle radius and the interparticle distance. The spectra are characterized by several resonances whose frequencies are close to the cutoff frequencies for the shear wave diffraction orders. Arrays of heavy particles (lead and steel) exhibit a pronounced resonance which occurs when the lattice resonant frequency is close to the frequency of the rigid-body translation (dipole) resonance of an isolated sphere in an unbounded matrix. An approximate low-frequency theory is developed which assumes that the inclusions are rigid, but which takes into account the multiple-scattering effect. The comparison between theory and the experiment is found to be good for arrays with particle area fraction as high as 32%. © 1999 Acoustical Society of America. [S0001-4966(99)01612-4]

PACS numbers: 43.20.Gp, 43.20.Ks, 43.35.Cg [DEC]

INTRODUCTION

Wave scattering by a periodic array of scatterers has received a good deal of attention in optics, radio physics, and hydroacoustics,¹⁻⁵ where it is widely used in the analysis of diffraction gratings or antenna arrays. More recently, the propagation of an elastic wave through a single coplanar array of either a random or a periodic distribution of elastic inclusions has received some attention.⁶⁻⁸ Although this problem is worth examining in its own right because of its inherent value as a canonical problem in elastodynamics of materials with microstructure, it also has applications in geophysics, quantitative nondestructive evaluation, and in the design of ultrasound absorptive materials.

For the quasi-two-dimensional problem of an array of circular cylinders in an elastic slab, Lakhtakia *et al.*⁹ derived an exact solution for the reflection and transmission of incident longitudinal and shear waves. The presented theoretical and experimental reflection coefficient resembles the back-scattering form factor of the single cylinder modified by the presence of an infinite number of identical cylinders. In addition, it has sharp variations at frequencies where higher diffraction orders become propagative. In the context of diffraction gratings, a similar phenomenon was reported by Wood¹⁰ in 1902, and has been extensively studied in optics.¹

Achenbach and his co-workers have calculated the reflection and transmission of a plane wave by a coplanar array of bounded scatterers in random or periodic arrangements in an elastic medium using a plane space harmonic expansion and the boundary integral equation method. Numerical results were presented for elastic inclusions,^{6,7} cracks,^{11,12} and cavities.^{13,14} For the case of an array of cracks, the results revealed an absence of resonances below the cutoff frequency for the first shear wave diffraction order and a sharp

variation of the reflection coefficient at cutoff frequencies of shear wave diffraction orders. In the case of array of elastic inclusions in an elastic medium, numerical results were presented only for the special case when the density of the matrix and the inclusion are taken to be equal.^{6,7} As expected, in the frequency range from zero to the cutoff frequency of the first shear wave diffraction order, the transmission coefficient was found to vary very little from unity. In a recent article⁸ we have reported experimental transmission spectra of square grids of lead particles. These spectra were found to have a number of strong resonances attributed to the lattice periodicity. It was conjectured that the largest spectral dip was due to the dipole particle resonance modified by the presence of other particles in the array.

In this work, we report the experimental measurements of reflection/transmission of a normally incident plane longitudinal wave from/through a coplanar square periodic array of spherical particles embedded in a polyester matrix for three different inclusion materials: stiff and heavy (steel), stiff but relatively light (glass), and very heavy but not very stiff (lead). Motivated by the experimental results, we develop a simple theoretical model in which the inclusions are assumed to be rigid and small. Only the monopole and dipole scattering terms are retained. In spite of these approximations, excellent comparison is observed for the case of steel and glass inclusions, and a satisfactory comparison is observed for the case of lead inclusions in the frequency range of $k_p a \leq 1$.

I. EXPERIMENTAL PROCEDURES

A schematic of the apparatus is shown in Fig. 1. The specimen is immersed in water between a matched pair of Panametrics broadband piezoelectric transducers with a crystal diameter of 19 mm and a center frequency of 2.25 MHz. A short-duration pulse is applied to the transmitter by a Panametrics pulser/receiver (model 5052UA). The received

^{a)}Electronic mail: maslov@tamu.edu

^{b)}Electronic mail: kinra@tamu.edu

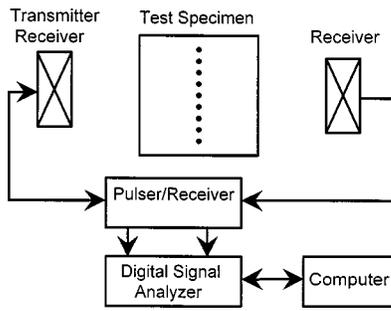


FIG. 1. Schematic of the water-immersion apparatus.

signal is amplified by the pulser/receiver and then digitized by a Tektronix DSA 601 digital analyzer at a sampling interval of 2 ns.

With the specimen removed, the transducers are aligned with respect to each other by maximizing the amplitude of the received signal. Additional fine angular adjustment is made by maximizing the bandwidth of the received signal. Similarly, the alignment of the sample relative to the transducers is accomplished by maximizing the signal reflected from the sample surface. Since the experimental results were found to remain independent of the distance between the transducers and the specimen over a range from 0 (direct contact) to 10 cm, all data reported in this work was collected at a fixed separation of 7 cm between the sample faces and the transducers.

In order to estimate the precision of the experiments, several measurements of each spectrum were made and the mean spectrum and the standard deviation were calculated.¹⁵ The standard deviation was found to be less than 0.05. The transmission coefficient, T , was determined by deconvolving the fast Fourier transform (FFT) of a signal transmitted through a layer of inclusions with respect to the FFT of the signal transmitted through a neat reference specimen (an identical specimen without inclusions). The reflection coefficient, R , was calculated by deconvolving the FFT of a signal reflected from a layer of inclusions with respect to the FFT of the signal reflected from the rear free surface of a reference specimen whose length was one-half the test specimen length.

It should be noted that in addition to the desired reflected or transmitted wave, both the transmitted and reflected signals contain spurious reflections from the specimen/water interfaces. In order to extract the response of a particle layer in an unbounded medium, the signal from the array was isolated by applying a rectangular time window of 25- μ s duration, the time taken by a P wave to complete a roundtrip travel between the specimen face and the particle



FIG. 2. Photograph of a typical periodic specimen containing steel particles with $A_f=0.32$, $d=2.63$ mm, and $a=0.58$ mm.

layer. This limits the spectral resolution of the FFT to about 40 kHz.

Manufacturing procedures for the random and periodic specimens have been described elsewhere in detail⁸ and for brevity will not be repeated here. Cylindrical polyester specimens 52 mm in diameter were manufactured, each containing a layer of identical spheres. All specimens were cut to a length of 7 cm and polished. Specimen nomenclature follows the pattern $AdMa$, where A denotes the layer arrangement (S =square), d is the lattice parameter (interparticle distance) in millimeters, M represents the inclusion material (G =glass, L =lead, S =steel), and a is the inclusion radius in millimeters. Specimens were manufactured with d equal to 1.63, 2.63, and 3.95 mm. The corresponding area fractions, $A_f = \pi a^2/d^2$, are approximately 0.32, 0.16, and 0.07 with a slight difference for glass, lead, and steel particles due to their small difference in particle radius (see Table I). A photograph of the periodic composite S1.86L0.6 is included in Fig. 2.

The acoustic properties of the matrix (mean \pm largest variation between specimens) and the particles are listed in Table I. The acoustic properties of the several neat polyester specimens were measured using the experimental procedure detailed previously.¹⁵ The precision in measuring longitudinal wave speed c_p , the shear wave speed c_s , and the mass density ρ of the matrix material was 0.02%, 0.2%, and 0.1%, respectively. Matrix and inclusion mass densities were measured using Archimedes' principle, and the particle radius was then calculated from the weight measurements. The acoustic properties of the inclusions are taken from Selfridge.¹⁶

As expected, the attenuation, α_p (longitudinal) and α_s (shear) were found to increase linearly with frequency within

TABLE I. Acoustical properties of the constituents.

Material	Radius (mm)	Wave speed (km/s)	Shear wave speed (km/s)	Density (g/cm ³)	Attenuation $\alpha_{p,s}$ @ 1 MHz, (cm ⁻¹)
Polyester	NA	2.49 \pm 0.01	1.18 \pm 0.04	1.22 \pm 0.01	0.17 \pm 0.04 (p) 0.35 \pm 0.1 (s)
Lead	0.6	2.21	0.86	11.3	0.13 (p)
Steel	0.585	5.94	3.2	7.8	Negligible
Glass	0.56	5.66	3.3	2.49	Negligible

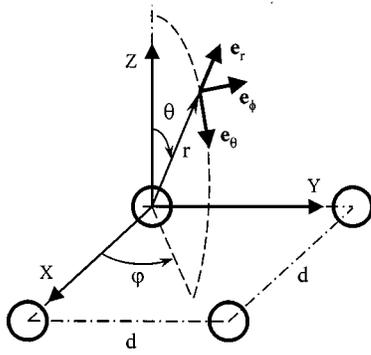


FIG. 3. Spherical coordinate system.

the measured range of 0.2 to 4 MHz. Unfortunately, the variation in α from sample to sample was relatively large, as much as 20%. Therefore, α_p of the composite specimen may be different from that of the reference specimen. As a result, the measured reflection and transmission spectra may vary significantly but systematically with frequency from their true value. However, the locations of the resonant extrema were found to be highly reproducible in spite of this systematic error.

II. RESULTS AND DISCUSSION

The experimentally measured transmission and reflection spectra of the square periodic array of particles in polyester matrix for normally incident plane longitudinal wave are shown in Figs. 6–12. A detailed discussion of the spectra will be presented later. In all figures, one can see a rapid variation of reflection and transmission—certain frequencies which in fact are cutoff frequencies of consecutive shear wave diffraction orders. In addition, a pronounced peak in reflection and a complementary dip in transmission, whose appearance strongly depends on particle mass density, can be seen below first cutoff frequency. An exact theoretical solution of the problem at hand is extremely complex and tedious. However, in the frequency range of the experiment, the particle radius a is much smaller than the acoustic wavelength ($a/\lambda < 0.1$). Therefore, for a normally incident plane longitudinal wave it is possible to derive a simple approximate low-frequency solution retaining only the monopole and dipole scattering terms of the solution of the Helmholtz equation. As we shall see, this solution compares well with the experimental result for $k_p a < 1$. Although the theoretical model was motivated by the experiments, we found that the paper is easier to follow when the analysis is presented before the experimental data.

A. Theoretical model

Consider a plane periodic array of spherical inclusions of infinite extent with the centers of particles located at $x = Pd$, and $y = Qd$, where P and Q are integers. With reference to Fig. 3, let us introduce a spherical coordinate system whose origin is located at the center of a particle. Here, angle of the declination measured from the z -axis is designated by θ , and the azimuthal angle measured from the x -axis is designated by ϕ .

In a spherical coordinate system, the displacement vector \mathbf{u} in the matrix can be decomposed into longitudinal (\mathbf{L}) and transverse (\mathbf{M} and \mathbf{N}) displacement vectors¹⁷

$$\mathbf{u} = \mathbf{L} + \mathbf{M} + \mathbf{N}. \quad (1)$$

Vectors \mathbf{L} , \mathbf{M} , and \mathbf{N} can be expressed in terms of scalar potentials¹⁷

$$\begin{aligned} \mathbf{L} &= \nabla \Phi, \\ \mathbf{M} &= \nabla \times (\mathbf{e}_r r \chi), \\ \mathbf{N} &= \nabla \times \nabla \times (\mathbf{e}_r r \Psi), \end{aligned} \quad (2)$$

where r is the radial coordinate, and \mathbf{e}_r is the unit vector in the radial direction. Scalar potentials Φ , Ψ , and χ satisfy scalar Helmholtz equations

$$(\nabla^2 + k_p^2)\Phi = 0, \quad (\nabla^2 + k_s^2)(\chi, \Psi) = 0, \quad (3)$$

where k_p and k_s are, respectively, the longitudinal and transverse wave numbers in the matrix. The solution of the Helmholtz equation for each potential can be expressed as a series of orthogonal spherical harmonics¹⁷

$$(\Phi, \chi, \Psi) = \sum_{n=0}^{\infty} \sum_{m=0}^n A_{mn}^{(\Phi, \chi, \Psi)} e^{im\phi} P_n^m(\cos \theta) Z_n(kr), \quad (4)$$

where $A_{mn}^{(\Phi, \chi, \Psi)}$ are coefficients of expansion for Φ , χ , and Ψ , respectively, n is the index for the radial function, $Z_n(kr)$, m is the index for the azimuthal function, $e^{im\phi}$, and $P_n^m(\cos \theta)$ is the Legendre polynomial. Radial functions $Z_n(kr)$ are linear combinations of the n th-order spherical Bessel function, $j_n(kr)$, which is bounded at $r \rightarrow 0$ and represents the field incident upon a particle,¹⁷ and the spherical Hankel function of first kind, $h_n(kr)$, which according to its asymptotic behavior at $r \rightarrow \infty$ represents the field scattered by the particle.¹⁷ Common term $\exp(i\omega t)$ is omitted in this and all following equations.

In the low-frequency range, we will neglect terms \mathbf{L}_{mn} , \mathbf{M}_{mn} , and \mathbf{N}_{mn} defined by (2) and (4) for $n \geq 2$. Moreover, for the problem at hand only three of the remaining nine terms are nonzero. From (4) and (2) it easily follows that \mathbf{M}_{00} and \mathbf{N}_{00} are always zero. For a normally incident plane longitudinal wave on a rectangular array, the problem has at least a twofold rotational symmetry with respect to the z -axis, and a reflection symmetry with respect to the x - z mirror plane. The first symmetry makes the three terms with $m = 1$ in (4) vanish. The second symmetry requires that the displacement in the \mathbf{e}_θ -direction should be an odd function of ϕ ; therefore, $\mathbf{M}_{01} = 0$. Accordingly, within the approximation introduced in the foregoing, for the problem at hand, the displacement vector \mathbf{u} can be expressed in terms of two scalar potentials Φ and Ψ (analogous to that for a single particle in an unbounded elastic medium^{18,19})

$$\begin{aligned} \mathbf{u} &= \mathbf{e}_r \left(\frac{\partial \Phi}{\partial r} + \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \Psi \right) \right) \right) \\ &+ \mathbf{e}_\theta \left(\frac{1}{r} \frac{\partial \Phi}{\partial \theta} - \frac{\partial}{\partial \theta} \left(\frac{1}{r} \frac{\partial}{\partial r} (r \Psi) \right) \right), \end{aligned} \quad (5)$$

where

$$\begin{aligned} \Phi \approx & j_0(k_p r) + 3i j_1(k_p r) \cos \theta + A_0 h_0(k_p r) \\ & + A_1 h_1(k_p r) \cos \theta + C_0 j_0(k_p r) + C_1 j_1(k_p r) \cos \theta, \end{aligned} \quad (6)$$

$$\Psi \approx + B_1 h_1(k_s r) \cos \theta + D_1 j_1(k_s r) \cos \theta. \quad (7)$$

In (6) and (7), the Legendre polynomials have been replaced by their explicit expressions: $P_0(\cos \theta) = 1$ and $P_1(\cos \theta) = \cos \theta$, and A_n , B_n , C_n , and D_n are unknown constants. The first two terms in (6) are the first two terms of a spherical harmonic expansion of a plane, time-harmonic, longitudinal incident wave traveling in the positive z -direction, $\Phi^{(i)} = \exp(ik_p z)$. The third and fourth terms in (6) and the first term in (7) represent longitudinal, $\Phi^{(s)}$, and shear, $\Psi^{(s)}$, spherical waves scattered by the particle

$$\begin{aligned} \Phi^{(s)} \approx & A_0 h_0(k_p r) + A_1 h_1(k_p r) \cos \theta, \\ \Psi^{(s)} \approx & B_1 h_1(k_s r) \cos \theta. \end{aligned} \quad (8)$$

The remaining terms represent the longitudinal, $\Phi^{(o)}$, and shear, $\Psi^{(o)}$, waves incident upon the particle under consideration from the remaining particles in the array

$$\begin{aligned} \Phi^{(o)} \approx & C_0 j_0(k_p r) + C_1 j_1(k_p r) \cos \theta, \\ \Psi^{(o)} \approx & D_1 j_1(k_s r) \cos \theta. \end{aligned} \quad (9)$$

Equation (9) contains zero-order and the first-order terms of the spherical harmonic expansion of the sum of scattered fields of all particles in the array, except the particle under consideration. In view of the fact that for a square lattice of infinite extent, the scattered field of every particle is exactly the same, the displacement in the matrix produced by any particle can be found by using (5) and (8) in a coordinate system whose origin is located at the center of that particle. The mathematical expressions for the coefficients of a spherical harmonic expansion of spherical waves can be found in Morse and Feshbach²⁰ and Tversky.^{3,4} The summation of these coefficients over all particles relates constants C_n and D_n to A_n and B_n . The relation between C_n and A_n can also be found directly using translational invariance of the gradient operator in (2) and orthogonality of spherical harmonics in (9). In the following, we will derive these relations using simple physical considerations. As it follows from (9) and (5), the first order terms in (9) define uniform displacement in the z -direction on a spherical surface with its center at $r = 0$. As $r \rightarrow 0$, these become

$$\mathbf{u}_z^{(p)} = (1/3)k_p C_1, \quad \mathbf{u}_z^{(s)} = -(2/3)k_s D_1. \quad (10)$$

These displacements are, respectively, the sums at a point $r = 0$ of the displacement vectors in the longitudinal and shear waves scattered by all particles. From (8) and (5), these are

$$\mathbf{u}_z^{(p)} = k_p A_1 \sum_q \left(\frac{h_1(k_p r_q)}{k_p r_q} \right) = k_p A_1 S_1^{(p)}(k_p d), \quad (11)$$

$$\mathbf{u}_z^{(s)} = -k_s B_1 \sum_q \left(h_0(k_s r_q) - \frac{h_1(k_s r_q)}{k_s r_q} \right) = -k_s B_1 S_1^{(s)}(k_s d),$$

where r_q is the distance to the q th particle, and summation is over all particles in the lattice except the particle under con-

sideration. Combining (10) and (11) one can relate C_1 to A_1 , and D_1 to B_1 as follows:

$$C_1 = 3A_1 S_1^{(p)}(k_p d), \quad D_1 = (3/2)B_1 S_1^{(s)}(k_s d). \quad (12)$$

The zero-order term in (9) defines uniform radial displacement on a spherical surface, which in the limit $r \rightarrow 0$, becomes equal to the dilatation component, $\nabla \times \mathbf{u}$, of the waves incident on the particle from all the remaining particles in the array. A straightforward calculation using (5), (8), and (9) gives

$$C_0 = A_0 \sum_q h_0(k_p r_q) \equiv A_0 S_0^{(p)}(k_p d). \quad (13)$$

Equation (12) accounts for the dipole–dipole interaction of the particles in the array, whereas Eq. (13) accounts for the monopole–monopole interaction. The sums in (11) and (13) are the so-called lattice sums^{3–5} that depend only on the lattice geometry, and are independent of the acoustical properties of the inclusions. For a square array, using the properties of the Hankel functions, the lattice sums $S_0^{(p)}(k_p d)$, $S_1^{(p)}(k_p d)$, and $S_1^{(s)}(k_s d)$ are given by

$$\begin{aligned} S_0^{(p)}(k_p d) &= \sum_q h_0(k_p r_q) \\ &= 4 \sum_{l=1, m=0}^{\infty} \frac{\exp(ik_p d \sqrt{l^2 + m^2})}{ik_p d \sqrt{l^2 + m^2}}, \\ S_1^{(p)}(k_p d) &= -4 \sum_{l=1, m=0}^{\infty} \frac{\exp(ik_p d \sqrt{l^2 + m^2})}{(k_p d)^2 (l^2 + m^2)} \\ &\quad \times \left[1 - \frac{1}{ik_p d \sqrt{l^2 + m^2}} \right], \\ S_1^{(s)}(k_s d) &= 4 \sum_{l=1, m=0}^{\infty} \frac{\exp(ik_s d \sqrt{l^2 + m^2})}{ik_s d \sqrt{l^2 + m^2}} \\ &\quad \times \left[1 - \frac{1}{ik_s d \sqrt{l^2 + m^2}} - \frac{1}{(k_s d)^2 (l^2 + m^2)} \right]. \end{aligned} \quad (14)$$

Since we were unable to find a closed-form solution for these lattice sums, these had to be calculated numerically. Boundary conditions at the surface of the particle give the remaining equations necessary to find the six unknown constants in (6) and (7).

The discussion so far applies to any spherical inclusion. We now assume that the inclusion is *rigid*. As we shall see, it turns out to be a reasonable assumption for stiff inclusions (glass and steel) in a compliant matrix (polyester) tested in this investigation. It also works sufficiently well for lead particles. In addition to the considerable simplification of the model, this assumption assigns a direct physical sense to the translational displacement of a particle; this we will exploit in the following. At the boundary of a rigid sphere, i.e., at $r = a$, the monopole term of the displacement vector in the matrix is equal to zero. The dipole components of the displacement vector at $r = a$ are related to the rigid-body trans-

lation of a spherical particle, U . Due to symmetry, the only nonzero component of U is in the z -direction. It follows that

$$u_r = U \cos(\theta), \quad u_\theta = -U \sin(\theta), \quad (15)$$

where U is governed by Newton's second law,

$$\frac{-4\pi a^3 \rho_i}{3} U \omega^2 = \iint (\tau_{rr} \cos \theta - \tau_{r\theta} \sin \theta) \times a^2 \sin \theta d\theta d\varphi|_{r=a}. \quad (16)$$

The surface integral in (16) is performed over the entire particle surface, ρ_i is the particle mass density, and the tractions, τ_{rr} and $\tau_{r\theta}$, are given by¹⁹

$$\begin{aligned} \tau_{rr} &= 2\mu \left\{ -\frac{k_T^2}{2} \Phi - \frac{2}{r} \frac{\partial \Phi}{\partial r} - \frac{1}{r^2} D_\theta \Phi + \frac{D_\theta}{r} \left(\frac{\partial \Psi}{\partial r} - \frac{\Psi}{r} \right) \right\}, \\ \tau_{r\theta} &= 2\mu \frac{\partial}{\partial \theta} \left\{ \frac{2}{r} \frac{\partial \Phi}{\partial r} - \frac{1}{r^2} \Phi + k_T^2 \Psi + \frac{2}{r} \frac{\partial \Psi}{\partial r} + \frac{2}{r^2} \Psi \right. \\ &\quad \left. + \frac{2}{r^2} D_\theta \Psi \right\}, \quad D_\theta = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right). \end{aligned} \quad (17)$$

Substituting (12) and (13) into (6) and (7), then (6) and (7) into (17), then (17) into (16), and solving (16) and (15), we find

$$\begin{aligned} A_0 &= -\frac{j_1(k_p a)}{h_1(k_p a) + S_0^{(p)}(k_p d) j_1(k_p a)}, \\ A_1 &= -3i \frac{(9\bar{\rho} E_{10} - (\bar{\rho} + 2)) j_1(k_p a) - ((2\bar{\rho} + 1) E_{10} - 1) k_p a j_0(k_p a)}{(9\bar{\rho} E_{10} - (\bar{\rho} + 2)) E_1 - ((2\bar{\rho} + 1) E_{10} - 1) E_0}, \end{aligned} \quad (18)$$

where $\bar{\rho} = \rho/\rho_i$ is the ratio of the mass density of the matrix to that of a particle, and

$$\begin{aligned} E_0 &= k_p a (3S_1^{(p)}(k_p d) j_0(k_p a) + h_0(k_p a)), \\ E_1 &= 3S_1^{(p)}(k_p d) j_1(k_p a) + h_1(k_p a), \\ E_{10} &= \frac{3S_1^{(s)}(k_s d) j_1(k_s a) + 2h_1(k_s a)}{k_s a (3S_1^{(s)}(k_s d) j_0(k_s a) + 2h_0(k_s a))}. \end{aligned} \quad (19)$$

For a normally incident plane wave, particles in a square periodic array have an exactly identical scattering field. Therefore, the amplitude of a specularly reflected plane wave, $\Phi_{(z)}$, can be calculated as an integral of the field scattered by one particle over a remote plane surface parallel to the array of particles, multiplied by the number of particles per unit area in the lattice.²¹ Recalling that $h_0(x) = \exp(ix)/(ix)$ and $h_1(x) = -\partial h_0(x)/\partial x$ from Eq. (8), we have

$$\begin{aligned} \Phi_{(z)} &\approx \frac{1}{d^2} \iint_s (A_0 h_0(k_p r) + A_1 h_1(k_p r) \cos \theta) \\ &\quad \times \exp(-ik_p z) ds \\ &= \frac{2\pi}{d^2} \int_{z_0}^\infty \left\{ A_0 \frac{\exp(ik_p r)}{ik_p r} + A_1 \frac{-\partial h_0(k_p r)}{k_p \partial r} - \frac{z}{r} \right\} r dr \\ &= \frac{2\pi}{d^2 k_p^2} (A_0 + iA_1) \exp(-ik_p z). \end{aligned} \quad (20)$$

The ratio $\Phi_{(z)}/\Phi^{(l)}$ at ($z=0$) is the plane wave reflection coefficient, $R(\omega)$, of a layer at normal incidence. The transmission coefficient, $T(\omega)$, can be found in a similar way by including the propagating incident field. Together with Eqs. (18) and (19),

$$R(\omega) \approx \frac{2\pi(A_0 + iA_1)}{d^2 k_p^2}, \quad T(\omega) \approx 1 + \frac{2\pi(A_0 - iA_1)}{d^2 k_p^2}, \quad (21)$$

provide an explicit solution for the reflection and transmission coefficients of a square array of rigid spheres in an elastic matrix insonified by a plane longitudinal wave at normal incidence. In the Rayleigh limit as $k_p d \rightarrow 0$, Eqs. (18) and (19) reduce to that for the single particle scattering

$$A_0 = -\frac{i(k_p a)^3}{3}, \quad A_1 = \frac{(\bar{\rho} - 1)(k_p a)^3}{3\bar{\rho}}, \quad (22)$$

and the equation for the reflection coefficient of the layer reduces to its familiar form²²

$$R = -i \frac{k_L V_f \rho_i}{2\rho}, \quad (23)$$

where $V_f = (4/3)\pi a^3/d^2$ is the volume fraction of particles in a layer.

It can be shown by setting all lattice sums—with the exception of $S_1^{(s)}(k_s d)$ —equal to zero, that the salient features of the reflection and transmission spectra (for example, the position of the maxima and minima) almost entirely depend on shear wave interaction between the particles. It is a natural consequence of the well-known fact that for a longitudinal incident wave, at $k_p a < 1$, the amplitude of the shear wave scattered in any transverse direction (i.e., in the x - y plane) is significantly larger than that of the longitudinal wave scattered in any transverse direction.²³

In Fig. 4 the magnitude and phase of the lattice sum $S_1^{(s)}(k_s d)$ are presented versus the dimensionless frequency $\Omega = k_s d/2\pi$. Note that due to a constructive interference of scattered waves, the lattice sum has sharp maxima at lattice resonance frequencies $\Omega = 1, \sqrt{2}, 2, \sqrt{5}, \sqrt{8}, \dots$. These frequencies are the cutoff frequencies for the appropriate shear wave

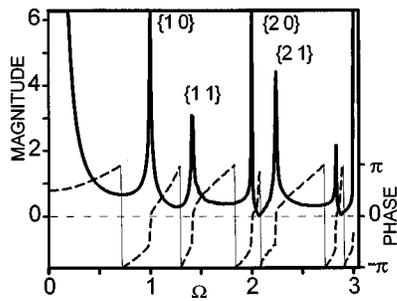


FIG. 4. Magnitude and phase of the grid sum $S_1^{(s)}(k_s d)$ versus dimensionless frequency Ω .

diffraction orders of the periodic array of particles. At these frequencies, shear waves can propagate in the lattice plane.²⁴ Due to the translational symmetry of the lattice, these are Floquet waves with the x - y components of the wave vector \mathbf{k} constrained by

$$k_x d = 2\pi P, \quad k_y d = 2\pi Q, \quad \text{where } P, Q = 0, \pm 1, \dots \quad (24)$$

The dimensionless frequency for the $[P \ Q]$ transverse wave mode is given by

$$\Omega = \sqrt{P^2 + Q^2}. \quad (25)$$

The integer pairs $[P \ Q]$ play the role of Miller indices for the square array. The lowest resonance frequency, $[1 \ 0]$, corresponds to a shear wave in the $[1 \ 0]$ direction. The group of equivalent due to the symmetry of a square lattice directions P, Q will be referred to as $\{P \ Q\}$.

Referring to (19), the rigid-body translation of a particle normalized by the displacement of the incident plane longitudinal wave may be derived as

$$U_N \approx -i \frac{\bar{\rho}}{(k_p a)^2} \times \frac{3(3E_{10} - 1)}{(9\bar{\rho}E_{10} - (\bar{\rho} + 2))E_1 - ((2\bar{\rho} + 1)E_{10} - 1)E_0}. \quad (26)$$

This equation serves as a measure of the resonance effects (if any) of an array. In effect, it is the amplification of the particle displacement with respect to that of the incident wave. The magnitude of U_N for periodic arrays of steel particles used in the experiments ($a = 0.58 \text{ mm}, \rho = 7.8 \text{ g/cm}^3$) is

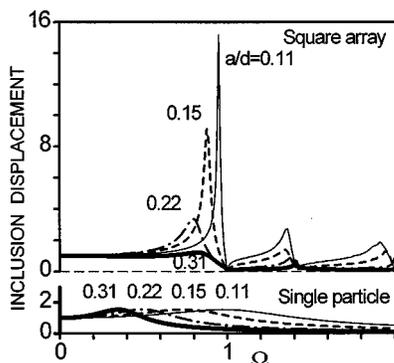


FIG. 5. Normalized translational particle displacement in a square coplanar array of steel particles insonified by a normally incident plane longitudinal wave.

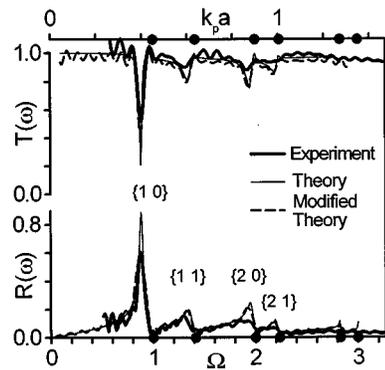


FIG. 6. Transmission and reflection coefficients of the steel specimen with $a/d = 0.15, A_f = 0.07$, and $\bar{\rho} = 0.16$ (S3.95S0.58). Maximum frequency corresponds to 1 MHz.

shown in Fig. 5 for interparticle distances, d , of 1.86, 2.63, 3.95 and 5.26 mm. Corresponding dimensionless ratios a/d are 0.31, 0.22, 0.15, and 0.11, respectively. For comparison purposes, the normalized amplitude of the displacement of an isolated particle drawn to the same scale for the same ratios a/d is also presented in Fig. 5. For the isolated steel particle, there is a peak in the displacement amplitude at $k_p a \sim 0.3$. This is the so-called translational rigid-body resonance (RBR) of a single particle in an unbounded elastic medium.²⁵ One can see that as the RBR frequency approaches the lattice resonance frequency, the vibration amplitude of a particle in an array increases monotonically, and at $a/d = 0.11$ it is one order of magnitude larger than the displacement of the isolated particle, and much larger than that of the incident wave. Normalized particle displacement was also calculated for arrays of lead and glass particles in square and hexagonal arrangements. The results are similar to those presented in Fig. 5, and for brevity are not included.

B. Experimental results

In the following, we present results in the increasing order of area fractions from a rather dilute suspension ($A_f = 0.07$) to a fairly concentrated mixture ($A_f = 0.32$). Reflection and transmission coefficients for the steel specimen (S3.95S0.58) are presented in Fig. 6; here, $a/d = 0.15, A_f = 0.07$, and $\bar{\rho} = 0.16$. The lattice resonance frequencies are

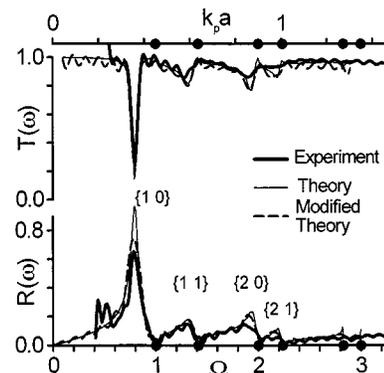


FIG. 7. Transmission and reflection coefficients of the lead specimen with $a/d = 0.15, A_f = 0.07$, and $\bar{\rho} = 0.11$ (S3.95L0.6). Maximum frequency corresponds to 1 MHz.

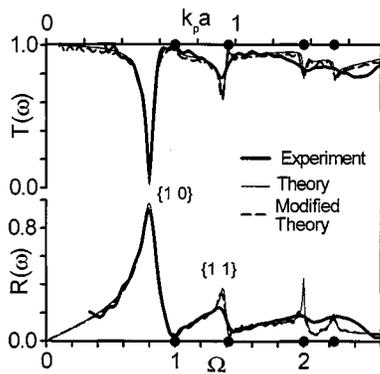


FIG. 8. Transmission and reflection coefficients of the steel specimen with $a/d=0.23$, $A_f=0.16$, and $\bar{\rho}=0.16$ (S2.63S0.58). Maximum frequency corresponds to 1.2 MHz.

marked by filled circles in Figs. 6–12. Below each lattice resonance frequency there is a peak in $R(\omega)$ and a corresponding dip in $T(\omega)$. With reference to Fig. 5, the frequency at which the particle displacement reaches a maximum corresponds fairly closely to these frequencies. Because this peak is by far the most dominant resonance, we will refer to it as the “primary resonance.” Good comparison between the experiment and the theory is evident up to $\Omega \sim 2$.

The amplitude of the primary resonance peak of $R(\omega)$ in the experimental spectrum is less than that in the theoretical spectrum. As was noted in the description of the experimental procedures, the frequency resolution of the measurement procedure is about 40 kHz, which in some cases is less than the width of the theoretical resonance peak. It follows that for resonance peaks narrower than 40 kHz, the experiment can yield peaks of magnitude lower than predicted. Since the measured spectrum is the true spectrum of a layer convolved with the Fourier transform of the rectangular time gate, it is more appropriate to compare the measured spectrum with the convolution of the theoretical spectrum with the Fourier transform of the 25- μ s rectangular window used in the experiment. Theoretical spectra, modified in this manner, are shown as dashed lines in Fig. 6 and in the following figures. One can see that this modification results in a much better comparison between theory and experiment.

Reflection and transmission spectra for a lead specimen

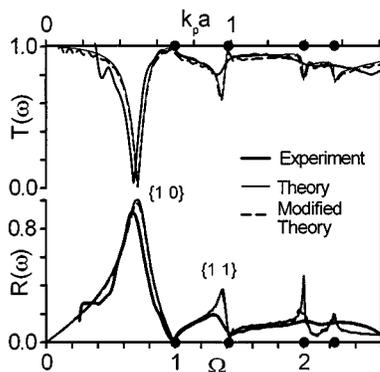


FIG. 9. Transmission and reflection coefficients of the lead specimen with $a/d=0.22$, $A_f=0.16$, and $\bar{\rho}=0.11$ (S2.63L0.6). Maximum frequency corresponds to 1.2 MHz.

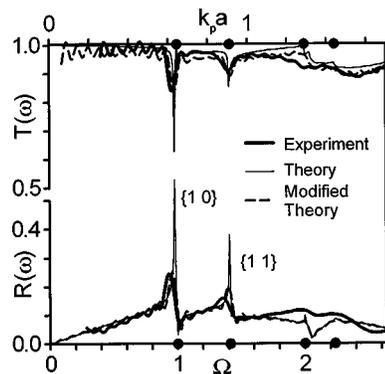


FIG. 10. Transmission and reflection coefficients of the glass specimen with $a/d=0.21$, $A_f=0.14$, and $\bar{\rho}=0.49$ (S2.63G0.56). Maximum frequency corresponds to 1.2 MHz.

(S3.95L0.6) are presented in Fig. 7; here again $a/d=0.15$, $A_f=0.07$, but $\bar{\rho}=0.11$. In this case ($\rho=11.3 \text{ g/cm}^3$) the primary resonance occurs at $\Omega \sim 0.8$, which compares with $\Omega \sim 0.9$ for steel particles ($\rho=7.8 \text{ g/cm}^3$).

Reflection and transmission spectra for the steel, lead, and glass specimens (S2.63S0.58, S2.63L0.6, and S2.63G0.56; $A_f \sim 0.16$, $a/d \sim 0.22$) are presented in Figs. 8, 9, and 10, respectively. The primary resonance occurs at $\Omega = 0.67$, $\Omega = 0.8$, and $\Omega \sim 1$, respectively. At primary resonance, reflection from the layer of heavy steel and lead particles, both in theory and experiment, is nearly unity. The depth of the corresponding transmission dip has a magnitude of about -40 dB . This may find application in the design of selective filters for elastic waves. Good comparison between experimental and theoretical data in the range where $\Omega < 2$ is evident for the steel and the glass samples. On the other hand, in the case of the lead sample, there is a slight discrepancy between the theoretically predicted ($\Omega = 0.71$) and the measured ($\Omega = 0.67$) frequency of the primary resonance. This shift may be attributed to the fact that since the longitudinal and shear wave velocity in lead are less than the corresponding velocities in the matrix, the rigid inclusion assumption may not be justifiable.

For the highest area fraction studied ($A_f \sim 0.3$, $a/d = 0.31$) spectra for the steel (S1.86S0.58, $\bar{\rho}=0.16$) and glass (S1.86G0.56, $\bar{\rho}=0.49$) specimens are plotted in Figs. 11 and 12, respectively. For this relatively high-area fraction, $k_L a$

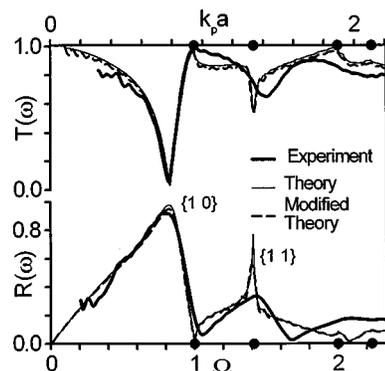


FIG. 11. Transmission and reflection coefficients of the steel specimen with $a/d=0.31$, $A_f=0.32$, and $\bar{\rho}=0.16$ (S1.86S0.58). Maximum frequency corresponds to 1.5 MHz.

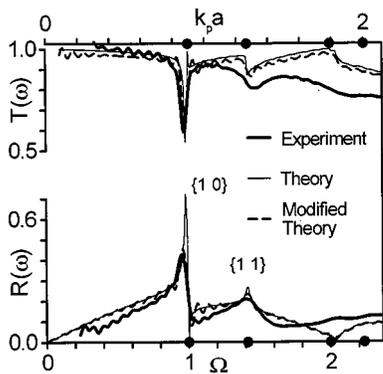


FIG. 12. Transmission and reflection coefficients of the glass specimen with $a/d=0.30$, $A_f=0.28$, and $\bar{\rho}=0.49$ (S1.86G0.56). Maximum frequency corresponds to 1.5 MHz.

of the order of 1 at $\Omega \sim 1$; hence, in the spherical harmonic expansion of the scattered field, terms which are higher than the dipole term ($n=1$) should be taken into account. Nevertheless, there is a good agreement between theory and experiment for $\Omega < 1$ for both the heavy (steel, $\bar{\rho}=0.16$) and light (glass, $\bar{\rho}=0.49$) particles, i.e., the model works outside the range of the underlying assumptions. Finally, the good comparison between theory and experiment suggests that for the particular arrays studied, at $\Omega < 1$ our very simple model captures the essential physics of the interparticle interaction.

III. SUMMARY

Transmission and reflection spectra of a normally incident plane longitudinal wave for a coplanar periodic array of spherical particles in a polyester matrix exhibit a series of lattice resonances (Wood's anomalies). The dipole-dipole interaction between particles in shear is responsible for the uncommonly large magnitude of resonance, Wood's anomaly, which occurs when the frequency of the rigid-body (dipole) resonance of an isolated sphere in an unbounded matrix is close to the lattice resonance frequency. At resonance, an almost complete reflection, and -40 -dB transmission occurs in a sparse ($A_f=0.16$) array of steel particles. Therefore, a rather thin ($k_p a \sim 0.5$) layer of particles can serve as a narrow-band reject acoustic filter. This effect may find application in the acoustic filter design, for example, when the second harmonic of a carrier frequency needs to be selectively filtered out of a signal.

¹A. Hessel and A. Oliner, "A new theory of Wood's anomalies on optical gratings," *Appl. Opt.* **4**, 1275–97 (1965).

²V. Twersky, "On scattering of waves by the infinite grating of circular cylinders," *IRE Trans. Antennas Propag.* **AP-10**, 737–65 (1962).

- ³V. Twersky, "Multiple scattering of plane waves by the doubly periodic array of obstacles," *J. Math. Phys.* **16**, 633–643 (1975).
- ⁴V. Twersky, "Lattice sums and scattering coefficients for the rectangular planar array," *J. Math. Phys.* **16**, 644–657 (1975).
- ⁵I. Tolstoy and A. Tolstoy, "Line and plane arrays of resonant monopole scatterers," *J. Acoust. Soc. Am.* **87**, 1038–1043 (1990).
- ⁶M. Kitahara and J. D. Achenbach, "BIE calculations for harmonic waves in solid with periodic distributed inhomogeneities," in *Pros., IUTAM Symp., Advanced Boundary Integral Methods*, edited by Th. A. Crause (Springer, Berlin, 1988), pp. 197–205.
- ⁷J. D. Achenbach, M. Kitahara, Y. Mikata, and D. A. Sotiropoulos, "Reflection and transmission of plane waves by a layer of compact inhomogeneities," *PAGEOPH* **128**, 101–118 (1988).
- ⁸V. K. Kinra, N. A. Day, K. Maslov, B. K. Henderson, and G. Diderich, "The transmission of a longitudinal wave through a layer of spherical inclusions with a random or periodic arrangement," *J. Mech. Phys. Solids* **46**, 153–165 (1998).
- ⁹A. Lakhtakia, V. V. Varadan, and V. K. Varadan, "Reflection characteristics of an elastic slab containing a periodic array of circular elastic cylinders: P and SV wave analysis," *J. Acoust. Soc. Am.* **83**, 1267–1275 (1988).
- ¹⁰R. W. Wood, "On a remarkable case of uneven distribution of light in a diffraction grating spectrum," *Philos. Mag.* **4**, 396–402 (1902).
- ¹¹D. A. Sotiropoulos and J. D. Achenbach, "Ultrasonic reflection by a planar distribution of cracks," *J. Nondestruct. Eval.* **7**, 123–129 (1988).
- ¹²Y. C. Angel and J. D. Achenbach, "Reflection and transmission of elastic waves by a periodic array of cracks," *J. Appl. Mech.* **52**, 33–41 (1985).
- ¹³M. Kitahara and J. D. Achenbach, "BIE method to analyze wave motion in solids with periodically distributed cavities," *Comput. Methods Appl. Mech. Eng.* **64**, 523–36 (1987).
- ¹⁴J. D. Achenbach and M. Kitahara, "Reflection and transmission of an obliquely incident wave by an array of spherical cavities," *J. Acoust. Soc. Am.* **80**, 1209–14 (1986).
- ¹⁵V. K. Kinra and V. R. Iyer, "Ultrasonic measurement of the thickness, phase velocity, density or attenuation of a thin viscoelastic plate. Part I: The forward problem," *Ultrasonics* **33**, 95–109 (1995).
- ¹⁶A. R. Selfridge "Approximate material properties in isotropic materials," *IEEE Trans. Sonics Ultrason.* **3**, 381–94 (1985).
- ¹⁷N. G. Einspruch, E. J. Witterholt, and R. Truell, "Scattering of plane transverse wave by a spherical obstacle in an elastic medium," *J. Appl. Phys.* **31**, 806–818 (1960).
- ¹⁸C. F. Ying and R. Truell, "Scattering of a plane longitudinal wave by a spherical obstacle in an isotropically elastic solid," *J. Acoust. Soc. Am.* **27**, 1086–1097 (1956).
- ¹⁹Y.-H. Pao and C. C. Mow, "Scattering of a plane compressional waves by a spherical obstacle," *J. Appl. Phys.* **34**, 493–499 (1963).
- ²⁰P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part II* (McGraw-Hill, New York, 1953), p. 1875.
- ²¹L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1975), Chap. 4.
- ²²V. Twersky, "On the nonspecular reflection of plane waves of sound," *J. Acoust. Soc. Am.* **22**, 539–546 (1950).
- ²³V. V. Varadan, Y. Ma, V. K. Varadan, and A. Lakhtakia, "Scattering of waves by spheres and cylinders," in *Field Representation and Introduction to Scattering*, edited by V. V. Varadan, A. Lakhtakia, and V. K. Varadan (Elsevier, Amsterdam, 1991), pp. 211–324.
- ²⁴K. Maslov, V. K. Kinra, and B. K. Henderson, "Lattice resonances of a planar array of spherical inclusions: An experimental study," *Mech. Mater.* **31**, 175–186 (1999).
- ²⁵A. Wolf, "Motion of a rigid sphere in an acoustic wave field," *Geophysics* **10**, 91–109 (1945).

Effect on ultrasonic signals of viscous pore fluids in unconsolidated sand

Patricia K. Seifert,^{a)} Bruno Kaelin,^{b)} and Lane R. Johnson

Department of Geology and Geophysics, University of California, and Center for Computational Seismology, Lawrence Berkeley National Laboratory, Berkeley, California 94720

(Received 18 May 1998; revised 10 July 1999; accepted 24 August 1999)

Ultrasonic attenuation measurements in unconsolidated sand with pore fluids ranging in viscosity between 0.001 and 1 Pa·s were compared with the predictions of fluid flow and scattering theories. Laboratory experiments were performed for *P* waves propagating through sand samples saturated with water, castor oil and two different silicone oils. The attenuation shows a frequency squared dependence for all measurements, regardless of viscosity, in the range between 100 and 1000 kHz. The results show that for unconsolidated sand, fluid flow models which imply significant effects of the viscous pore fluids on ultrasonic waves cannot explain the laboratory measurements. The main attenuation effects observed in the laboratory can be simulated with a three-dimensional generalized dynamic composite elastic medium model, which includes scattering from the pores and grains as well as intrinsic attenuation caused by the viscous pore fluids. For the studied sand samples, scattering is the main attenuation mechanism for ultrasonic *P* waves. © 1999 Acoustical Society of America. [S0001-4966(99)03612-7]

PACS numbers: 43.20.Jr, 43.35.Cg, 43.35.Fj, 43.35.Mr [DC]

INTRODUCTION

Over the last three decades, enormous strides have been made in understanding the connections between physical properties of rocks and elastic wave propagation. In particular, the dependence of attenuation on material properties of different fluids and their interaction with the matrix has been the topic of various studies.¹⁻⁵

The first major breakthrough in predicting the elastic moduli of porous media at low frequencies was achieved by Gassmann.⁶ Gassmann's equations relate elastic moduli of fluid saturated rocks to the properties of the dry frame and the fluid and are still widely used, but they provide little insight into the physics of wave propagation. Biot developed a theory of wave propagation in fluid saturated porous media^{7,8} that focuses on macroscopic fluid flow. Biot's theory shows that acoustic waves create relative motion between the fluid and the solid matrix due to inertial effects, resulting in viscous dissipation of acoustic energy. However, at seismic frequencies the predicted attenuation is usually small compared to measured effects.¹ Another fluid flow mechanism, often called "local-flow" or "squirt-flow," is based upon microscopic fluid motion between pores and cracks and within cracks.⁹ The combination of Biot's theory and local-flow models can usually be fit to experimental data,^{10,11} but this approach does not have the predictive power of Gassmann's or Biot's theory because it is highly dependent on details of the microstructure. A different approach to the study of elastic waves in porous media is to calculate the elastic scattering from pores and grains. Although scattered energy is not absorbed by the rock as heat, it is similar to intrinsic attenuation in that energy is lost from the primary pulse.¹²⁻¹⁵

This study examines the role of pore fluids in determining attenuation at ultrasonic frequencies by comparing experimental laboratory data with fluid flow theories and scattering theories. The major difference between these two theories lies in the importance of the fluid viscosity. While the fluid flow theories depend strongly upon the viscosity, it has little impact on scattering theory at the frequencies under consideration. We conducted *P* wave measurements through unconsolidated quartz sand samples saturated with different pore fluids ranging in viscosity from 0.001 to 1 Pa·s. After describing the experimental method, the measured data are analyzed in terms of physical models of elastic wave propagation in unconsolidated sand at ultrasonic frequencies.

I. EXPERIMENTAL PROCEDURE AND RESULTS

We measured the attenuation of ultrasonic *P* waves through an unconsolidated quartz sand in order to determine the effect of different fluid viscosities. For the laboratory measurements, a subrounded quartz sand sieved between 60 and 70 mesh (210 and 250 μm openings, respectively) was chosen. The four different pore fluids were water, two different silicone oils and castor oil, which had a three decade variation in viscosity ranging from 0.001 to 1 Pa·s (Table I). To saturate the specimens, the fluids were first poured onto the transducer, which was jacketed by a flexible Teflon sleeve. Dissolved air was extracted from the fluids by a vacuum pump before the washed and dried sand was carefully poured into the fluid to avoid trapped air in the sample. Finally, the specimens were packed to a porosity of about 35%. The sand samples were cylindrical with a diameter of 50.5 mm and an average length of 36.5 mm. They were placed between a transmitter and receiver in a triaxial cell where confining and axial stresses on the samples were

^{a)}Now at McKinsey and Company, Munich, Germany.

^{b)}Now at Swiss Federal Office of Metrology, Bern-Wabern, Switzerland.

TABLE I. P wave velocity v_p , density ρ and dynamic viscosity η for different fluids after Selfridge (Ref. 26), if not otherwise noted.

Pore fluid	v_p (m/s)	ρ (kg/m ³)	η (Pa·s)
Water	1497	998	0.001
Oil, silicone10 Dow 200 (cs10)	968	940	0.01
Oil, silicone100 Dow 200 (cs100)	980	968	0.1
Oil, castor	1507 ^a	942 ^a	1.0

^aMeasured properties.

maintained at 70 and 140 kPa, respectively. More information about the geometry of the triaxial cell can be found in Geller and Myer.¹⁶

We used the pulse transmission technique to measure travel time and amplitude of P waves propagating through the samples fully saturated with each of the fluids. Ultrasonic pulses had a passband between 100 and 2000 kHz with a central frequency of approximately 900 kHz. Sample dimensions and porosities for the samples saturated with different fluids are listed in Table II. Figure 1(a) shows the waveforms for the sand samples saturated with the four different pore fluids. The amplitudes have been normalized by the first arrival.

The spectral ratio method^{17,18} was used to calculate the frequency dependent attenuation and the low frequency P wave and S wave velocity. As the reference standard we chose water with the same length as the samples to avoid geometrical spreading effects.¹⁹ The spectrum U_{ref} of the reference standard was calculated with half a cycle of the first arriving pulse

$$U_{\text{ref}} = A_{\text{ref}} \exp(i\phi_{\text{ref}}) = A_{\text{ref}} \exp\left(-i \frac{\omega}{V_{\text{ref}}} L_{\text{ref}}\right), \quad (1)$$

where A_{ref} is the amplitude spectrum, ϕ_{ref} the phase spectrum, ω the angular frequency, V_{ref} the phase velocity and L_{ref} the length of the reference standard. Similarly, the spectrum U of the P wave and S wave was calculated with half a cycle of the first arriving pulse

$$U = A \exp(i\phi) = A_{\text{ref}} \exp(-\alpha L) \exp\left(-i \frac{\omega}{V} L\right), \quad (2)$$

where A is the amplitude spectrum, ϕ the phase spectrum, α the attenuation, V the phase velocity and L the length of the sample. The spectral ratio of Eqs. (1) and (2) is

TABLE II. Sample length, porosity and measured low frequency P wave velocity and matrix S wave velocity for the different samples of saturated sand. The S wave velocity of the samples and matrix P wave velocity were calculated with Berryman's equations (Ref. 20). All samples were 50.5 mm in diameter.

Saturated sand	L (mm)	ϕ (%)	v_p^a (m/s)	v_s^b (m/s)	Matrix v_p^b	Matrix v_s^a
Water	37.0	36.4	1825	694	2522	1029
cs10 oil	36.5	35.5	1310	511	2093	742
cs100 oil	36.5	35.5	1346	501	2322	725
Castor oil	35.5	33.7	1802	719	2335	1012

^aMeasured properties.

^bCalculated properties.

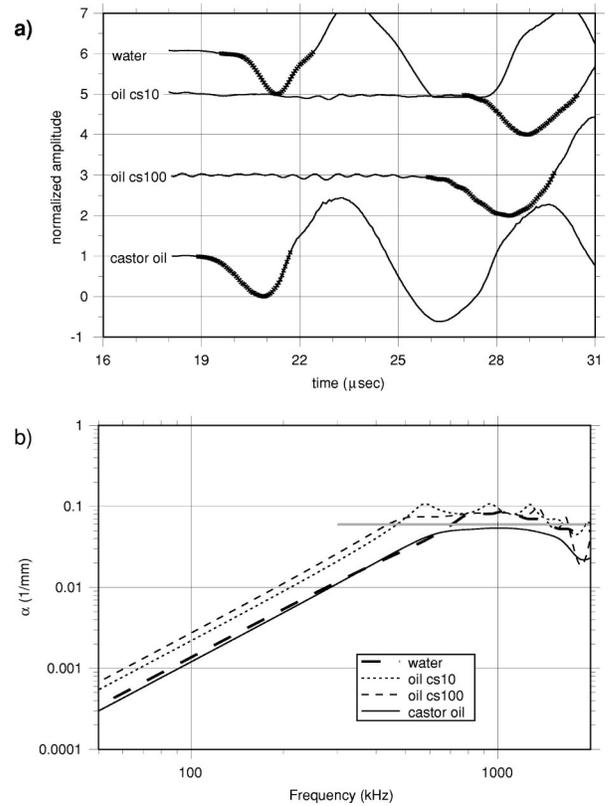


FIG. 1. Panel (a) shows the measured P wave traces for sand samples saturated with the pore fluids water, cs10, cs100 and castor oil. The highlighted time windows are those used to calculate the frequency dependent attenuation which is shown in panel (b). Attenuation is considered significant only if $\alpha < 0.6 \text{ mm}^{-1}$, which is shown as a gray line.

$$\frac{U}{U_{\text{ref}}} = \frac{A}{A_{\text{ref}}} \exp[i(\phi - \phi_{\text{ref}})] \quad (3)$$

or

$$\frac{U}{U_{\text{ref}}} = \exp(-\alpha L) \exp\left[-i\omega\left(\frac{L}{V} - \frac{L_{\text{ref}}}{V_{\text{ref}}}\right)\right]. \quad (4)$$

Comparing Eqs. (3) and (4), the attenuation is

$$\alpha = -\frac{1}{L} \ln \frac{A}{A_{\text{ref}}}, \quad (5)$$

and the phase velocity is

$$V = \left[\frac{L_{\text{ref}}}{L} \frac{1}{V_{\text{ref}}} - \frac{\phi - \phi_{\text{ref}}}{L\omega} \right]^{-1}. \quad (6)$$

In Eq. (5) we have assumed that the attenuation of water is negligible. The phase velocity obtained in Eq. (6) is not equivalent to velocities determined from the first break of the signal, which are higher, but it is consistent with the low frequency velocity derived by composite elastic medium theories.^{15,20,21} The measured P wave velocities in the low frequency limit for the different samples are listed in Table II. The P wave velocities are slightly different from the water reference, which could cause geometrical spreading effects, but because we chose the same length for the samples and the reference, this effect is actually very small.¹⁹ Also, the near source effects are not significant because the wave-

lengths are much smaller than the sample length at the 900 kHz central frequency.

At low frequencies the attenuation shows a frequency squared dependence regardless of the pore filling fluids [Fig. 1(b)]. At higher frequencies the signals are highly attenuated, and we found that the signal to noise ratio becomes too small for reliable amplitude measurements if the amplitudes are smaller than about 10% of the reference signal. With an average sample length of 36.5 mm, the measured attenuations are therefore significant only if $\alpha < 0.06 \text{ mm}^{-1}$ [Fig. 1(b)].

II. EVALUATION OF DIFFERENT THEORIES OF ATTENUATION CAUSED BY PORE FLUIDS

A. Fluid flow theories

Fluid flow theories predict attenuation of the incident elastic wave due to fluid flow in pores and cracks, which leads to viscous dissipation of acoustic energy. All of the different mechanisms can be characterized by critical frequencies where the attenuation is greatest and which are inversely proportional to the viscosity of the pore fluid. If fluid flow theories are applied to our experiment, one would predict large changes in attenuation as the fluid viscosity is varied by three orders of magnitude. However, Fig. 1(b) shows no such relationship. The samples saturated with silicon oils show slightly stronger attenuation than the water saturated sample, but the sand saturated with the 1000 times more viscous castor oil shows almost identical attenuation. Hence, fluid flow theories are not capable of explaining our experimental results, at least in the frequency range considered in this study, and have to be rejected as a possible physical model.²²

B. Scattering theories

A comparison of the results in Fig. 1(b) with the fluid properties in Table I suggests that the attenuation is more likely to be related to the fluid velocity and density than the fluid viscosity. Thus a scattering mechanism would appear to be a plausible physical model, as such mechanisms depend strongly on the impedance contrast between the fluid and the matrix.

Depending on the ratio of wavelength to the size of the scatterer, different scattering regimes exist. In our experiment, pore and grain sizes are much smaller than the wavelength at the central frequency of the reference. The scattering effects can therefore be described by the Rayleigh scattering regime, which predicts that the attenuation is proportional to the fourth power of frequency.^{12,13,15} However, Kaelin and Johnson have shown that exact scattering functions are required to adequately describe scattering phenomena.¹⁵ They have also shown that the coherent signal is not confined to one dominant pulse, but widens and becomes oscillatory due to scattering. This makes it especially difficult to extract the coherent part of the wave front from experimental data. In our experimental setup the second half of the first cycle already includes reflections from the back of the transducers. Hence, we used the dynamic composite elastic medium theory (DYCEM),¹⁵ which includes multiple forward scattering from spherical pores and grains as well as

intrinsic attenuation of viscous fluids, to compute full waveforms that could be compared with the experimental data.

Before the waveforms for the different samples can be calculated, we first have to determine the P velocities and S velocities of the matrix. In unconsolidated porous media, the velocities of elastic waves in the matrix are generally smaller than those of the single crystals.^{6,23} It has been generally accepted that this effect is caused by cracks within the grains and grain contacts. All our samples showed very slow low frequency S wave velocities, which made it impossible to determine the exact S wave velocity. However, we were able to determine the high frequency S wave velocity, i.e., the S velocity of the matrix, by using smaller sample lengths and gradually increasing the sample lengths. The measured S wave velocities of the matrix for the different samples are shown in Table II.

In order to obtain the P wave velocity of the matrix and the S wave velocity of the samples we used Berryman's theory,²⁰ which is equivalent to the low frequency limit of the dynamic composite elastic medium theory. Thus with the measured porosities and phase velocities in Table II and the fluid properties in Table I, the elastic moduli of the matrix and of the samples can be calculated using the equations of Berryman.²⁰

$$\begin{aligned} \mu &= \left[\frac{\phi}{F} - \frac{(1-\phi)}{\mu_{Ma} + F} \right]^{-1} - F, \\ K_{Ma} &= (1-\phi) \left[\frac{1}{K + 4\mu/3} - \frac{\phi}{K_f + 4\mu/3} \right]^{-1} - 4\mu/3, \\ F &= \frac{\mu}{6} \frac{9K + 8\mu}{K + 2\mu}, \\ \rho &= (1-\phi)\rho_{Ma} + \phi\rho_f, \\ \mu_{Ma} &= \rho_{Ma}(v_s)_{Ma}^2, \\ K &= \rho v_p^2 - 4\mu/3. \end{aligned} \quad (7)$$

Here, K_f is the bulk modulus and ρ_f the density of the fluid, while K_{Ma} , μ_{Ma} , ρ_{Ma} and K , μ , ρ are the bulk modulus, shear modulus and density of the matrix and the saturated sample, respectively. The results are shown in Tables II and III.

TABLE III. Calculated P wave and measured S wave velocities of the matrix after Berryman (Ref. 20) for the different samples. The adhesion tension between pore fluids and quartz grains has been calculated with the measured surface tension and the capillary rise in a glass tube with 1 mm diameter. The mean spherical pore radius has been calculated with the dynamic composite elastic medium theory.

Saturated sand	Spherical					
	Matrix v_p^a (m/s)	Matrix v_s^b (m/s)	pore radius ^a (μm)	Surface tension ^b (dyne/cm)	Capillary rise ^b (mm)	Adhesion tension ^a (dyne/cm)
Water	2522	1029	120	70.5	13	33.1
cs10 oil	2093	742	95	23.1	6	14.5
cs100 oil	2322	725	115	23.8	6	14.5
Castor oil	2335	1012	130	40.7	13	31

^aCalculated properties.

^bMeasured properties.

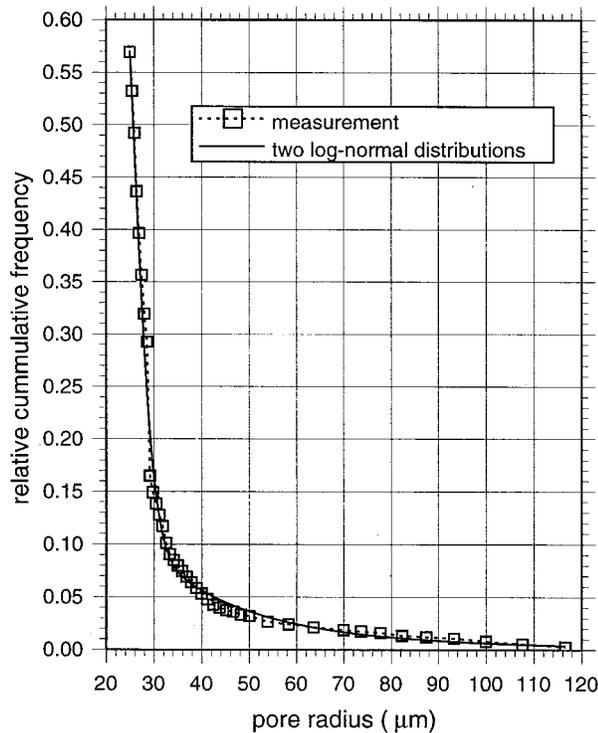


FIG. 2. Pore size distribution from capillary pressure measurements on the primary drainage curve, assuming the pores can be represented by the capillary bundle model. To fit the measurements, we have used two log-normal distributions with a mean pore radius of $26.3 \mu\text{m}$ and standard deviations of 2.63 and $38.2 \mu\text{m}$, respectively.

The calculated matrix velocities in Table III vary significantly for the different samples, even though we have used the same quartz sand. We suspect that differences in the physical fluid properties have caused changes in the effective elastic matrix properties. Table III shows that the variation in velocity, particularly the S velocity, correlates well with the adhesion tension between the different fluids and the quartz grains. Adhesion tension was calculated for the different fluids from the measured surface tension and the capillary rise in a glass tube of 1 mm diameter. We hypothesize that the fluids form thin films between the quartz grains and that the fluids with high adhesion tension provide better contact between the grains, thus increasing the effective stiffness of the matrix. However, a detailed study of the adhesion tension effect on the ultrasonic waves is beyond the scope of this study.

The pore size distribution was determined from a fully water saturated sand sample by progressively desaturating it while increasing the capillary pressure. By using the capillary bundle model,²⁴ we can relate the pore radius to the cumulative relative frequency (Fig. 2). Since the capillary bundle model is an idealization of the pore geometry represented by a collection of capillary tubes of different radii, we have used only the measured pore size distribution and left the mean pore radius as an adjustable parameter. We found that the sum of two log-normal distributions can explain the measured pore size distribution reasonably well (Fig. 3). The log-normal distributions both had a mean of $26.3 \mu\text{m}$ and standard deviations of 2.63 and $38.2 \mu\text{m}$, respectively. The

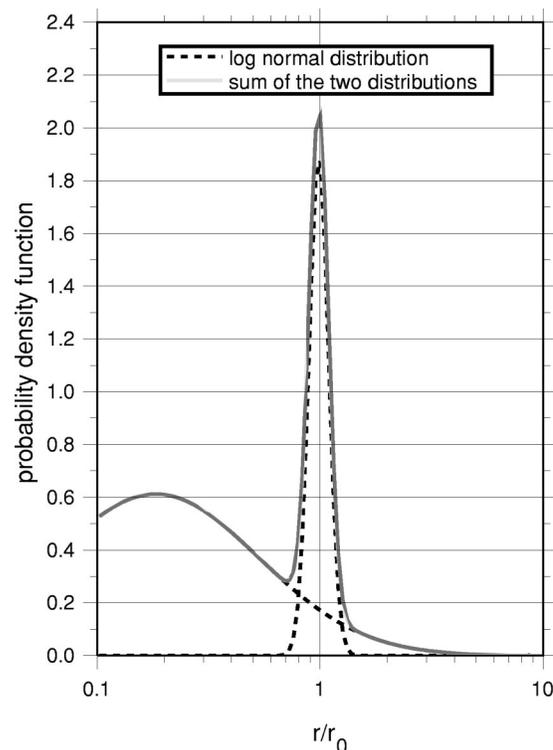


FIG. 3. Two log-normal distributions representing the measured pore radii distribution of Fig. 2. The first distribution takes up 47%, and the second 53%, of the summed distribution.

weight of the distributions was 47% and 53%. The grain size was determined from the sieve openings between 210 and $250 \mu\text{m}$, which corresponds to 230 ± 20 grain diameter or $115 \pm 10 \mu\text{m}$ grain radius.

To obtain the mean pore radius of spherical pores, we calculated the frequency dependent attenuation with the dynamic composite elastic medium theory and the material properties in Table I, the matrix velocities in Table III and the pore size distribution in Fig. 3. By minimizing the difference between the calculated and the measured attenuation [Fig. 1(b)], the mean pore radius for each sample was obtained (Table III). Since each sample represents one realization of the grain and pore size distribution only, we used the average pore radius and its standard deviation of the four samples, which is $115 \pm 15 \mu\text{m}$.

For comparison of the synthetic data and the laboratory measurements, we calculated a Green function using the dynamic composite medium method and convolved it with the reference spectrum. Figure 4 shows the calculated and the measured P waves after propagation through sand samples saturated with fluids of different viscosity. The amplitudes have been normalized by the first arrival. The frequency dependent attenuation has been computed with half a cycle of the first arrival, analogous to the measured traces. Figure 5 shows the attenuation as a function of frequency for calculated and measured data for (a) water, (b) 10cs silicone oil, (c) 100cs silicone oil and (d) castor oil. The mean synthetic attenuation and the uncertainty limits shown in Fig. 5 were obtained from the average pore radius of $115 \mu\text{m}$ and two standard deviations of $30 \mu\text{m}$, respectively.

The measurements lie within the uncertainty limits for

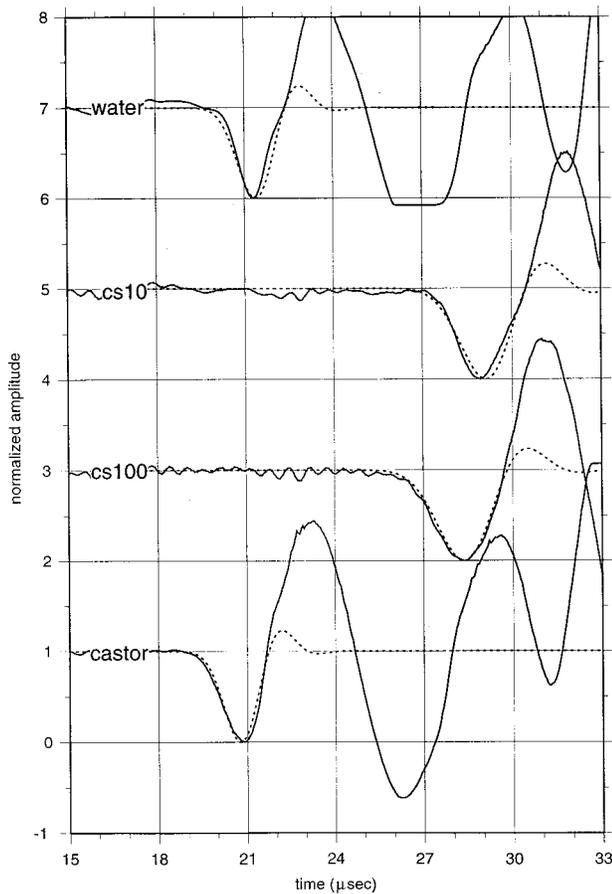


FIG. 4. Synthetic (dashed lines) and measured (solid lines) ultrasonic waves through a sand sample fully saturated with fluids of different viscosity. The synthetic traces were calculated with the dynamic composite elastic medium theory.

all fluids and the attenuation always shows a frequency squared dependence with small shifts caused by viscosity. The differences between the different samples is mainly caused by weaker or stronger scattering due to the velocity differences between matrix and fluids. Since the dynamic composite elastic medium theory accounts for both scattering and viscous intrinsic attenuation, we can conclude that the viscous attenuation is much weaker than the scattering attenuation at the studied ultrasonic frequencies. Hence, scattering is the dominant attenuation mechanism for unconsolidated sand with grain diameters of $230 \mu\text{m}$ or larger in the frequency range between 100 and 1000 kHz.

III. CONCLUSIONS

The P wave attenuation of an unconsolidated sand with different viscous fluids shows a frequency squared dependence for frequencies between 100 and 1000 kHz. The attenuation shows no correlation with the viscosity of the different pore fluids and thus theories that depend upon fluid flow cannot explain these data. It has been shown that a generalized scattering theory can predict the correct frequency dependence, with the calculated values lying within the uncertainty of the measurements. Thus we conclude that for grain diameters larger than $230 \mu\text{m}$ the P wave attenuation in this frequency range is caused mainly by scattering

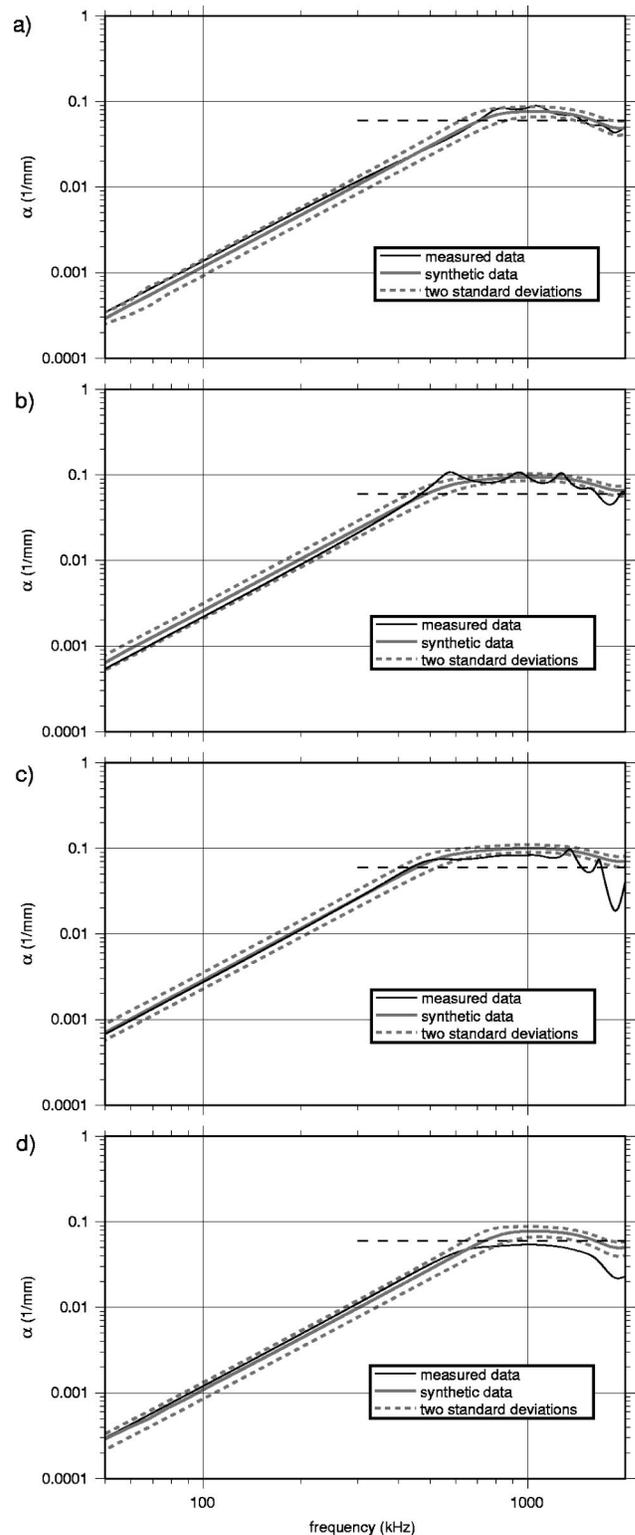


FIG. 5. Attenuation calculated as a function of frequency for the synthetic and measured data shown in Fig. 4 for (a) water, (b) cs10 silicone oil, (c) cs100 oil and (d) castor oil. Attenuation values are considered significant only when less than 0.6 mm^{-1} , which is marked by the horizontal dashed line.

attenuation, with the viscosity of the pore fluid contributing only a minor effect. Note, however, that at lower frequencies scattering attenuation becomes weaker and the effect of viscous fluids may become more important in a relative sense.²⁵ This possibility that the dominant attenuation effects are dif-

ferent at high and low frequencies might explain some of the differences observed when comparing laboratory and field measurements.

ACKNOWLEDGMENTS

The authors thank Tom Daley and Kurt Nihei for their helpful insight in the preparation of this paper. This research was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Geosciences Program, through U.S. Department of Energy Contract No. DE-AC03-76F00098, the U.S. Environmental Protection Agency Geophysics Program, Las Vegas, Nevada and by the Air Force, through Scientific Research Contract No. AFOSR-960010.

- ¹K. W. Winkler and W. F. Murphy III, "Rock physics and phase relations," AGU reference shelf **3**, 20–34 (1995).
- ²W. F. Murphy, K. W. Winkler, and R. L. Kleinberg, "Acoustic relaxation in sedimentary rocks: Dependence on grain contacts and fluid saturation," *Geophysics* **51**, 757–766 (1986).
- ³R. Burridge and J. Keller, "Poroelasticity equations derived from microstructure," *J. Acoust. Soc. Am.* **70**, 1140–1146 (1981).
- ⁴G. M. Mavko, "Frictional attenuation: An inherent amplitude dependence," *J. Geophys. Res.* **84**, 4769–4775 (1979).
- ⁵G. A. Gist, "Fluid effects on velocity and attenuation in sandstones," *J. Acoust. Soc. Am.* **96**, 1158–1173 (1997).
- ⁶F. Gassmann, "Über die Elastizität elastischer Medien," *Vierteljahresschrift der Naturforschenden Gesellschaft Zürich* **96**, 1–21 (1951).
- ⁷M. 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).
- ⁸M. A. Biot, "Theory of propagation of elastic waves in a fluid saturated porous solid: II. Higher frequency range," *J. Acoust. Soc. Am.* **28**, 179–191 (1956).
- ⁹R. J. O'Connell and B. Budiansky, "Viscoelastic properties of fluid-saturated cracked solids," *J. Geophys. Res.* **82**, 5719–5735 (1977).
- ¹⁰J. Dvorkin and A. Nur, "Dynamic poroelasticity: A unified model with the squirt and the Biot mechanisms," *Geophysics* **58**, 525–533 (1993).
- ¹¹M. S. Sams, J. P. Neep, M. H. Worthington, and M. S. King, "The measurement of velocity dispersion and frequency-dependent intrinsic attenuation in sedimentary rocks," *Geophysics* **62**, 1456–1464 (1997).
- ¹²C. M. Sayers, "Ultrasonic velocity dispersion in porous materials," *J. Phys. D* **74**, 413–420 (1981).
- ¹³K. W. Winkler, "Frequency dependent ultrasonic properties of high-porosity sandstones," *J. Geophys. Res.* **88**, 9493–9499 (1983).
- ¹⁴P. K. Seifert, J. T. Geller, and L. R. Johnson, "Effect of *P*-wave scattering on velocity and attenuation in unconsolidated sand saturated with immiscible liquids," *Geophysics* **63**, 161–170 (1998).
- ¹⁵B. Kaelin and L. R. Johnson, "Dynamic composite elastic medium theory, Part II. Three dimensional media," *J. Appl. Phys.* **84**, 5458–5468 (1998).
- ¹⁶J. T. Geller and L. R. Myer, "Ultrasonic imaging of organic contaminants in unconsolidated porous media," *J. Contam. Hydrol.* **19**, 85–104 (1995).
- ¹⁷M. N. Toksöz, D. H. Johnston, and A. Timur, "Attenuation of seismic waves in dry and saturated rocks: I. Laboratory measurements," *Geophysics* **44**, 681–690 (1979).
- ¹⁸J. Remi, M. Bellanger, and F. Homand-Etienne, "Laboratory velocity and attenuation of *P*-waves in limestones during freeze-thaw cycles," *Geophysics* **59**, 245–251 (1994).
- ¹⁹X. M. Tang, M. N. Toksöz, P. Tarif, and R. H. Wilkens, "A method for measuring acoustic wave attenuation in the laboratory," *J. Acoust. Soc. Am.* **83**, 453–462 (1988).
- ²⁰J. G. Berryman, "Long-wavelength propagation in composite elastic media. I. Spherical inclusions," *J. Acoust. Soc. Am.* **68**, 1809–1819 (1980).
- ²¹G. Kuster and M. Toksöz, "Velocity and attenuation of seismic waves in two-phase media: Part I. Theoretical formulations," *Geophysics* **39**, 587–606 (1974).
- ²²P. K. Seifert, "Effects of pore fluids in the subsurface on ultrasonic wave propagation," Ph.D. thesis, University of California, Berkeley, 1998.
- ²³J. E. White, *Seismic Waves: Radiation Transmission and Attenuation* (McGraw-Hill, New York, 1965).
- ²⁴E. C. Childs and N. Collis-George, "The permeability of porous materials," *Proc. R. Soc. London, Ser. A* **201**, 392–405 (1950).
- ²⁵T. D. Jones, "Pore fluids and frequency-dependent wave propagation," *Geophysics* **51**, 1939–1953 (1986).
- ²⁶A. R. Selfridge, "Approximate material properties in isotropic materials," *IEEE Trans. Sonics Ultrason.* **32**, 381–394 (1985).

Computation of transient radiation in semi-infinite regions based on exact nonreflecting boundary conditions and mixed time integration

Lonny L. Thompson and Runnong Huan

Department of Mechanical Engineering, Clemson University, Clemson, South Carolina 29634

(Received 4 March 1999; revised 17 July 1999; accepted 30 August 1999)

Transient radiation in a semi-infinite region, bounded by a planar infinite baffle with a local acoustic source is considered. The numerical simulation of the transient radiation problem requires an artificial boundary Γ , here chosen to be a hemisphere, which separates the computational region from the surrounding unbounded acoustic medium. Inside the computational region we use a semidiscrete finite element method. On Γ , we apply the exact nonreflecting boundary condition (NRBC) first derived by Grote and Keller for the free-space problem. Since the problem is symmetric about the infinite planar surface, in order to satisfy the rigid baffle condition it is sufficient to restrict the indices in the spherical harmonic expansion which defines the NRBC and scale the radial harmonics which drive auxiliary equations on the boundary. The Fourier expansion in the circumferential angle appearing in the NRBC may be used to efficiently model axisymmetric problems in two dimensions. A new mixed explicit-implicit time integration method which retains the efficiency of explicit pressure field updates without the need for diagonal matrices in the auxiliary equations on Γ is presented. Here, the interior finite element equations are integrated explicitly in time while the auxiliary equations are integrated implicitly. The result is a very natural and highly efficient algorithm for large-scale wave propagation analysis. Numerical examples of fully transient radiation resulting from a piston transducer mounted in an infinite planar baffle are compared to analytical solutions to demonstrate the accuracy of the mixed time integration method with the NRBC for the half-space problem. © 1999 Acoustical Society of America.

[S0001-4966(99)04712-8]

PACS numbers: 43.20.Px, 43.20.Bi [ANN]

INTRODUCTION

We consider the problem of determining the transient acoustic field radiated from an arbitrary shaped transducer or vibrating structure in a semi-infinite three-dimensional region, bounded by a planar infinite baffle. Modeling of local acoustic sources in a half-space has broad application including numerical simulation of piezoelectric transduction systems; ultrasonics and nondestructive testing, and noninvasive therapeutic applications such as high-intensity focused ultrasound. Examples include ring transducers used in sonar devices¹ and geometrically focused transducers.² Often, the transducer must deliver a precise acoustic near-field radiation pattern which is difficult to measure experimentally.³ For this reason, in recent years, there has been increased interest in the use of numerical simulation to predict the acoustic radiation field and to aid in the design of optimal transduction systems, e.g., Refs. 4 and 5.

When modeling transient radiation/scattering from structures in an acoustic medium which extends to infinity with finite element/difference/volume methods, the computational domain must be truncated at a finite distance from the structure. The impedance of the unbounded domain exterior to the artificial truncation boundary is then represented on this boundary by either absorbing boundary conditions, infinite elements, or matched "sponge" layers. Survey articles of various boundary treatments are given in Ref. 6. If accurate boundary treatments are used, the finite computational region

can be reduced so that the truncation boundary is relatively close to the radiator/scatterer, and fewer acoustic elements than otherwise would be possible may be used, resulting in considerable savings in both cpu time and memory. In the frequency domain, several accurate and efficient methods for representing the impedance of the far field are well understood, including the Dirichlet-to-Neumann (DtN) map,^{7,8} and infinite elements.⁹ However, efficient evaluation of accurate boundary treatments for the time-dependent wave equation on unbounded spatial domains has long been an obstacle for the development of reliable solvers for time domain simulations. Ideally, the artificial boundary would be placed as close as possible to the source, and the radiation boundary treatment would be capable of arbitrary accuracy at a cost and memory not exceeding that of the interior solver.

A standard approach is to apply local (differential) boundary operators which annihilate leading terms in the radial multipole expansion for outgoing wave solutions. A well-known sequence of boundary conditions developed for a spherical truncation boundary are the local operators derived by Bayliss and Turkel.¹⁰ Because these operators involve only local spatial derivatives, while derived for a spherical boundary in free space, they may be used without alteration for semi-infinite regions, such as the problem of a transducer mounted in an infinite half-space. However, these and other approximate local boundary conditions exhibit significant spurious reflection for high-order wave harmonics, especially as the position of the truncation boundary ap-

proaches the source of radiation.^{11,12,13} In addition, as the order of these local non-reflecting boundary conditions increases, they become increasingly difficult to implement in standard semidiscrete finite element formulations due to the occurrence of high-order time derivatives on the truncation boundary.

In recent years, new boundary treatments have been developed which dramatically improve both the accuracy and efficiency of time domain simulations compared to approximate local radiation boundary conditions. One promising approach is the application of the “perfectly matched layer” (PML) technique^{14,15} which introduces an external layer designed to absorb outgoing waves. In Refs. 16 and 17, exact nonreflecting boundary conditions (NRBCs) are derived for the free-space problem involving solution of an auxiliary Cauchy problem for linear first-order systems of time-dependent differential equations on a spherical boundary for each harmonic. In Ref. 18, the NRBC is rederived based on direct application of a result given in Lamb,¹⁹ with improved scaling of the first-order system of equations associated with the NRBC. This rescaling improves the numerical conditioning of the first-order system of equations when solved with implicit methods. Formulation of the NRBC in standard semidiscrete finite element methods with several alternative implicit and explicit time-integrators is reported in Refs. 18 and 20. When implemented in the finite element method, NRBC requires inner products of spherical harmonics and standard C^0 continuous basis functions with compact support, appearing in the force vector. As a result, the NRBC may be implemented efficiently and does not disturb the symmetric and banded/sparse structure of the finite element matrix equations. In Ref. 21, an efficient method is described for calculating far-field solutions concurrently with the near-field solution based on the exact NRBC.

In Ref. 20, a modified version of the exact NRBC for the free-space problem with improved accuracy for high-order harmonics is formulated. To obtain a symmetric finite element variational equation, an additional auxiliary function is introduced on the artificial truncation boundary. This modified version may be viewed as an extension of the second-order local boundary operator derived by Bayliss and Turkel.¹⁰ In Ref. 20 an implicit time discretization scheme is developed to integrate the semidiscrete finite element equations. However, in three dimensions because of the difficulty in obtaining diagonal matrices for the auxiliary equations, a fully explicit time discretization which uncouples the system of equations was not possible.

Motivated by the excellent accuracy of the NRBC for the free-space problem, it is natural to extend these ideas to the problem of radiation in a semi-infinite acoustic domain resulting from transducers or vibrating structures mounted in a half-plane. However, as a result of the nonlocal spherical transform and expansion on a spherical boundary in free space, the NRBC must be modified for the infinite half-space problem. In this work, we give the extension of the exact NRBC originally derived by Grote and Keller¹⁶ for the free-space problem for application to semi-infinite problems defined by an infinite planar baffle. Two alternative forms of the NRBC which satisfy the symmetry condition imposed by

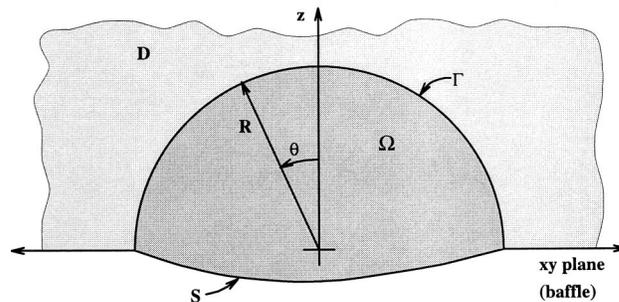


FIG. 1. Illustration of semi-infinite region lying on one side of a boundary composed of an arbitrary radiation surface S and the xy plane. The computational domain Ω is surrounded by a semi-spherical truncation boundary Γ . Exterior region denoted by D .

the rigid baffle are possible; in the first we orient the baffle normal (perpendicular) to a z axis of revolution defined in spherical coordinates, while in the second the baffle is aligned (parallel) with the z axis. The advantage of the first approach is that axisymmetric radiation in a half-space may be modeled efficiently in two dimensions.

For the symmetric form of the modified NRBC we give a new mixed explicit-implicit time integration method which retains the efficiency of explicit time discretization for the finite element matrix equations, without the need for diagonal matrices in the auxiliary equations on the artificial truncation boundary. Here, the interior finite element equations are integrated explicitly in time while the auxiliary equations are integrated implicitly in time. By treating the auxiliary equations on the boundary implicitly, a very natural and highly efficient algorithm is developed for large-scale wave propagation analysis which allows the pressure field to be updated without assembling or factoring the interior finite element matrices.

In Refs. 18 and 20 numerical experiments for radiation from a sphere in free-space are presented which demonstrate the accuracy of the NRBC compared to steady-state analytical solutions and standard local absorbing boundary conditions. In this work, numerical results for fully transient solutions for a circular transducer mounted in an infinite rigid planar baffle are compared to analytical solutions. The numerical results are used to assess the accuracy of the mixed explicit-implicit time integration method with the NRBC restricted for the half-space problem.

I. TRANSIENT RADIATION IN ACOUSTIC HALF-SPACE

We consider time-dependent scattering/radiation in a three-dimensional semi-infinite region bounded by a boundary composed of an arbitrary shaped radiation surface S and a planar infinite baffle (see Fig. 1). We denote the space above this plane as the semi-infinite region \mathcal{R} . The numerical simulation of the transient radiation problem requires an artificial boundary Γ , here chosen to be a semi-sphere of radius $||\mathbf{x}|| = R$, which separates the computational region Ω from the surrounded unbounded acoustic medium. At R we impose an absorbing boundary condition to reduce spurious reflection from it. Inhomogeneities and nonlinear sources

may be incorporated within the computational domain while the remainder of the problem is treated as a homogeneous acoustic medium occupying an infinite half-space and is dealt with through the domain truncation boundary.

Within Ω the acoustic pressure $p(\mathbf{x}, t)$ satisfies the scalar wave equation:

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = -f, \quad \mathbf{x} \text{ in } \Omega, \quad t > 0, \quad (1)$$

with initial conditions

$$p(\mathbf{x}, 0) = p_0(\mathbf{x}), \quad \frac{\partial p}{\partial t}(\mathbf{x}, 0) = \dot{p}_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (2)$$

and driven by a normal velocity $v = -\mathbf{v} \cdot \mathbf{n}$, prescribed on the radiation boundary \mathcal{S} :

$$\frac{\partial p}{\partial n} = \rho_0 \dot{v}(\mathbf{x}, t), \quad \mathbf{x} \in \mathcal{S}, \quad t > 0. \quad (3)$$

In the above, $c(\mathbf{x})$ is the velocity of sound in the acoustic medium, $\rho_0(\mathbf{x})$ is the mass density, a superimposed dot denotes a time derivative, and \mathbf{n} is an outward pointing normal vector. The normal velocity represents the time-dependent motion of the transducer. The normal velocity $v(\mathbf{x}, t)$, acoustic source $f(\mathbf{x}, t)$, and initial data are assumed to be confined to the interior of the region Ω , so that in the infinite half-space, i.e., the region outside Γ , the acoustic pressure field $p(\mathbf{x}, t)$ satisfies the homogeneous form of the wave equation with constant wave speed c ,

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0 \quad \mathbf{x} \text{ in } \mathcal{D}, \quad t > 0, \quad (4)$$

$$p(\mathbf{x}, 0) = 0, \quad \frac{\partial p}{\partial t}(\mathbf{x}, 0) = 0, \quad \mathbf{x} \in \mathcal{D}, \quad (5)$$

and the homogeneous Neumann boundary condition imposed by the rigid baffle,

$$\frac{\partial p}{\partial n} = 0, \quad \mathbf{x} \in \mathcal{S}, \quad t > 0. \quad (6)$$

In the following, we introduce spherical coordinates (r, θ, φ) ,

$$x = r \cos \varphi \sin \theta, \quad (7)$$

$$y = r \sin \varphi \sin \theta, \quad (8)$$

$$z = r \cos \theta, \quad (9)$$

such that the z axis is aligned *perpendicular* (normal) to the planar baffle.

With this parametrization, the artificial boundary is defined by the hemisphere,

$$\Gamma := \{r = R, 0 < \theta \leq \pi/2, 0 < \varphi \leq 2\pi\},$$

and a general solution to the wave equation (4) in the exterior region $\mathcal{D} = \{r \geq R, 0 \leq \theta \leq \pi/2, 0 \leq \varphi < 2\pi\}$ may be expanded as

$$p(r, \theta, \varphi, t) = \sum_{n=0}^{\infty} \sum_{m=0}^n {}' P_n^m(\cos \theta) (p_{nm}^c(r, t) \cos m\varphi + p_{nm}^s(r, t) \sin m\varphi). \quad (10)$$

Here P_n^m is the associated Legendre function of the first kind, and the prime on the sum indicates that a factor of $\frac{1}{2}$ multiplies the term with $m=0$. The radial harmonics associated with the even and odd trigonometric functions are computed from

$$p_{nm}^c = \frac{2}{N_{nm}} \int_0^{2\pi} \int_0^{\pi/2} p(r, \theta, \varphi, t) P_n^m \times (\cos \theta) \cos m\varphi \sin \theta \, d\theta d\varphi, \quad (11)$$

$$p_{nm}^s = \frac{2}{N_{nm}} \int_0^{2\pi} \int_0^{\pi/2} p(r, \theta, \varphi, t) P_n^m \times (\cos \theta) \sin m\varphi \sin \theta \, d\theta d\varphi, \quad (12)$$

where N_{nm} is the normalization factor for the orthogonal spherical harmonics:

$$N_{nm} = \frac{2\pi(n+m)!}{(2n+1)(n-m)!}. \quad (13)$$

For the semi-infinite region defined by the half-space \mathcal{D} , the multiplier 2 appearing in (11) and (12) results from integration restricted over the hemisphere only, the limits of integration ranging from $0 \leq \theta \leq \pi/2$.

Since the problem is symmetric about the rigid planar baffle at $\theta = \pi/2$, i.e.,

$$\left. \frac{\partial p}{\partial \theta} \right|_{\theta=\pi/2} = -r \left. \frac{\partial p}{\partial z} \right|_{z=0} = 0, \quad r \geq R, \quad (14)$$

it is sufficient to restrict the expansion (10) in spherical harmonics to indices $n+m$ even. While this modification is trivial, it is not necessarily obvious. To prove this result, we evaluate the expansion (10) at $\theta = \pi/2$,

$$\left. \frac{\partial p}{\partial \theta} \right|_{\theta=\pi/2} = \sum_{n=0}^{\infty} \sum_{m=0}^n {}' P_n^m(0) (p_{nm}^c(r, t) \cos m\varphi + p_{nm}^s(r, t) \sin m\varphi). \quad (15)$$

From properties of the associated Legendre functions,²²

$$P_n^m(0) = \begin{cases} 0, & n+m = \text{even}, \\ (-1)^{(m+n-1)/2} \frac{1 \cdot 3 \cdot 5 \cdots (n+m)}{2 \cdot 4 \cdot 6 \cdots (n-m-1)}, & n+m = \text{odd}, \end{cases} \quad (16)$$

only the combination $n+m = \text{even}$ satisfies $P_n^m(0) = 0$, and it follows that expansion (10) satisfies (15) only with the restriction $n+m = \text{even}$.

II. EXACT NRBCs ON A HEMISPHERE

On the artificial boundary Γ the radial functions p_{nm}^c and p_{nm}^s appearing in (10) satisfy the boundary condition derived in Refs. 16 and 18 for a spherical boundary in free-space:

$$B_1[p_{nm}] = -\frac{1}{r} \mathbf{c}_n \cdot \mathbf{z}_{nm}(t), \quad r=R, \quad (17)$$

where

$$B_1[p_{nm}] := \left(\frac{\partial}{\partial r} + \frac{1}{c} \frac{\partial}{\partial t} + \frac{1}{r} \right) p_{nm} \quad (18)$$

is the ‘‘first-order’’ local boundary operator of Bayliss and Turkel.¹⁰ The constant n -component vector $\mathbf{c}_n = \{c_n^j\}$ is defined with coefficients

$$c_n^j = n(n+1)j/2R, \quad j=1,2,\dots,n, \quad (19)$$

while the vector functions $\mathbf{z}_{nm}(t) = \{z_{nm}^j(t)\}$, $j=1,\dots,n$, of order n satisfy the first-order system of ordinary differential equations,

$$\begin{aligned} \frac{d}{dt} \mathbf{z}_{nm}(t) &= \mathbf{A}_n \mathbf{z}_{nm}(t) + c \Phi_{nm}(t), \\ \mathbf{z}_{nm}(0) &= 0, \end{aligned} \quad (20)$$

with constant $n \times n$ matrices $\mathbf{A}_n = \{A_n^{ij}\}$ defined with coefficients¹⁸

$$A_n^{ij} = \begin{cases} \frac{-n(n+1)c}{2R}, & \text{if } i=1, \\ \frac{(n+i)(n-i+1)c}{2iR}, & \text{if } i=j+1, \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

For the semi-infinite half-space problem considered here, the system (20) is driven by the time-dependent vector function,

$$\Phi_{nm}(t) = [cp_{nm}|_{r=R}, 0, \dots, 0]^T, \quad (22)$$

with radial modes evaluated at $r=R$, $p_{nm}|_{r=R}$, computed from (11) and (12).

The exact nonreflecting boundary condition (NRBC) for the half-space problem on the hemisphere Γ is obtained by multiplying (17) by spherical harmonics, summing over n and m , setting $r=R$ and using (10),

$$\begin{aligned} B_1[p] &= -\frac{1}{R} \sum_{n=1}^{\infty} \sum_{m=0}^n P_n^m(\cos \theta) (w_{nm}^c(t) \cos m\varphi \\ &\quad + w_{nm}^s(t) \sin m\varphi), \quad n+m = \text{even}, \end{aligned} \quad (23)$$

where $w_{nm}^c = \mathbf{c}_{nm} \cdot \mathbf{z}_{nm}^c$ and $w_{nm}^s = \mathbf{c}_{nm} \cdot \mathbf{z}_{nm}^s$ are scalar functions defined by the even and odd harmonics in φ . This condition is the same as the free-space problem derived in Refs. 16 and 18, except that for the rigid baffle symmetry condition, the indices are restricted to $n+m = \text{even}$, and the radial harmonics include a factor of 2 resulting from integration over a hemisphere.

Alternatively, the z axis may be *aligned* (parallel) with the infinite baffle such that the hemisphere is defined by $\Gamma := \{r=R, 0 < \theta \leq \pi, 0 < \phi \leq \pi\}$. With this orientation, the symmetry condition imposed by the rigid planar baffle is satisfied by restricting the Fourier expansion in (10) to even functions $\cos m\phi$. In this case, the exact nonreflecting boundary condition for the hemisphere may be written as

$$\begin{aligned} B_1[p] &= -\frac{1}{R} \sum_{n=1}^{\infty} \sum_{m=0}^n \mathbf{c}_n \cdot \mathbf{z}_{nm}(t) P_n^m(\cos \theta) \cos m\phi \quad \text{on } \Gamma, \end{aligned} \quad (24)$$

where the system of equations (20) for \mathbf{z}_{nm} is driven by the radial modes,

$$\begin{aligned} p_{nm}|_{r=R} &= \frac{2}{N_{nm}} \int_0^\pi \int_0^\pi p(R, \theta, \phi, t) P_n^m \\ &\quad \times (\cos \theta) \cos m\phi \sin \theta d\theta d\phi, \end{aligned} \quad (25)$$

with integration restricted over the range $0 \leq \phi \leq \pi$.

III. AXISYMMETRIC PROBLEMS

For general problems in three dimensions, the two forms of the NRBC (23) and (24) have the same storage requirements and operation counts. An advantage in expressing the NRBC in the form (23) is that axisymmetric problems in a half-space defined by a planar baffle may be solved efficiently with a periodic Fourier expansion in φ and imposing the planar symmetry condition in the Legendre function expansion in θ . In this case an efficient solution is obtained by reducing the axisymmetric problem to a sequence of uncoupled two-dimensional problems with a Fourier expansion in the circumferential direction φ about a z axis of revolution perpendicular to the planar baffle. This reduction is not possible with the NRBC expressed in the form (24) since the Fourier harmonics are restricted by the baffle condition. An example of axisymmetric radiation in a half-space is given by the classic model of transient radiation from a circular piston mounted in an infinite planar baffle (see numerical examples in Sec. V).

To be specific, for radiation surfaces \mathcal{S} with axisymmetric geometry about a z axis *perpendicular* (normal) to the planar baffle, and driven by acoustic sources (3) which are periodic in the angle of revolution φ , i.e.,

$$\dot{\mathbf{v}}(\mathbf{x}, t) = \sum_{m=0}^{\infty} [g_m^c(r, \theta, t) \cos m\varphi + g_m^s(r, \theta, t) \sin m\varphi], \quad (26)$$

then the pressure may be expressed by the Fourier series,

$$p = \sum_{m=0}^{\infty} [p_m^c(r, \theta, t) \cos m\varphi + p_m^s(r, \theta, t) \sin m\varphi]. \quad (27)$$

In this case, the pressure field decouples for different Fourier harmonics m due to the orthogonality of the trigonometric functions and the problem simplifies to solving for the Fourier modes $p_m^c(r, \theta, t)$ and $p_m^s(r, \theta, t)$ in a two-dimensional half-plane defined by the cylindrical coordinates (ρ, z) , with $\rho = r \sin \theta$ and $z = r \cos \theta$. Outgoing solutions for the modes p_m are absorbed exactly by imposing the NRBC in the form (23) with the variation in φ suppressed, i.e.,

$$B_1[p_m] = -\frac{1}{R} \sum_{n=1}^{\infty} \mathbf{c}_n \cdot \mathbf{z}_{nm}(t) P_n^m(\cos \theta), \quad n+m = \text{even}. \quad (28)$$

In the above $\mathbf{c}_n \cdot \mathbf{z}_{nm}(t) = w_{nm}^c(t)$ for modes $p_m = p_m^c$, and $\mathbf{c}_n \cdot \mathbf{z}_{nm}(t) = w_{nm}^s(t)$ for modes $p_m = p_m^s$. For the rigid baffle

condition, the indices in the Legendre function expansion remain restricted to $n+m$ even, and the system of equations (20) for \mathbf{z}_{nm} are driven by radial modes p_{nm} computed from the restricted Legendre transform,

$$p_{nm}|_{r=R} = \frac{2\pi}{N_{nm}} \int_0^{\pi/2} p_m(r, \theta, t) P_n^m(\cos \theta) \sin \theta d\theta. \quad (29)$$

When driven by sources $\dot{v} = g_0$ which are independent of the angle of revolution φ , the pressure field is defined by the single mode $p = p_0(r, \theta, t)$, and the exact NRBC reduces naturally by setting the index $m=0$ in (23), with the result

$$B_1[p] = -\frac{1}{R} \sum_{n=2,4,\dots}^{\infty} \mathbf{c}_n \cdot \mathbf{z}_{n0}(t) P_n(\cos \theta). \quad (30)$$

Here the system of equations (20) for \mathbf{z}_{n0} is driven by the radial modes,

$$p_{n0}|_{r=R} = (2n+1) \int_0^{\pi/2} p(R, \theta, t) P_n(\cos \theta) \sin \theta d\theta, \quad (31)$$

with integration restricted over the quarter circle, $0 \leq \theta < \pi/2$.

IV. MODIFIED NRBCs ON A HEMISPHERE

In practice, the infinite sum over n in (23) or (24) is truncated at a finite value N . In this case, we denote the boundary condition by NR1(N), where N defines the number of harmonics included in the truncated series. Use of NR1(N) on a hemisphere with boundary Γ will exactly represent all harmonics $p_{nm}(r, t)$, for $n \leq N$ in the outgoing solution to the initial-boundary value problem for the half-space. For $n > N$, then NR1(N) approximates the harmonics with the local operator $B_1[p] = 0$ on Γ , with leading error of order, $B_1[p] = O(1/R^3)$. Accuracy of the approximated harmonics, $n > N$, may be improved by increasing the radius of the truncation boundary R , but at the added expense of a larger computation region Ω , resulting in increased memory and cpu times.

To improve the approximation to the truncated harmonics $n > N$, without affecting the modes $n \leq N$, the second-order local boundary operator,

$$B_2[p_{nm}] := \left(\frac{\partial}{\partial r} + \frac{1}{c} \frac{\partial}{\partial t} + \frac{3}{r} \right) B_1[p_{nm}], \quad (32)$$

of Bayliss and Turkel¹⁰ may be used to obtain a modified boundary condition for the radial modes:^{17,20}

$$B_2[p_{nm}] = \frac{1}{r} \tilde{\mathbf{c}}_n \cdot \mathbf{z}_{nm}(t), \quad r=R. \quad (33)$$

Here the coefficient vector $\tilde{\mathbf{c}}_n = \{\tilde{c}_n^j\}$ is given by

$$\tilde{c}_n^j = n(n+1)j(j-1)/2R^2, \quad j=1,2,\dots,n, \quad (34)$$

and the vector functions $\mathbf{z}_{nm}(t)$ appearing in (33) satisfy the same first-order system of ordinary differential equations (20), driven by (22). This modified condition was first derived in Ref. 17 for a spherical boundary Γ in free-space and modified in Ref. 20 with improved scaling $\tilde{\mathbf{c}}_{nm}$.

To obtain an equivalent but more tractable form for finite element implementation, the second-order radial derivative appearing in the local B_2 operator defined in (32) is eliminated using the radial wave equation for the modes p_{nm} ,

$$\frac{\partial^2 p_{nm}}{\partial r^2} = \frac{1}{c^2} \frac{\partial^2 p_{nm}}{\partial t^2} - \frac{2}{r} \frac{\partial p_{nm}}{\partial r} + \frac{n(n+1)}{r^2} p_{nm}, \quad (35)$$

with the result

$$\frac{r}{c} \frac{\partial}{\partial t} B_1[p_{nm}] + B_1[p_{nm}] + \frac{n(n+1)}{2r} p_{nm} = \frac{1}{2} \tilde{\mathbf{c}}_n \cdot \mathbf{z}_{nm}(t), \quad r=R. \quad (36)$$

With the z axis oriented perpendicular (normal) to the infinite planar baffle, then multiplying (36) by the spherical harmonics defined in (10) gives the modified NRBC:

$$\hat{B}_2[p] = \frac{1}{2} \sum_{n=2}^N \sum_{m=0}^n P_n^m(\cos \theta) (\tilde{w}_{nm}^c(t) \cos m\varphi + \tilde{w}_{nm}^s(t) \sin m\varphi), \quad n+m = \text{even}, \quad (37)$$

$$\hat{B}_2[p] := \frac{R}{c} \frac{\partial}{\partial t} B_1[p] + B_1[p] - \frac{1}{2R} \Delta_\Gamma[p], \quad (38)$$

$$\Delta_\Gamma[p] := \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial p}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 p}{\partial \varphi^2}, \quad (39)$$

where $\tilde{w}_{nm} = \tilde{\mathbf{c}}_n \cdot \mathbf{z}_{nm}(t)$. This modified condition takes the same form as the free-space problem derived in Ref. 20. Here, the indices are restricted to $n+m = \text{even}$ in order to satisfy the rigid baffle condition.

We denote (37) by NR2(N). Use of NR2(N) will exactly represent all harmonics $p_{nm}(r, t)$, for $n \leq N$ on a semi-spherical truncation boundary for the half-plane. For $n > N$, the truncated condition (37) reduces to $B_2[p] = 0$ on Γ . This condition approximates the harmonics $n > N$, with leading error of the order, $B_2[p] = O(1/R^5)$. Therefore, when truncated at a finite value N , the modified condition approximates the truncated harmonics $n > N$ with greater accuracy than NR1.

In Ref. 20, we show how a modified boundary condition in the form (37) can be implemented in a symmetric finite element variational formulation for the free-space problem by introducing additional auxiliary functions $q_{nm}(t)$ and $\psi(\theta, \varphi, t)$, such that

$$B_1[p] - \frac{1}{2R} \Delta_\Gamma[\psi] = \frac{1}{2} \sum_{n=2}^N \sum_{m=0}^n P_n^m(\cos \theta) \times (q_{nm}^c(t) \cos m\varphi + q_{nm}^s(t) \sin m\varphi), \quad (40)$$

$$\left(\frac{R}{c} \frac{\partial}{\partial t} + 1 \right) \Delta_\Gamma[\psi] = \Delta_\Gamma[p], \quad \psi(\theta, \varphi, 0) = 0, \quad (41)$$

and q_{nm}^c and q_{nm}^s satisfy

$$\left(\frac{R}{c} \frac{d}{dt} + 1\right) q_{nm}(t) = \tilde{\mathbf{c}}_n \cdot \mathbf{z}_{nm}(t), \quad q_{nm}(0) = 0. \quad (42)$$

The three equations (40)–(42) define an equivalent form of the exact NRBC (37), suitable for implementation in a symmetric finite element formulation. With the z axis oriented perpendicular (normal) to the planar rigid baffle, it is sufficient to restrict the expansion in spherical harmonics given in (40) to indices $n + m$ even.

Alternatively, with the z axis aligned (parallel) with the planar baffle, the rigid condition is satisfied with the expansion

$$B_1[p] - \frac{1}{2R} \Delta_\Gamma[\psi] = \frac{1}{2} \sum_{n=2}^N \sum_{m=0}^n q_{nm}(t) P_n(\cos \theta) \cos m\phi. \quad (43)$$

Again, the advantage of the form (40), based on a Fourier expansion in φ about a z axis of revolution oriented normal to the planar baffle, is that axisymmetric problems in a half-space can be modeled efficiently in two dimensions.

V. FINITE ELEMENT FORMULATION

Finite element discretization of the bounded acoustic region Ω allows for a natural coupling to an elastic radiator on the surface \mathcal{S} . The finite element formulation of the NR1(N) defined in (23) or (24), and the symmetric form of NR2(N) defined in (40) or (43) for the half-space problem posed on a rigid planar baffle, follows the same form as given in Refs. 18 and 20 for the free-space problem. By introducing finite element approximations, a system of ordinary differential equations are obtained which must then be integrated in time. In the following, we summarize the semidiscrete matrix equations resulting from the symmetric form of NR2(N) and then present a new mixed explicit/implicit time-integration method for advancing the solution.

A. Finite element discretization

The finite element discretization is obtained by approximating the variational equation associated with the wave equation and the nonreflecting boundary condition. The variational equation within Ω is obtained by multiplying (1) with a weighting function δp and using the divergence theorem. For the symmetric NR2(N) condition, an auxiliary equation on Γ is obtained by multiplying (41) with a different weighting function $\delta\psi$, then integrating by parts. Using independent finite element approximations,

$$p(\mathbf{x}, t) \approx \mathbf{N}_p(\mathbf{x}) \mathbf{p}(t) \quad \text{in } \Omega \cup \partial\Omega, \quad (44)$$

$$\psi(\mathbf{x}, t) \approx \mathbf{N}_\psi(\mathbf{x}) \boldsymbol{\psi}(t) \quad \text{on } \Gamma, \quad (45)$$

where \mathbf{N}_p and \mathbf{N}_ψ are standard vector arrays of C^0 basis functions with compact support associated with each node of the finite element mesh in $\Omega \cup \partial\Omega$, and on the boundary Γ , results in the following coupled, symmetric system of semidiscrete matrix equations,

$$\mathbf{M}_p \frac{d^2 \mathbf{p}(t)}{dt^2} + \mathbf{C}_p \frac{d \mathbf{p}(t)}{dt} + \mathbf{K}_p \mathbf{p}(t) = \mathbf{F}(t) - \mathbf{A} \boldsymbol{\psi}(t), \quad (46)$$

$$\mathbf{C}_\psi \frac{d \boldsymbol{\psi}(t)}{dt} + \mathbf{K}_\psi \boldsymbol{\psi}(t) = \mathbf{A}^T \mathbf{p}(t). \quad (47)$$

The time-dependent vector $\mathbf{p}(t)$ determines the global solution at each node in the mesh, while $\boldsymbol{\psi}(t)$ is a vector of auxiliary parameters associated with the nodes on the truncation boundary. The symmetric arrays associated with the pressure field take the same form as the free-space problem given in Ref. 20:

$$\mathbf{M}_p = \int_\Omega \frac{1}{c^2} \mathbf{N}_p^T \mathbf{N}_p d\Omega, \quad (48)$$

$$\mathbf{C}_p = \int_\Gamma \frac{1}{c} \mathbf{N}_p^T \mathbf{N}_p d\Gamma, \quad (49)$$

$$\mathbf{K}_p = \int_\Omega (\nabla \mathbf{N}_p)^T (\nabla \mathbf{N}_p) d\Omega + \frac{1}{R} \int_\Gamma \mathbf{N}_p^T \mathbf{N}_p \Gamma. \quad (50)$$

Similarly, the symmetric arrays associated with the auxiliary function ψ take the form

$$\mathbf{C}_\psi = \frac{R^2}{2} \int_\Gamma \frac{1}{c} (\nabla^s \mathbf{N}_\psi)^T (\nabla^s \mathbf{N}_\psi) d\Gamma, \quad (51)$$

$$\mathbf{K}_\psi = \frac{R}{2} \int_\Gamma (\nabla^s \mathbf{N}_\psi)^T (\nabla^s \mathbf{N}_\psi) d\Gamma. \quad (52)$$

The coupling array between p and ψ is given by

$$\mathbf{A} = \frac{R}{2} \int_\Gamma (\nabla^s \mathbf{N}_p)^T (\nabla^s \mathbf{N}_\psi) d\Gamma, \quad (53)$$

where ∇^s denotes the surface gradient on a sphere and $d\Gamma = R^2 \sin \theta d\theta d\varphi$.

In deriving the matrix equations, second-order tangential derivatives appearing in Δ_Γ were reduced to first-order derivatives using integration-by-parts on the hemisphere Γ , i.e.,

$$\int_\Gamma \delta p \Delta_\Gamma \psi d\Gamma = -R^2 \int_\Gamma \nabla^s \delta p \cdot \nabla^s \psi d\Gamma. \quad (54)$$

For the z axis normal to the planar baffle, the above result follows from the symmetry condition in the angle θ , i.e., $\partial\psi/\partial\theta = 0$ at $\theta = \pi/2$, and the periodic condition in φ , i.e., $\delta p(R, \theta, 0) = \delta p(R, \theta, 2\pi)$.

The auxiliary functions ψ only appear as a surface gradient $\nabla^s \psi$. As a result, a family of solutions for ψ that differ by a constant will satisfy (47). To obtain a unique solution, the function ψ may be constrained at one arbitrary node on the truncation boundary. The value used to constrain the auxiliary function ψ at that node is inconsequential to the unique solution for p , and may be set to zero.²⁰

The semidiscrete equations are driven by the initial conditions and discrete force vector, $\mathbf{F}(t) = \mathbf{F}_S + \mathbf{F}_\Gamma$ composed of a standard load vector,

$$\mathbf{F}_S(t) = \int_\Omega \mathbf{N}_p^T f(\mathbf{x}, t) d\Omega + \int_S \mathbf{N}_p^T \rho_0 \dot{v}(\mathbf{x}, t) dS, \quad (55)$$

and a part associated with the auxiliary functions appearing in the NRBC,

$$\mathbf{F}_\Gamma(t) = \frac{1}{2} \sum_{n=2}^N \sum_{m=0}^n (q_{nm}^c(t) \mathbf{f}_{nm}^c + q_{nm}^s(t) \mathbf{f}_{nm}^s), \quad (56)$$

where

$$\mathbf{f}_{nm}^c := \int_\Gamma \mathbf{N}_p^T P_n^m(\cos \theta) \cos m \varphi d\Gamma, \quad (57)$$

$$\mathbf{f}_{nm}^s := \int_\Gamma \mathbf{N}_p^T P_n^m(\cos \theta) \sin m \varphi d\Gamma. \quad (58)$$

With the z axis normal to the rigid planar baffle, the indices appearing in (56) are restricted to $n+m = \text{even}$. The functions $q_{nm}(t)$ are solutions to the first-order equation (42), driven by the auxiliary variables $\mathbf{z}_{nm}(t)$. The vector functions \mathbf{z}_{nm} in turn satisfy (20) driven by the radial modes on the hemisphere:

$$p_{nm}^c(R, t) = \frac{2}{R^2 N_{nm}} \mathbf{f}_{nm}^{cT} \cdot \mathbf{p}_\Gamma(t), \quad (59)$$

$$p_{nm}^s(R, t) = \frac{2}{R^2 N_{nm}} \mathbf{f}_{nm}^{sT} \cdot \mathbf{p}_\Gamma(t), \quad (60)$$

where $\mathbf{p}_\Gamma(t) = \{p_I(t)\}$, $I = 1, 2, \dots, N_\Gamma$, is a vector of nodal solutions on the artificial boundary Γ with N_Γ nodes.

Implementation of the nonreflecting boundary condition only requires inner products of spherical harmonics and finite element basis functions with compact support within the force vector \mathbf{f}_{nm} . As a result, the nonreflecting boundary condition is easy to implement using standard force vector assembly over each boundary element on Γ , and does not disturb the symmetric and banded/sparse structure of the finite element matrix equations.

For axisymmetric radiation from a rigid baffle, the force vector (56) naturally specializes with the index restricted to $m=0$, i.e.,

$$\mathbf{F}_\Gamma(t) = \pi R^2 \sum_{n=2,4,\dots}^N q_{n0}(t) \int_0^{\pi/2} \mathbf{N}_p^T P_n(\cos \theta) \sin \theta d\theta, \quad (61)$$

where the system of equations (20) for \mathbf{z}_{n0} is driven by the radial modes given in (31).

VI. MIXED-TIME INTEGRATION ALGORITHM

Both implicit and explicit time marching schemes have been developed in Ref. 18 to integrate the semi-discrete equations associated with the NR1(N) form of the exact non-reflecting condition on a spherical boundary in free-space. These time-integration algorithms may be used to implement the boundary condition (23) or (24) for the half-space problem with no significant modifications. For NR2(N), *implicit* time marching procedures developed in Ref. 20 may be applied directly to the coupled system of equations (46) and (47). However, direct application of *explicit* time stepping schemes which uncouple the system of equations is not possible due to the difficulty in generating an accurate diagonal matrix \mathbf{C}_ψ appearing in the auxiliary equations (47). Fully explicit time discretization with diagonal matrices drastically

reduces computational cost and memory requirements. To obtain the efficiency of explicit time discretization without the need for a diagonal matrix \mathbf{C}_ψ , we present a new mixed explicit-implicit time integration method for solving the coupled system. Here, the interior finite element equations (46) are integrated explicitly in time and the auxiliary equations (47) on Γ are integrated implicitly in time. By treating the auxiliary equations implicitly, a very natural algorithm results which avoids the need to update either the pressure solutions or the auxiliary functions at intermediate time steps.

Let $\mathbf{F}^k = \mathbf{F}(t_k)$ be the force at time step $t_k = k\Delta t$. To compute the solution $\mathbf{p}^{k+1} = \mathbf{p}(t_{k+1})$, we apply the second-order accurate, *explicit* central difference method to the interior finite element matrix equations given in (46), with the result

$$\hat{\mathbf{M}}\mathbf{p}^{k+1} = \mathbf{R}^k, \quad (62)$$

with effective mass matrix

$$\hat{\mathbf{M}} = \frac{1}{\Delta t^2} \mathbf{M}_p + \frac{1}{2\Delta t} \mathbf{C}_p \quad (63)$$

and

$$\begin{aligned} \mathbf{R}^k = & \mathbf{F}^k - \mathbf{A}\boldsymbol{\psi}^k - \left(\mathbf{K} - \frac{2}{\Delta t^2} \mathbf{M}_p \right) \mathbf{p}^k \\ & - \left(\frac{1}{\Delta t^2} \mathbf{M}_p - \frac{1}{2\Delta t} \mathbf{C}_p \right) \mathbf{p}^{k-1}. \end{aligned} \quad (64)$$

The algebraic equations given in (62) are decoupled using standard lumping techniques to diagonalize \mathbf{M}_p and \mathbf{C}_p , e.g., using nodal (Lobatto) quadrature, row-sum technique, or the HRZ lumping scheme defined in Ref. 23. Using nodal lumping the effective mass $\hat{\mathbf{M}}$ is diagonal, and the system of equations (62) can be solved without factorizing a matrix; i.e., only matrix multiplications are required to obtain the right-hand-side effective load vector \mathbf{R}^k , after which the nodal pressures p_I can be updated using

$$p_I^{k+1} = \frac{R_I^k}{\hat{m}_{II}}, \quad (65)$$

where p_I^{k+1} and R_I^k denote the I th components of the vectors \mathbf{p}^{k+1} and \mathbf{R}^k , respectively, and \hat{m}_{II} is the I th diagonal element of the effective mass matrix obtained from the lumped mass and damping matrices. Furthermore, the matrix-vector products $\mathbf{A}\boldsymbol{\psi}^k$ and $\mathbf{K}_p\mathbf{p}^k$ can be evaluated at the element level by summing the contributions from each element to the effective load vector, without matrix assembly of \mathbf{A} or \mathbf{K}_p , rendering a highly efficient algorithm for large-scale wave propagation analysis.

The system of equations (47) for the auxiliary parameters $\boldsymbol{\psi}$ are not readily decoupled because of the difficulty generating an accurate diagonal matrix for \mathbf{C}_ψ using standard lumping techniques. In this case, we compute $\boldsymbol{\psi}^{k+1}$ using the *implicit*, second-order Adams-Moulton method (trapezoidal rule), i.e.,

$$\hat{\mathbf{C}} \boldsymbol{\psi}^{k+1} = \left(\mathbf{C}_\psi - \frac{\Delta t}{2} \mathbf{K}_\psi \right) \boldsymbol{\psi}^k + \frac{\Delta t}{2} \mathbf{A}^T (\mathbf{p}^{k+1} + \mathbf{p}^k), \quad (66)$$

with

$$\hat{\mathbf{C}} = \mathbf{C}_\psi + \frac{\Delta t}{2} \mathbf{K}_\psi. \quad (67)$$

Using a direct solver, and a constant time step Δt , the banded/sparse matrix $\hat{\mathbf{C}}$ is factorized only once into $\hat{\mathbf{C}} = \mathbf{L}\mathbf{D}\mathbf{L}^T$, where \mathbf{L} is a lower triangle and \mathbf{D} is a diagonal matrix. For constant wave speed c on the artificial boundary Γ , then $\mathbf{K}_\psi = (c/R)\mathbf{C}_\psi$ and the system (66) may be solved even more efficiently with the following procedure:

$$\text{Compute: } \mathbf{r}^{k+1} = \mathbf{A}^T (\mathbf{p}^{k+1} + \mathbf{p}^k), \quad (68)$$

$$\text{Solve: } \mathbf{C}_\psi \mathbf{y}^{k+1} = \mathbf{r}^{k+1}, \quad (69)$$

$$\text{Update: } \boldsymbol{\psi}^{k+1} = \frac{b}{a} \boldsymbol{\psi}^k + \frac{\Delta t}{2a} \mathbf{y}^{k+1}, \quad (70)$$

where $a = 1 + \gamma$, $b = 1 - \gamma$, and $\gamma = c\Delta t/2R$.

Similarly, the numerical solution \mathbf{z}_{nm}^{k+1} to the first-order system (20) and the solution q_{nm}^{k+1} to the first-order equation (42) may be computed concurrently using the implicit and unconditionally stable second-order Adams-Moulton method, i.e.,

$$\mathbf{B}_n \mathbf{z}_{nm}^{k+1} = \left(\mathbf{I} + \frac{\Delta t}{2} \mathbf{A}_n \right) \mathbf{z}_{nm}^k + \frac{\Delta t}{2} (\Phi_{nm}^{k+1} + \Phi_{nm}^k) \quad (71)$$

with

$$\mathbf{B}_n = \mathbf{I} - \frac{\Delta t}{2} \mathbf{A}_n, \quad (72)$$

and then update

$$q_{nm}^{k+1} = \frac{b}{a} q_{nm}^k + \frac{\gamma}{a} \tilde{\mathbf{c}}_n \cdot (\mathbf{z}_{nm}^{k+1} + \mathbf{z}_{nm}^k). \quad (73)$$

After the initial conditions are established, the complete mixed time-integration algorithm proceeds as follows for a fixed time step size, Δt :

- (1) Calculate effective loads at time t from (64).
- (2) Update the pressure field at time $t + \Delta t$ from (65).
- (3) Solve for auxiliary functions $\boldsymbol{\psi}$ at time $t + \Delta t$ from (66).
- (4) For each mode, solve the functions \mathbf{z}_{nm} at time $t + \Delta t$ from (71).
- (5) For each mode, update the functions q_{nm} at time $t + \Delta t$ from (73).
- (6) Update the time step, and repeat.

The key to the effectiveness of this algorithm is that the pressure update relies only on the auxiliary functions at the current time step, i.e., $\boldsymbol{\psi}^k$ and q_{nm}^k , and the update of the auxiliary functions relies only on the most recently computed pressure at time step t_{k+1} . The result is a very natural algorithm which avoids the need for intermediate updates between equations as would be the case in a staggered-step time integration. We also note that this mixed time-integration method also provides an efficient way to implement the local B_2 boundary condition of Bayliss and

Turkel¹⁰ in symmetric form; in this case the functions \mathbf{z}_{nm} and q_{nm} are not used, so that steps (4) and (5) may be skipped in the above algorithm.

VII. NUMERICAL STUDIES

In Refs. 18 and 20 numerical experiments for radiation from a sphere in free-space are presented which demonstrate the accuracy of the NRBC compared to steady-state analytical solutions and standard local absorbing boundary conditions. In this work, numerical studies of fully transient solutions for a circular piston transducer mounted in an infinite rigid planar baffle are presented. Numerical results are used to assess the accuracy of the mixed explicit-implicit time integration method and the NRBC defined in (40)–(42) for a half-space problem defined by a rigid baffle. Both sinusoidal and Gaussian pulse surface velocities are used to drive the transient solutions. A circular transducer radiating into an acoustic fluid is considered since this case has been widely studied and is important to many researchers.

A. Circular piston in a rigid planar baffle

We first consider a circular transducer of radius a , oscillating perpendicular to the plane of a rigid infinite baffle. The sound pressure field is determined by the wave equation and boundary conditions,

$$\frac{\partial p}{\partial z} = \begin{cases} -\rho_0 \dot{v}(t) H(t), & \text{on piston P} = \{0 \leq r \leq a, \theta = \pi/2\}, \\ 0, & \text{on baffle B} = \{r > a, \theta = \pi/2\}, \end{cases} \quad (74)$$

where $H(t)$ is the Heaviside (unit step) function. Here, z is the coordinate normal (perpendicular) to the piston and baffle, $v(t)$ is the normal velocity of the piston, and a superimposed dot denotes a time derivative.

The sound field $p(r, \theta, t)$ is rotationally symmetric about the z axis normal to the center of the piston and independent of φ . Since the problem is axisymmetric, it is convenient to introduce cylindrical coordinates (ρ, z) , where $\rho = r \sin \theta$ is the polar radius.

1. Sinusoidal input

We first assume that the piston velocity in Eq. (74) is sinusoidal for $t \geq 0$, with

$$v(t) = 1 - \cos \omega t, \quad (75)$$

resulting in a pressure gradient for $t \geq 0$,

$$\frac{\partial p}{\partial z} = -\rho_0 \omega \sin \omega t \quad \text{on piston P}, \quad (76)$$

where ω is a radian frequency.

The steady-state response along the z axis for the sinusoidal input (75) is available in a closed-form analytical solution given in Ref. 24:

$$p(0, z, t) = \text{Imag} \{ i \rho_0 c e^{i\omega t} [e^{-ikz} - e^{-ik\sqrt{z^2 + a^2}}] \}, \quad (77)$$

where c is the speed of sound and $k = \omega/c$ is the wave number. A study of the pressure amplitude on the axis of this piston reveals that the axial response exhibits strong interference effects, fluctuating between 0 and $2\rho_0 c$. These zero

pressure amplitudes occur at points z_m satisfying the condition

$$\frac{z_m}{a} = \frac{1}{m} \frac{a}{\lambda} - \frac{m}{4} \frac{\lambda}{a}, \quad (78)$$

where $\lambda = 2\pi/k$ and $m = \text{even}$.

Immediately after the circular piston (transducer) is switched on, the acoustic field will undergo a transient solution that is quite different from the steady-state condition; the radiation impedance consists of high-frequency components only, and then rapidly approaches the steady-state value. Since the rigid vibrating piston can be considered to be a distribution of point sources, the sound field occupies a region in space which is obtained by locating spheres of radius ct from each point on the piston. Thus the pressure on the piston itself is transient for the first $2a/c$ seconds, which is the time required for a signal to propagate from one edge of the piston across to the opposite edge. The pressure on the piston after the first $2a/c$ seconds is the same as in the steady state.

The transient sound field is available in a closed-form expression that can be integrated numerically.²⁵ The velocity potential field ϕ is represented as the time convolution of the normal velocity of the piston and a radiation impulse response:

$$\phi(\rho, z, t) = v(t) * h(\rho, z, t), \quad (79)$$

where the asterisk is used to denote the convolution operation, h is the velocity potential impulse response, t is the time, and a is the radius of the circular piston. The acoustic velocity in the medium is the negative gradient of the velocity potential ϕ , given by $\mathbf{v} = -\nabla\phi$. The pressure may then be obtained from the velocity potential using the relationship $p = \rho_0 \dot{\phi}$.

The impulse response function $h(\rho, z, t)$ is the time-dependent velocity potential field resulting from a Dirac impulsive z -velocity of the piston:²⁵

$$h(\rho, z, t) = \begin{cases} c, & \rho < a, \quad z < ct < R_1, \\ \frac{c}{\pi} \cos^{-1} \left(\frac{c^2 t^2 - z^2 + \rho^2 - a^2}{2\rho \sqrt{c^2 t^2 - z^2}} \right), & R_1 < ct < R_2, \\ 0, & \text{elsewhere,} \end{cases} \quad (80)$$

where $R_1 = \sqrt{z^2 + (a - \rho)^2}$ and $R_2 = \sqrt{z^2 + (a + \rho)^2}$ are the shortest and longest distances, respectively, from the observation point to the circumference of the piston.

For observation points on the z axis, $R_1 = R_2 = \sqrt{z^2 + a^2}$, so that h is a delayed rectangular pulse, and the time convolution given in (79) may be evaluated in closed-form:

$$p(0, z, t) = \rho_0 c \left\{ v \left(t - \frac{z}{c} \right) H \left(t - \frac{z}{c} \right) - v \left(t - \frac{\sqrt{z^2 + a^2}}{c} \right) H \left(t - \frac{\sqrt{z^2 + a^2}}{c} \right) \right\}, \quad (81)$$

where $H(t)$ is the Heaviside function. The pressure on the z axis is thus seen to consist of two signals of opposite

strength equal to the characteristic impedance $\rho_0 c$.

Since the problem is axisymmetric, it is sufficient to compute the finite element solution in the two-dimensional domain Ω defined by the (r, θ) plane for $0 \leq r \leq R$ and $0 \leq \theta \leq \pi/2$. The pressure field is approximated with four-node bilinear axisymmetric acoustic elements with a nonreflecting boundary applied to a quarter-circle truncation boundary $\Gamma := \{r = R, 0 < \theta \leq \pi/2\}$.

For the finite element solution, the truncation boundary Γ is positioned close to the radius of the piston at $R/a = 1.25$, resulting in a relatively small computational domain. The piston radius and wave speed are normalized such that $a = 1$ and $c = 1$. The calculation is then driven with a normalized frequency $\omega a/c = 4\pi$ on a mesh with 150 evenly spaced elements along the z axis from $0 \leq z \leq 1.25$, and 90 evenly spaced elements from $0 < \theta \leq \pi/2$. For this frequency and piston radius, the zero amplitude points along the z axis occur at $z_2 = 0.75$, and at the origin $z_4 = 0.0$.

A time-harmonic solution is obtained by starting from rest with initial data p_0 and \dot{p}_0 equal to zero and driving the solution to steady state with a time step $\Delta t = 0.003$ s. The mesh and time step size are small relative to the wavelength $\lambda = 0.5$, so that numerical error is due primarily to the radiation boundary condition on Γ .

Figure 2 shows time-dependent solutions at several observation points along the z axis obtained using the local B_1 and B_2 operators, and the nonreflecting boundary condition NR2(20), compared to the analytical solution. Figure 3 shows time histories at several locations on the truncation boundary Γ . The numerical solution obtained using NR2(20) can barely be distinguished from the analytical solution at all locations, including the interference point $z = 0.75$, where the steady-state amplitude is zero. The solution using B_2 exhibits relatively accurate solutions for points off-axis, however, significant amplitude errors occur for points on the z axis. The B_1 operator yields both large amplitude and phase errors, at all observation points except the piston origin $z = 0$, where all the operators accurately represent the solution, even during the zero steady-state amplitude phase. This interesting result is expected since at this location and for a piston of infinite radius it can be shown that the pressure and normal velocity at $z = 0$ are related by the simple plane-wave relationship $p = \rho_0 c v$, which is accurately represented by all of the boundary conditions considered.

The instantaneous error measured in L_2 norm on the z axis is defined as

$$E(t) = \left\{ \int_0^R [p^h(z, t) - p(z, t)]^2 dz \right\}^{1/2}, \quad (82)$$

where p^h is the approximate finite element solution, p is the exact steady-state solution, and $R = 1.25$. The magnitude of the absolute error $E(t)$ can be scaled by any constant, and will only serve as a means to compare the relative accuracy of the different boundary conditions. Figure 4 shows the instantaneous error on the piston axis and over the steady-state time interval $4.7 < t < 5.5$. The results illustrate the reduction in error obtained from NR2(N) by increasing the number of terms in the truncated series from $N = 10$ to $N = 20$. The accuracy of the NR2(20) solution is significantly improved

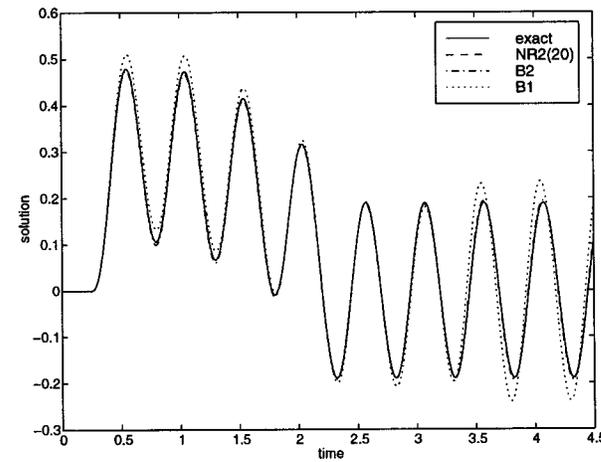
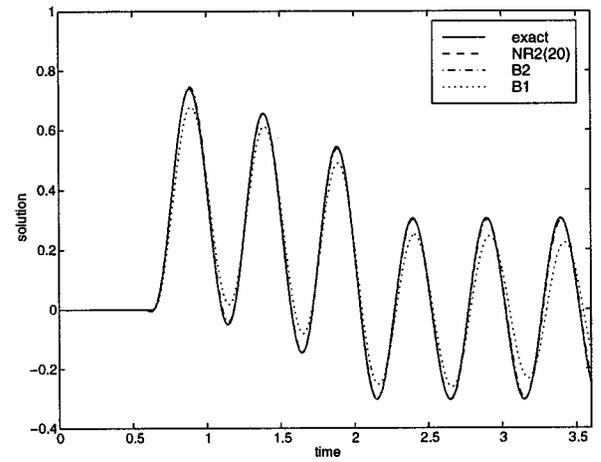
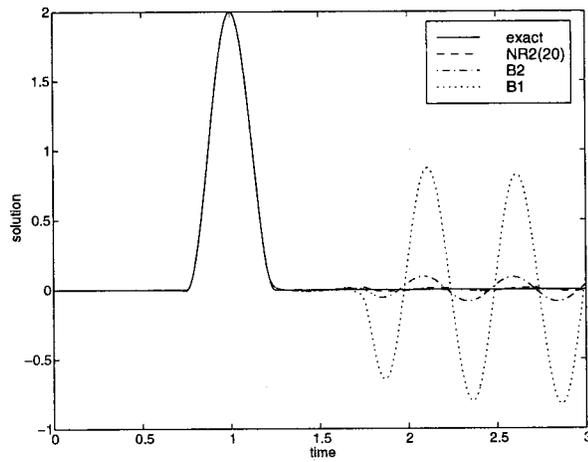
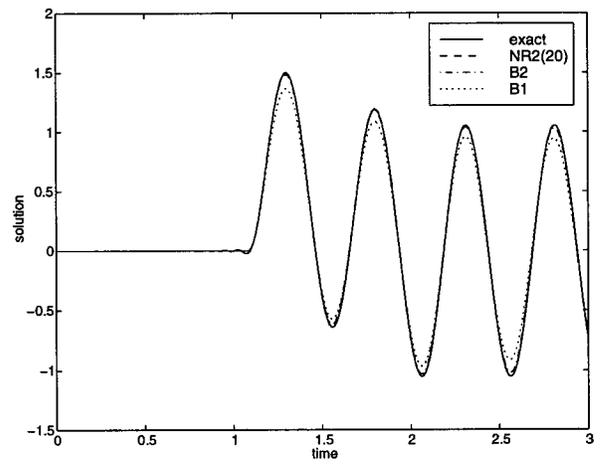
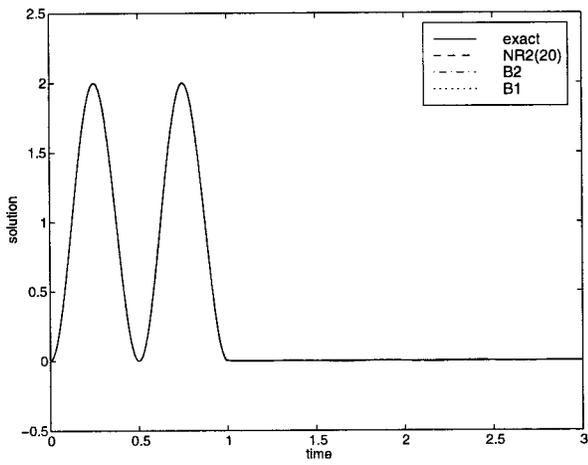


FIG. 2. Sinusoidal input. Time histories at on-axis observation points: (a) $z=0.0$; (b) $z=0.75$; (c) $z=1.0$.

FIG. 3. Sinusoidal input. Time histories at off-axis observation points on truncation boundary at $R=1.25$, and $\theta=30, 60, 90$ degrees.

over the local B_2 operator. We also observe that the error using the nonreflecting boundary condition NR2(20) reduced by an order of magnitude compared to the local B_1 boundary condition.

2. Gaussian pulse

We next study the response due to a transient pulse which excites a range of frequencies. The piston velocity is assumed to be the Gaussian pulse,

$$v(t) = e^{-0.5f_0^2(t-t_0)^2} H(t), \quad (83)$$

where $t_0=0.5$ s and $f_0=8$. The frequency spectrum of this wavelet is

$$v(\omega) = \frac{\sqrt{2\pi}}{f_0} e^{-0.5\omega^2/f_0^2}. \quad (84)$$

Figure 5 shows the pulse and its amplitude spectrum.

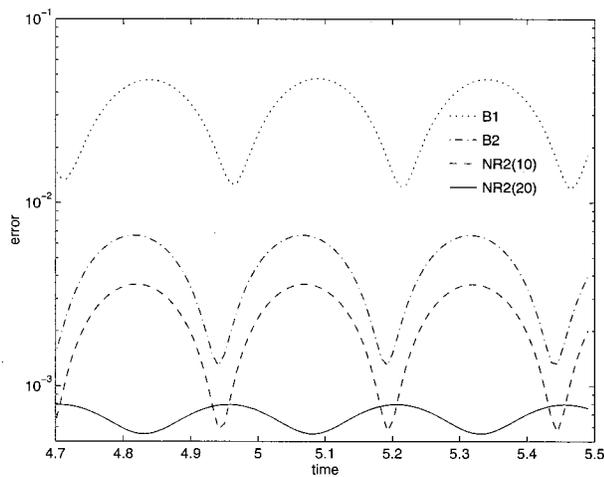


FIG. 4. Instantaneous error $E(t)$ along the z axis and at steady-state due to a sinusoidal radiating circular piston with normalized frequency $\omega a/c = 4\pi$ and truncation boundary radius $R/a = 1.25$. Results compared for local operators B_1 , B_2 , and boundary condition $\text{NR2}(N)$, with series truncation $N = 10$ and $N = 20$.

For this input, the pressure on the z axis consists of two Gaussian pulses of opposite amplitude. The time delay of the initial pulse corresponds to the propagation time from the center of the piston to the spatial point, and the time delay of the second pulse corresponds to the propagation time from the edge of the piston to the spatial point.

The computations are performed with the element size, and time-step unchanged from the previous study. Figures 6 and 7 show contours of the pressure field solution using $\text{NR2}(20)$ at several time steps. Figure 8 shows transient solutions at several observation points on the z axis. Figure 9 shows time histories at different locations on the truncation boundary Γ . Comparisons are made between the analytical solution and finite element solutions using the local operators B_1 and B_2 and the nonreflecting condition $\text{NR2}(N)$, with series truncation $N = 20$. The solution obtained using $\text{NR2}(20)$ can barely be distinguished from the analytical solution at all observation points. The early time response is accurately represented on the z axis using any of the boundary conditions studied. However, the numerical solution for B_1 exhibits large errors during the second pulse, both overshooting and undershooting the exact solution. The solution using B_2 shows significant error at the trailing end of the second pulse. On the truncation boundary, at angles off the piston axis, the local boundary operators exhibit spurious reflections during both the initial and secondary pulses. It is observed that both the local operators and nonreflecting condition $\text{NR2}(N)$ accurately represent the solution at the center of the piston $z = 0$, for all time.

VIII. CONCLUSIONS

An extension of the exact nonreflecting boundary condition (NRBC), first derived by Grote and Keller¹⁷ for the free-space problem, is formulated on a hemisphere for the time-dependent wave equation in a half-space defined by a planar baffle. Since the problem is symmetric about the infinite planar surface, we show that it is sufficient to restrict the spherical harmonic expansion which defines the NRBC to indices

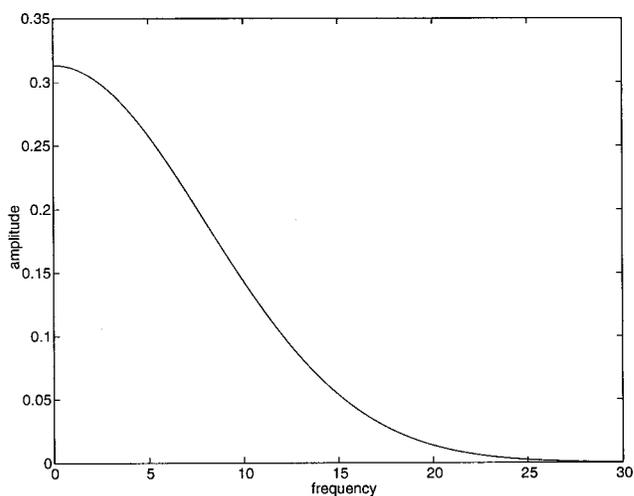
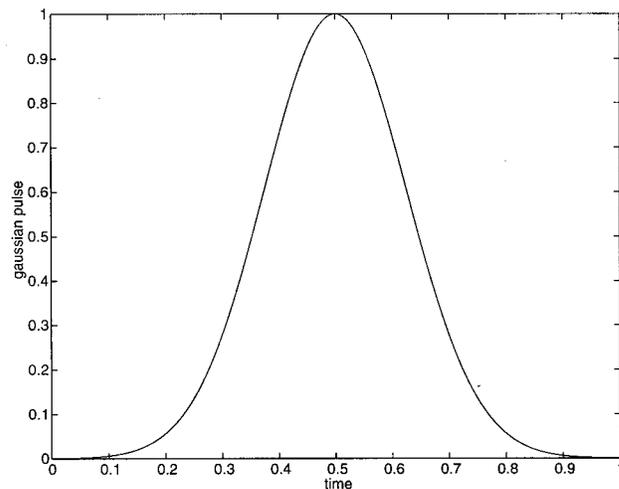


FIG. 5. Gaussian pulse and its Fourier transform: (a) Gaussian pulse versus time; (b) the amplitude spectrum versus frequency ω .

$n + m$ even, and scale the radial harmonics by a factor of 2, in order to satisfy the rigid baffle condition. In this case, we position the baffle normal (perpendicular) to a z axis in spherical coordinates. Since the symmetry condition is imposed with the Legendre function expansion in θ , the Fourier expansion in the circumferential angle φ may be used to efficiently model axisymmetric problems in two dimensions. Alternatively, the NRBC for the half-space problem may be formulated based on a z axis aligned parallel to the planar baffle. In this case the Fourier expansion in the circumferential angle is restricted to even functions in order to satisfy the rigid baffle symmetry condition. While this alternative form of the NRBC has the same number of operation counts and memory requirements in general three-dimensional problems, it cannot be used for the axisymmetric problem since the Fourier harmonics are restricted by the baffle condition. Symmetry conditions for the planar baffle may also be applied to recently developed asymptotic radiation boundary conditions given in Ref. 26 which have improved efficiency over the exact condition.

For the symmetric form of the modified NRBC we de-

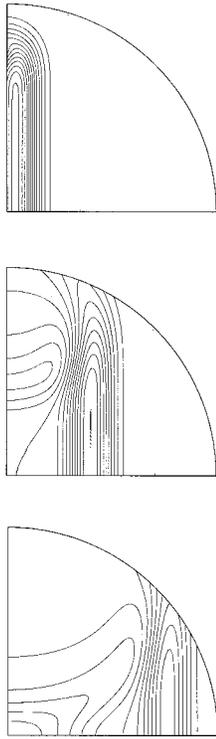


FIG. 6. Solution contours of pressure field using NR2(20) for transient radiation from a circular piston in infinite planar baffle with Gaussian normal velocity at time (a) $t=0.45$; (b) $t=0.9$; (c) $t=1.35$.

veloped a new mixed explicit-implicit time integration method which retains the efficiency of explicit time discretization for the finite element matrix equations, without the need for diagonal matrices in the auxiliary equations on the

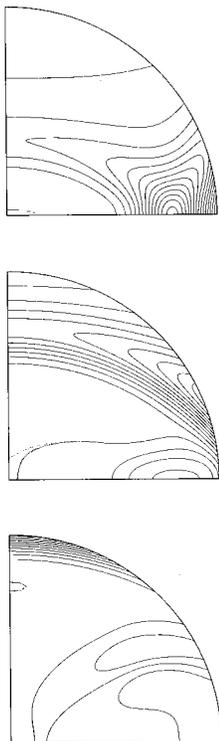


FIG. 7. Solution contours of pressure field using NR2(20) for transient radiation from a circular piston in infinite planar baffle with Gaussian normal velocity at time (a) $t=1.8$; (b) $t=2.25$; (c) $t=2.7$.

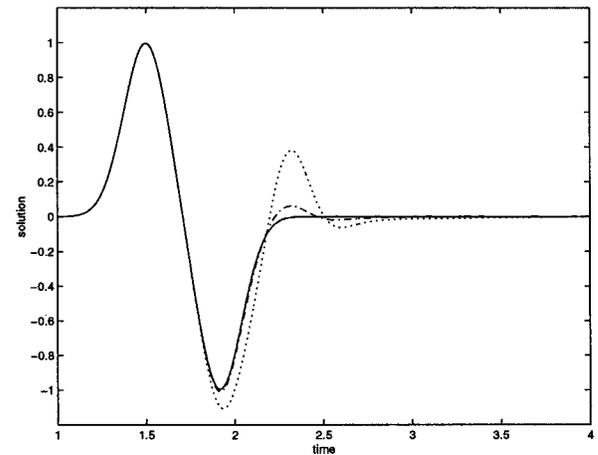
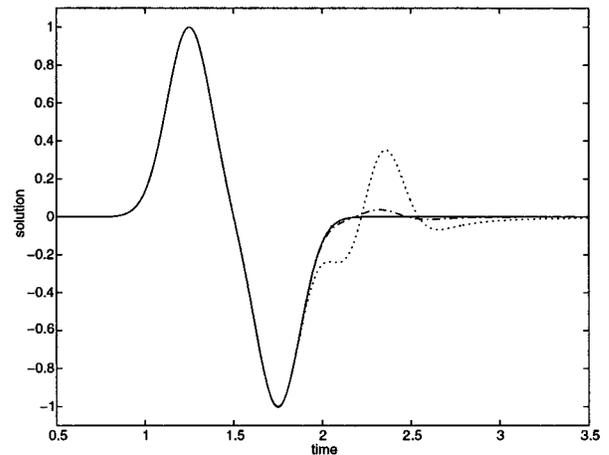
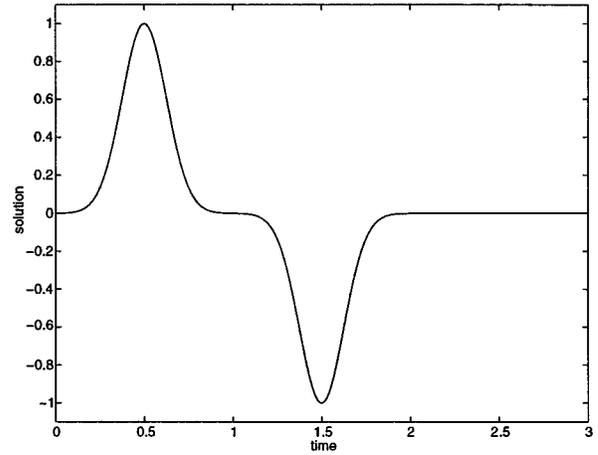


FIG. 8. Gaussian pulse. Time histories at on-axis observation points: (a) $z=0.0$; (b) $z=0.75$; (c) $z=1.125$. Solid lines denote analytic solution; dashed lines denote NR2(20); dash-dotted lines denote B_2 ; dotted lines denote B_1 .

artificial truncation boundary. Here, the interior finite element equations are integrated explicitly in time while the auxiliary equations on Γ are integrated implicitly in time. By treating the auxiliary equations implicitly, a very natural algorithm results which avoids the need to update either the pressure solutions or the auxiliary functions at intermediate

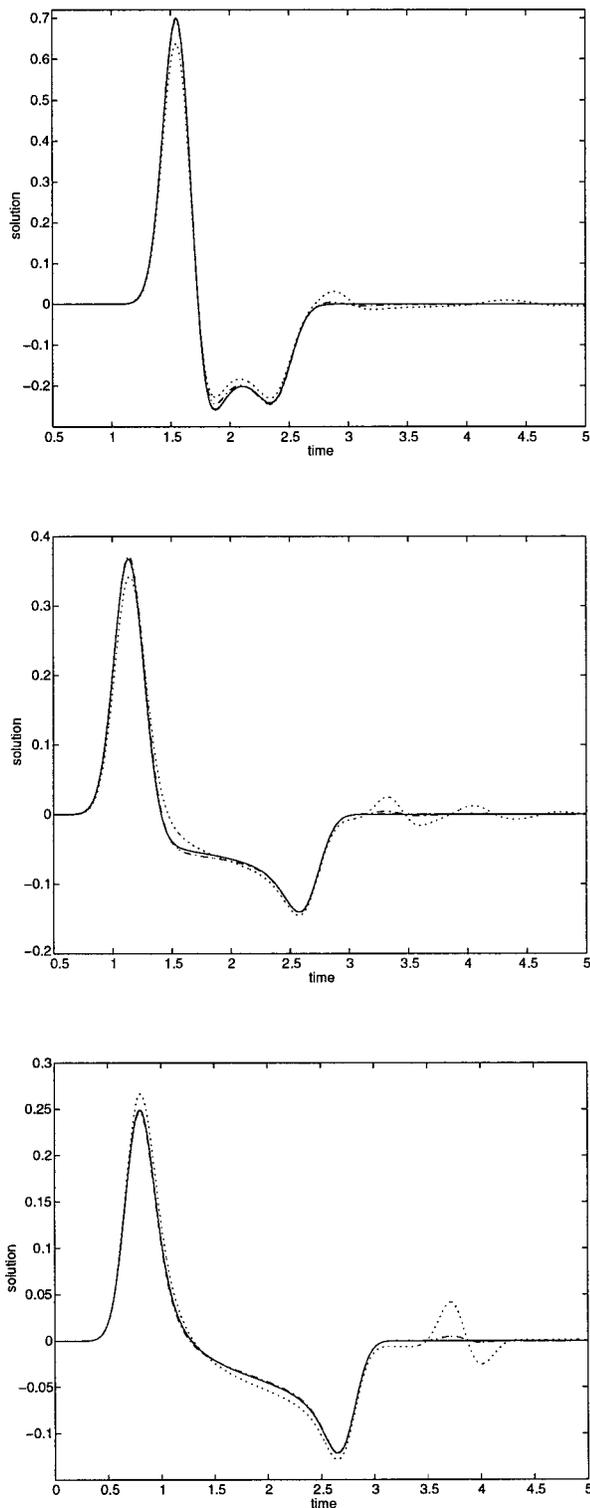


FIG. 9. Gaussian pulse. Time histories at off-axis observation points on truncation boundary at $R=1.25$: (a) $\theta=30$ degrees; (b) $\theta=60$ degrees; (c) $\theta=90$ degrees. Solid lines denote analytic solution; dashed lines denote NR2(20); dash-dotted lines denote B_2 ; dotted lines denote B_1 .

time steps and which retains the second-order accuracy of the underlying methods. The key to the effectiveness of the scheme is that the pressure update relies only on the auxiliary functions at the current time step, and the update of the auxiliary functions relies only on the most recently computed pressure at time step t_{k+1} . The result is a very natural and

highly efficient algorithm for large-scale wave propagation analysis which allows the pressure field to be updated without assembling or factoring the interior finite element matrices. This mixed time-integration method also provides an efficient way to implement the local B_2 boundary condition of Bayliss and Turkel.

Numerical examples of fully transient radiation from a circular piston transducer mounted in an infinite rigid planar baffle demonstrate the improved accuracy of the NRBC and the new mixed explicit-implicit time integration method compared to the first- and second-order local boundary conditions of Bayliss and Turkel. For a sinusoidal normal velocity input, the transient solution obtained using NR2(20) can barely be distinguished from the analytical solution at all points in the computational domain, including the difficult interference point on the piston axis where the steady-state amplitude is zero. The solution using the local second-order operator exhibits relatively accurate solutions for points off-axis, however significant amplitude errors occur for the difficult points on the piston axis. The local first-order operator yields both large amplitude and phase errors, at all observation points except the piston origin. Similar results are found for a Gaussian pulse input. The accuracy of the local operators may be improved by moving the artificial truncation boundary further away from the source of radiation at the expense of a larger computational domain with increased work. In other numerical studies,¹⁷ it is shown that the extra work in employing the local conditions with a large mesh is several times larger than the work required to compute the spherical harmonics in the exact NRBC on a smaller mesh. An important area of future work is to determine the benefit of increased accuracy using the modified NRBC with the extra expense in computing the auxiliary functions ψ at each node on the artificial boundary.

ACKNOWLEDGMENTS

Support for this work was provided by the National Science Foundation under Grant CMS-9702082 in conjunction with a Presidential Early Career Award for Scientists and Engineers (PECASE), and is gratefully acknowledged. We would also like to thank the anonymous reviewers for helping to clarify the alternative methods for satisfying the rigid baffle condition.

¹R. Bossut and J-N. Decarpigny, "Finite element modeling of radiating structures using dipolar damping elements," *J. Acoust. Soc. Am.* **86**, 1234–1244 (1989).

²F. 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).

³K. J. Hays-Stang and B. D. Cook, "Investigation of ultrasonic pulses with an optical probe," *J. Acoust. Soc. Am.* **97**, 2210–2217 (1995).

⁴G. Wojcik, J. Mould, F. Lizzi, N. Abboud, M. Ostromogilsky, and D. Vaughan, "Nonlinear modeling of therapeutic ultrasound," *IEEE Ultrasonics Symposium Proceedings* (1995), pp. 1–6.

⁵W. Friedrich, P. Krammer, and R. Lerch, "Finite element modeling of acoustic radiation from piezoelectric phased array antennas," *Proceedings IEEE Ultrasonics Symposium* (1990), Vol. 35, pp. 763–767.

⁶S. Tsynkov, "Numerical solution of problems on unbounded domains. A review," *Appl. Numer. Math.* **27**, 465–532 (1998).

⁷J. B. Keller and D. Givoli, "Exact non-reflecting boundary conditions," *J. Comput. Phys.* **82**, 172–192 (1989).

- ⁸L. L. Thompson, R. Huan, and C. Ianculescu, "Finite Element Formulation Of Exact Dirichlet-to-Neumann Radiation Conditions on Elliptic and Spheroidal Boundaries," *Proceedings of the 1999 ASME International Mechanical Engineering Congress and Exposition, Symposium on Computational Acoustics*, Nashville, TN, 14–19 Nov. 1999.
- ⁹D. Burnett, "A 3-D acoustic infinite element based on a generalized multipole expansion," *J. Acoust. Soc. Am.* **96**, 2798–2816 (1994).
- ¹⁰A. Bayliss and E. Turkel, "Radiation boundary conditions for wave-like equations," *Commun. Pure Appl. Math.* **33**, 707–725 (1980).
- ¹¹P. M. Pinsky and N. N. Abboud, "Finite element solution of the transient exterior structural acoustics problem based on the use of radially asymptotic boundary operators," *Comput. Methods Appl. Mech. Eng.* **85**, 311–348 (1991).
- ¹²P. M. Pinsky, L. L. Thompson, and N. N. Abboud, "Local high-order radiation boundary conditions for the two-dimensional time-dependent structural acoustics problem," *J. Acoust. Soc. Am.* **91**, 1320–1335 (1992).
- ¹³L. L. Thompson and P. M. Pinsky, "A space-time finite element method for structural acoustics in infinite domains, Part 2: Exact time-dependent non-reflecting boundary conditions," *Comput. Methods Appl. Mech. Eng.* **132**, 229–258 (1996).
- ¹⁴Q. H. Liu and J. Tao, "The perfectly matched layer (PML) for acoustic waves in absorptive media," *J. Acoust. Soc. Am.* **102**, 2072–2082 (1997).
- ¹⁵X. Yuan, D. Borup, J. W. Wiskin, M. Berggren, R. Eidsens, and S. A. Johnson, "Formulation and validation of Berenger's PML absorbing boundary for the FDTD simulation of acoustic scattering," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 816–822 (1997).
- ¹⁶M. J. Grote and J. B. Keller, "Exact non-reflecting boundary conditions for the time dependent wave equation," *SIAM (Soc. Ind. Appl. Math.) J. Appl. Math.* **55**, 280–297 (1995).
- ¹⁷M. J. Grote and J. B. Keller, "Nonreflecting boundary conditions for time dependent scattering," *J. Comput. Phys.* **127**, 52–65 (1996).
- ¹⁸L. L. Thompson and R. Huan, "Implementation of Exact Non-Reflecting Boundary Conditions in the Finite Element Method for the Time-Dependent Wave Equation," *Computer Methods in Applied Mechanics and Engineering* (in press).
- ¹⁹H. Lamb, *Hydrodynamics*, 4th ed. (Cambridge U. P., Cambridge, 1916), p. 517, Eq. 4.
- ²⁰L. L. Thompson and R. Huan, "Finite Element Formulation of Exact Non-Reflecting Boundary Conditions for the Time-Dependent Wave Equation," *Int. J. Numer. Methods Eng.* **45**, 1607–1630 (1999).
- ²¹L. L. Thompson and R. Huan, "Computation of far field solutions based on exact nonreflecting boundary conditions for the time-dependent wave equation," to appear in *Computer Methods in Applied Mechanics and Engineering*.
- ²²E. W. Hobson, *Spherical and Ellipsoidal Harmonics* (Cambridge U.P., Cambridge, 1955).
- ²³R. D. Cook, D. S. Malkus, and M. E. Plesha, *Concepts and Applications of Finite Element Analysis*, 3rd ed. (Wiley, New York, 1989).
- ²⁴P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Vol. II.
- ²⁵P. R. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," *J. Acoust. Soc. Am.* **49**, 1629–1638 (1971).
- ²⁶R. Huan and L. L. Thompson, "Accurate Radiation Boundary Conditions for the Time-Dependent Wave Equation on Unbounded Domains," to appear in *International Journal for Numerical Methods in Engineering*.

Characterization of multiple-sprung masses for wideband noise control

G. Maidanik and K. J. Becker

Naval Surface Warfare Center, Carderock Division (DTMB), 9500 MacArthur Boulevard, West Bethesda, Maryland 20817-5000

(Received 15 March 1999; revised 25 May 1999; accepted 4 August 1999)

The design of a wide frequency band neutralizer, vibration absorber and/or structural fuzzy, in the form of multiple-sprung masses, is extensively reported in the open literature. The action of the device is reported in terms of the joint point impedance of the sprung masses. This joint impedance is merely the sum over the impedances of the individual sprung masses at the common point to which the device is to be attached to a master structure. The normalized frequency bandwidth of a device composed of a single-sprung mass is proportional to the loss factor of that sprung mass. To increase this bandwidth, a device composed of more than one sprung mass, with distributed resonance frequencies, is utilized. To keep suppressed the undulations in the joint impedance of a set composed of a multiplicity of sprung masses, the loss factors are rendered larger than the normalized separations between adjacent antiresonance frequencies. This modal overlap condition, together with consideration of weight, are central to the design of the device. The analysis of the device is enriched by considering two distinct distributions of resonance frequencies for each set of sprung masses. Moreover, the ranges and parameters which specify that device are limited to reasonably moderate values; e.g., the useful frequency bandwidth of a given device is limited to one-third of its center frequency and the number of sprung masses in a device is restricted not to exceed one-score. In a set employing the first resonance frequency distribution, as the number of sprung masses is initially increased, an increase in the bandwidth is accompanied by an increase in the level of the joint impedance. As the number of sprung masses is further increased, the bandwidth and the level of the joint impedance become saturated. In a set incorporating the second resonance frequency distribution, an ongoing increase in the bandwidth, as the number of sprung masses increases, is accompanied by an ongoing decrease in the level of the joint impedance. The examination of these and other characteristics in the joint impedance of the sprung masses is provided by data obtained in computer experiments performed on a few selected sets of sprung masses. © 1999 Acoustical Society of America. [S0001-4966(99)00812-7]

PACS numbers: 43.20.Wd, 43.40.Tm, 43.50.Gf [DEC]

INTRODUCTION

Ever since structural fuzzies began reverberating within the noise control community, sprung masses, as vibration control devices, have become hot commodities.¹⁻⁵ The impedance $Z_j(\omega)$ that the j th sprung mass presents to a structure at the "point" x_j of attachment is simple enough. The impedance $Z_j(\omega)$ is constructed of the parallel combination of the impedance $i\omega M_j$ of the mass and the impedance $K_j/i\omega$ of the stiffness elements of the sprung mass; namely,

$$Z_j(\omega) = i\omega M_j \left[\{1 + \eta_j^2 - (\omega/\omega_j)^2\} - i(\omega/\omega_j)^2 \eta_j \right] \cdot \left[\{1 - (\omega/\omega_j)^2\}^2 + \eta_j^2 \right]^{-1};$$

$$(\omega_j)^2 = (K_{0j}/M_j); \quad K_j = K_{0j}(1 + i\eta_j),$$
(1)

where M_j , ω_j , and η_j are the mass, the resonance frequency, and the loss factor of the j th sprung mass.^{3,5} The loss factor is associated with the stiffness K_j of the spring. Consider initially a device consisting of a single-sprung mass.³⁻⁶ The resonance frequency of that sprung mass is designated ω_1 , yielding, thereby, a frequency of resonance given by

$$(\omega/\omega_1)^2 = (1 + \eta^2),$$
(2)

where η is the loss factor associated with that single-sprung-mass device.⁵ At and in the vicinity of the frequency of resonance, the impedance $Z(\omega)$ of the sprung mass is resistance controlled and of the approximate form

$$Z(\omega) \approx (\omega M/\eta); \quad |1 - (\omega/\omega_1)^2| \lesssim \eta \ll 1,$$
(3)

where M is the mass of this single-sprung mass device.⁵ The notion that a resistance-controlled impedance, that is inversely proportional to a loss factor, can be presented to a vibrating structure conjures a vision of the ultimate noise control device.^{3,5,6} Equation (3) states that in order to increase that resistance-controlled impedance, without the penalty of added weight, the loss factor η must be kept small. A small η implies a narrow-frequency-band device. To overcome this dilemma among others, e.g., the questions of fatigue and impedance matching, it is proposed to increase the number of sprung masses at a point.⁷ In this proposal the sprung masses are assigned a distribution of resonance frequencies to cover the desired wider frequency bandwidth.¹⁻⁵ Sprung masses that perceive a common response at their respective points of attachment can be considered to act at a point. A typical set of sprung masses at a point is sketched in Fig. 1(a).⁵ The expression for the joint impedance $Z_N(\omega)$

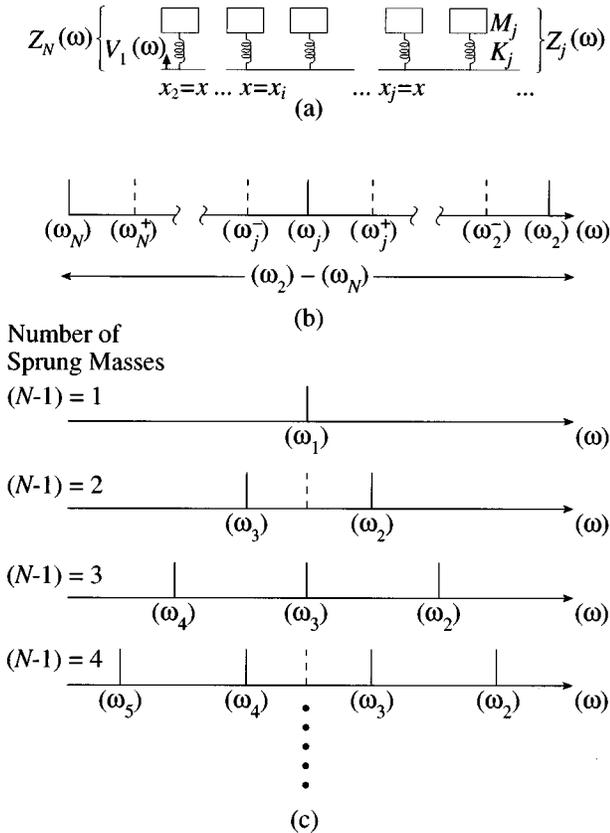


FIG. 1. (a) A generalized sketch of a set of sprung masses acting at a "point." (b) Locations of resonances designated by ω_j and adjacent antiresonances designated on the right by ω_j^+ and on the left by ω_j^- . (c) Number of sprung masses and their resonance frequency distributed with respect to the central frequency ω_1 .

that a set of $(N-1)$ -sprung masses presents at a *point* is readily derived, i.e., it is merely the sum of the individual impedances of the sprung masses,

$$Z_N(\omega) = \sum_{j=2}^N Z_j(\omega), \quad 2 \leq j \leq N, \quad (4)$$

where the index $j=1$ is reserved for identifying the master structure to which the $(N-1)$ -sprung masses are to be attached at a point.^{2-5,8} The overall normalized frequency bandwidth $\Delta\omega/\omega$ of the set of these sprung masses is related to the highest resonance frequency ω_2 and the lowest resonance frequency ω_N in the set. Formally, this normalized frequency bandwidth may be expressed in the form

$$(\Delta\omega/\omega) = \eta \delta_{2N} + (\Delta\omega/\omega_1)_{N \geq 3} (1 - \delta_{2N}), \quad N \geq 2, \quad (5)$$

where ω_1 is a normalizing frequency, η is the loss factor, and $\omega_1 \eta$ is the frequency bandwidth of a set composed of a single-sprung mass, δ_{2N} is the Kronecker delta and $(\Delta\omega/\omega_1)_{N \geq 3}$ is the frequency bandwidth of a set composed of two or more sprung masses [cf. Eqs. (2) and (3)]. It is noted that the normalized frequency bandwidth of a set composed of two or more sprung masses is

$$(\Delta\omega/\omega_1)_{N \geq 3} = [(\omega_2 - \omega_N)/\omega_1] + \sum_{j=2}^{N-1} [(\omega_j - \omega_{j+1})/\omega_1], \quad N \geq 3. \quad (6)$$

It took a few years, however, to relate appropriately the normalized *on the average* separation $\delta\omega/\omega$ between adjacent resonance frequencies and *on the average* loss factor η_c . Similar time elapsed to appropriately relate the normalized *local* separation $\delta\omega_j/\omega_j$ between adjacent antiresonance frequencies and the *local* loss factor η_j .^{5,9,10} These relationships define the condition of modal overlap.¹¹ The *on the average* condition of modal overlap is expressed in the form

$$\eta_c \geq (\delta\omega/\omega), \quad (7a)$$

$$(\delta\omega/\omega) = (N-2)^{-1} [(\omega_2 - \omega_N)/\omega_1], \quad N \geq 3,$$

where use is made of Eq. (6). On the other hand, the *local* condition of modal overlap is expressed in the form

$$\eta_j \geq (\delta\omega_j/\omega_j), \quad N \geq 3, \quad (7b)$$

and it is noted that the *local* separation $\delta\omega_j/\omega_j$ between adjacent antiresonance frequencies in a set composed of more than a single-sprung mass is

$$(\delta\omega_j/\omega_j) = [(\omega_j^+ - \omega_j^-)/\omega_1] (1 - \delta_{j2}) (1 - \delta_{jN}) + 2[(\omega_2 - \omega_2^-)/\omega_1] \delta_{j2} + 2[(\omega_N^+ - \omega_N)/\omega_1] \delta_{jN}, \quad 2 \leq j \leq N, \quad N \geq 3, \quad (8)$$

where ω_j^+ and ω_j^- are, respectively, the antiresonance frequencies on the higher and lower frequency sides of the resonance frequency ω_j , as illustrated in Fig. 1(b). Thus, the condition of modal overlap is defined with respect to either *on the average*, as stated in Eq. (7a), or *locally*, as stated in Eq. (7b). When the condition of modal overlap is not met, undulations in the joint impedance $Z_N(\omega)$, as a function of frequency, exist. The undulations are decreased as the condition of modal overlap is approached; the decreased undulations converge onto steady values as the condition of modal overlap is just reached.¹² When the condition of modal overlap is met, the undulations are suppressed altogether.

It has been speculated that the mean values of the undulations assume values that are independent of the loss factors of the individual sprung masses.^{3,4,12,13} Therefore, these values are commensurate with the values on which the undulations converge as the overlap condition is just met and beyond. This speculation is tested herein, and further in a companion paper, and the limits on its validity are formulated and revealed.¹⁴ Practitioners of noise control do not advocate large fluctuations in the response and, therefore, the condition of modal overlap is usually approached or reached in the design of a set of sprung masses. It is further held that an increase in damping, beyond that which just satisfies the modal overlap condition, is detrimental to this design.⁵ These notions are also to be tested and, if possible, quantified herein and in a companion paper.¹⁴ It is within this context that the influence of the damping, on the design of a set of sprung masses, is investigated. The investigation undertaken is facilitated by relating η and $\delta\omega/\omega$ to the designed *on the average* loss factor η_N , and η and $\delta\omega_j/\omega_j$ to the designed *local* loss factor η_{Nj} . The relationships are

$$\eta_N \Big\} = (b) \left[\eta \delta_{2N} + \left\{ \frac{(\delta\omega/\omega)}{(\delta\omega_j/\omega_j)} \right\} (1 - \delta_{2N}) \right], \quad (9a)$$

$$\eta_{Nj} \Big\} = (b) \left[\eta \delta_{2N} + \left\{ \frac{(\delta\omega/\omega)}{(\delta\omega_j/\omega_j)} \right\} (1 - \delta_{2N}) \right], \quad (9b)$$

where $\delta\omega/\omega$ and $\delta\omega_j/\omega_j$ are stated in Eqs. (7a) and (8), respectively, and it is convenient to define the *overlap factor* b in the form

$$b = \begin{cases} \frac{1}{3}, & \text{modal overlap condition is not satisfied,} \\ \frac{2}{3}, & \text{modal overlap condition is approached,} \\ \frac{3}{2}, & \text{modal overlap condition is just satisfied,} \\ 2, & \text{modal overlap condition is} \\ & \text{comfortably satisfied.} \end{cases} \quad (10)$$

To collate the investigation conducted in this paper, the distribution of the resonance frequencies of the sprung masses in a set is defined by

$$\begin{aligned} (\omega_j/\omega_1) &= [1 + \{2(j-1) - N\} \gamma_N]^{-1/2}; \\ \gamma_N(N-2) &< 1, \quad 2 \leq j \leq N, \end{aligned} \quad (11)$$

where γ_N is a universal parameter as yet to be assigned. Substituting Eq. (11) in Eqs. (6), (7a), and (8) one obtains

$$\begin{aligned} [(\omega_2 - \omega_N)/\omega_1] &= [\{1 - (N-2) \gamma_N\}^{-1/2} \\ &\quad - \{1 + (N-2) \gamma_N\}^{-1/2}], \end{aligned} \quad (12a)$$

$$\begin{aligned} (\delta\omega/\omega) &= [\{1 - (N-2) \gamma_N\}^{-1/2} - \{1 + (N-2) \gamma_N\}^{-1/2}] \\ &\quad \times (N-2)^{-1}, \end{aligned} \quad (12b)$$

$$\begin{aligned} (\delta\omega_j/\omega_j) &= [\{1 - \gamma_N[1 + \{2(j-1) - N\} \gamma_N]^{-1}\}^{-1/2} \\ &\quad - \{1 + \gamma_N[1 + \{2(j-1) - N\} \gamma_N]^{-1}\}^{-1/2}], \end{aligned} \quad (12c)$$

respectively. A practical limit needs to be imposed on the frequency bandwidth stated in Eq. (12a). This limit is then imposed in the form

$$[(\omega_2 - \omega_N)/\omega_1] \leq \frac{1}{3}, \quad N \geq 3, \quad (13)$$

which indirectly moderates the admissible values of γ_N and N . From Eqs. (12a) and (13) one finds that

$$[(\omega_2 - \omega_N)/\omega_1] \approx [\gamma_N(N-2)] \leq \frac{1}{3}, \quad N \geq 3. \quad (14)$$

Similarly, from Eqs. (12b) and (13) one finds that

$$(\delta\omega/\omega) \approx \gamma_N, \quad (15)$$

and from Eqs. (12c) and (13) one finds that

$$(\delta\omega_j/\omega_j) \approx (\omega_j/\omega_1)^2 \gamma_N, \quad (16)$$

where ω_j/ω_1 , and, therefore also $(\omega_j/\omega_1)^2$, are stated in Eq. (11). The investigation is enriched by considering two distinct distributions of resonance frequencies for each set of sprung masses. These distributions are distinguished by assigning to γ_N the two distinct forms

$$\gamma_N = \begin{cases} (\gamma/2)(N-1)^{-1}, & (17a) \\ \gamma_0, & (17b) \end{cases}$$

respectively. In Eqs. (17), γ and γ_0 are specified numerical increments that are independent of N . In both distributions, the center frequency ω_1 spans equal numbers of resonance

frequencies on the lower and higher frequency ranges. The center frequency ω_1 is the resonance frequency of a device that is composed of a single-sprung mass, e.g., as illustrated in Fig. 1(c). The center frequency ω_1 is also used as a normalizing frequency, e.g., in Eq. (5). It follows, from Eqs. (11), (17a), and (17b), that the difference between the two distributions of resonance frequencies is sustained by the duality of γ_N . However, if $N=N_0$, and γ and γ_0 are related in the form

$$(\gamma/2)(N_0-1)^{-1} = \gamma_0, \quad (17c)$$

then, for this value of N the value of γ_N is degenerate in the two distributions. It follows that for $N=N_0$ the corresponding parameters and quantities in the two distributions are identical and, therefore, for example, the joint impedances in the two distributions are identical for $N=N_0$.

The introduction of an auxiliary structure to a master structure often raises the question of weight. An equation for the designed weight of a set of sprung masses is stated relative to the effective weight of the master structure; namely

$$\sum_{j=2}^N M_j = (aM_1), \quad (18a)$$

where M_1 is an effective mass of the structure to which the set of sprung masses is to be attached and a is the ratio of the total mass in the set of sprung masses to this mass. An acceptable value for a is 10^{-1} . Were the individual masses in the sprung masses assumed to be equal, Eq. (18a) yields

$$M_j = M = (aM_1)[(N-1)]^{-1}, \quad M_j = M, \quad (18b)$$

where M is the mass of each of the sprung masses.

It is asserted that the loss factors η_N and η_{Nj} , as stated in Eq. (9), the distributions of the resonance frequencies, as stated in Eqs. (11) and (17), and the distribution of mass, as stated in Eq. (18b), adequately prescribe a set of sprung masses. Of course, the loss factors, the masses, and the distribution of the resonance frequencies may be defined variously from those stated in Eqs. (9), (11), (17), and (18). Here, however, the definitions in these equations prove convenient and varied enough. These definitions are, therefore, adopted as typical examples.¹⁰ Using these adopted definitions, computer experiments are conducted to examine the nature of the joint impedance $Z_N(\omega)$ for a few sets of sprung masses. Of particular interest in this examination are the influences on the characteristics of this joint impedance, by the variations and the restrictions just presented.

A major question remains: Given that the joint impedance $Z_N(\omega)$ of a set of sprung masses may be designed with specific characteristics, how is the potential noise control effectiveness of this set to be assessed? To answer this question, one needs to develop a noise control criterion that is based upon the performance of a specific set of sprung masses. The performance is to be assessed within the context of the response of the structure to which this set is attached, i.e., the master structure. Such a criterion may be expressed in terms of an overall gain $\Gamma_{1N}(\omega)$ defined in the form

$$\Gamma_{1N}(\omega) = |1 + [Z_N(\omega)/Z_1(\omega)]|^2, \quad (19)$$

where $Z_1(\omega)$ is the *point* impedance of the master structure in reference to the *point* of attachment.^{5,7} However, the detailed answer to the question just posed is deferred to a companion paper.¹⁴

I. COMPUTATIONS AND DISPLAYS OF NORMALIZED FREQUENCY BANDWIDTHS AND “ON THE AVERAGE” AND “LOCAL” LOSS FACTORS

The main events in this paper are related to exhibiting various criteria that can be used to design specific character-

istics and assess their benefits to a successful device. The device is a composition of several sprung masses at a *point*. (The device ultimately is intended to induce noise control in a master structure to which the device is to be attached at a point.) The sprung masses in a device are specified, as just established, in terms of a set of parameters which are defined for each of two distinct resonance frequency distributions. With the use of Eqs. (5), (9), (11), and (14–17), the dual sets of parameters are derived in the forms

$$(\Delta\omega/\omega) \approx \eta\delta_{2N} + \left\{ \frac{(\gamma/2)(N-2)(N-1)^{-1}}{\gamma_0(N-2)} \right\} (1 - \delta_{2N}), \quad (20a)$$

$$\eta_N \approx (b) \left[\eta\delta_{2N} + \left\{ \frac{(\gamma/2)(N-1)^{-1}}{\gamma_0} \right\} (1 - \delta_{2N}) \right], \quad (21a)$$

$$\eta_{Nj} \approx (b) \left[\eta\delta_{2N} + \left\{ \frac{(\gamma/2)(N-1)^{-1}}{\gamma_0} \right\} (\omega_j/\omega_1)^2 (1 - \delta_{2N}) \right], \quad (22a)$$

$$(\omega_j/\omega_1)^2 = \left\{ \frac{[1 + \{2(j-1) - N\}(\gamma/2)(N-1)^{-1}]^{-1}}{[1 + \{2(j-1) - N\}\gamma_0]^{-1}} \right\}. \quad (23a)$$

The first two parameters, i.e., $\Delta\omega/\omega$ and η_N , are displayed, as functions of N , in Figs. 2 and 3, and the last two parameters, i.e., η_j and $(\omega_j/\omega_1)^2$, are displayed as functions of (j) , in Figs. 4 and 5 for $N=6, 11$, and 18. Figures 2(a)–5(a) pertain to the first form of the resonance frequency distribution, as stated in Eq. (17a) and, in the same vein, Figs. 2(b)–5(b) pertain to the second form of the resonance frequency distribution, as stated in Eq. (17b). The appropriate values of $\{\eta, \gamma\}$ and $\{\eta, \gamma_0\}$ for Figs. 2–4 and the appropriate values of (γ, γ_0) for Fig. 5 are stated in the respective figure captions. Moreover, in Figs. 3 and 4 the value of the overlap factor b is selected to be $\frac{3}{2}$ [cf. Eq. (10)]. The asymptotic approaches of the displayed parameters in the dependent variable, N and j , respectively, are apparent in these figures. Another example: for $N \geq 3$, Eqs. (20a) and (20b) evaluate $\Delta\omega/\omega$ to be proportional to $(N-2)(N-1)^{-1}$ and to $N-2$, respectively. These proportionalities clearly emerge in Fig. 2(a) and (b), respectively. In particular, as N increases, the asymptotic plateau and the monotonic increase of $\Delta\omega/\omega$ are, respectively, discernible in these figures.

The numerical increments γ and γ_0 are selected in Figs. 2–5 to cause the dual distributions of the resonance frequencies to degenerate when $N=11=N_0$. Indeed, in Figs. 2–5, the values of two corresponding parameters, each pertaining to the one of the two distributions, are identical for $N=11$, e.g., the curve for $N=11$ in Fig. 5(a) is identical to the curve for $N=11$ in Fig. 5(b) [cf. Eq. (17c)].

II. NATURE OF THE IMPEDANCE OF A SET OF SPRUNG MASSES

The impedance $Z_j(\omega)$ of the j th sprung mass is stated in Eq. (1). The joint impedance $Z_N(\omega)$ of a set of sprung masses is stated, in terms of the impedances of the individual sprung masses, in Eq. (4). Utilizing Eqs. (9)–(12) to define the quantities and parameters of the set, the normalized joint impedance $\bar{Z}_N(\omega/\omega_1)$ of that set is derived for the *on the average* loss factor η_N in the form

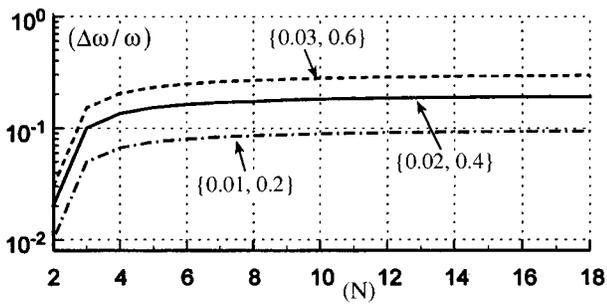
$$\begin{aligned} \bar{Z}_N(\omega/\omega_1) &= i(\omega/\omega_1)[a/(N-1)](1 + i\eta_N) \\ &\times \sum_{j=2}^N \{(1 + i\eta_N) - (\omega/\omega_1)^2 \\ &\times [1 + \{2(j-1) - N\}\gamma_N]\}^{-1}, \end{aligned} \quad (24a)$$

and for the *local* loss factor η_{Nj} in the form

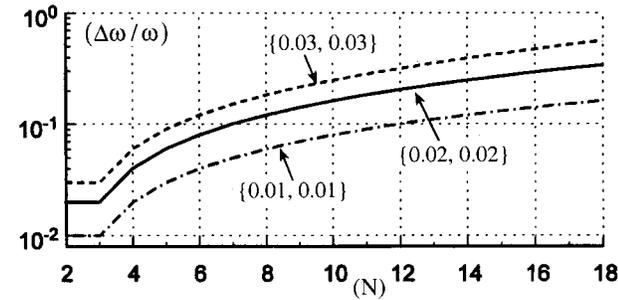
$$\begin{aligned} \bar{Z}_N(\omega/\omega_1) &= i(\omega/\omega_1)[a/(N-1)] \\ &\times \sum_{j=2}^N (1 + i\eta_{Nj}) \{(1 + i\eta_{Nj}) \\ &- (\omega/\omega_1)^2 [1 + \{2(j-1) - N\}\gamma_N]\}^{-1}, \end{aligned} \quad (24b)$$

where

$$\bar{Z}_N(\omega/\omega_1) = [Z_N(\omega)/(\omega_1 M_1)], \quad (25)$$



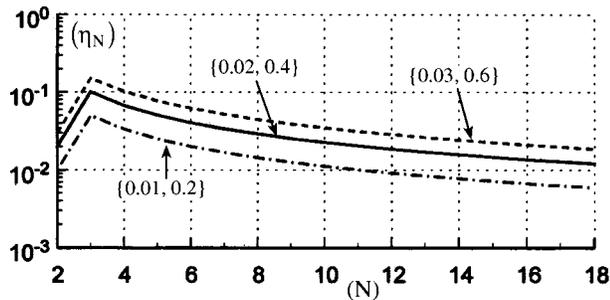
(a)



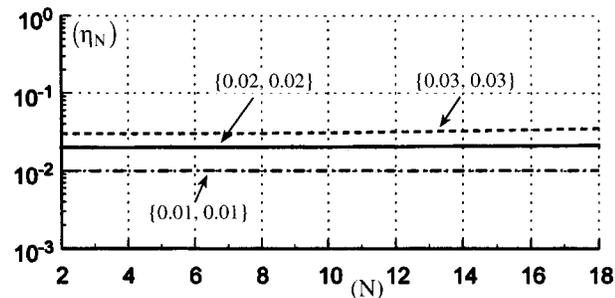
(b)

FIG. 2. The frequency bandwidth $\Delta\omega/\omega$ as a function of N . (a) $\{\eta, \gamma\}$ [cf. Eq. (20a)]. (b) $\{\eta, \gamma_0\}$ [cf. Eq. (20b)].

and $\omega_1 M_1$, being a real quantity, serves to normalize impedances. The behavior of the normalized joint impedance $\bar{Z}_N(\omega/\omega_1)$, as stated in Eq. (24), is investigated by presenting data that are acquired from computer experiments. As Eq. (24) indicates the computations of $\bar{Z}_N(\omega/\omega_1)$, as a function ω/ω_1 , for a set of sprung masses, require the specifica-

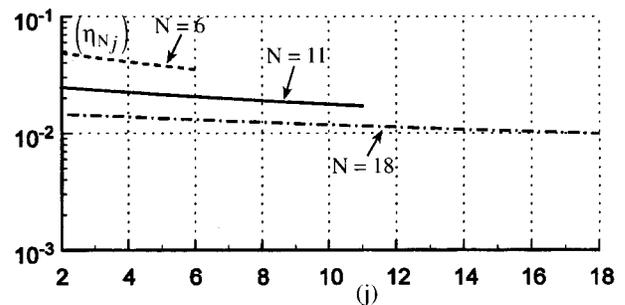


(a)

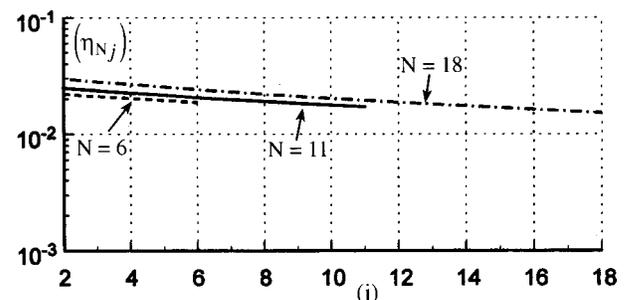


(b)

FIG. 3. The on the average loss factor η_N , as a function of N , with the overlap factor $b = \frac{3}{2}$ [cf. Eq. (10)]. (a) $\{\eta, \gamma\}$ [cf. Eq. (21a)]. (b) $\{\eta, \gamma_0\}$ [cf. Eq. (21b)].



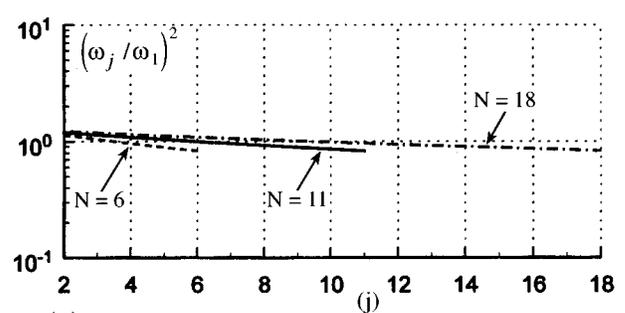
(a)



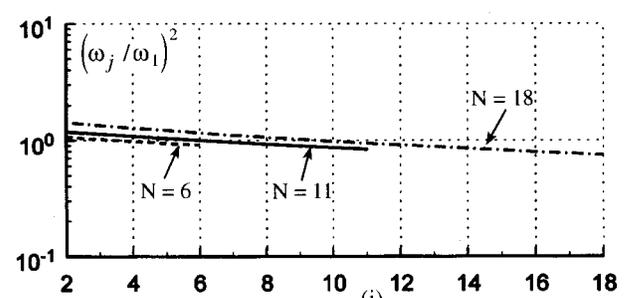
(b)

FIG. 4. The local loss factor η_{Nj} , as a function of j , with the overlap factor $b = \frac{3}{2}$ [cf. Eq. (10)]. (a) $\{\eta, \gamma\} = \{0.02, 0.4\}$ for $N = 6, 11,$ and 18 [cf. Eq. (22a)]. (b) $\{\eta, \gamma_0\} = \{0.02, 0.02\}$ for $N = 6, 11,$ and 18 [cf. Eq. (22b)].

tions of the number N , indicating that there are $(N-1)$ -sprung masses in the set; the mass ratio a ; the duality of the universal parameter γ_N , i.e., in terms of the numerical increments γ and γ_0 ; and, finally, the loss factors η_N and η_{Nj} . The natures of the loss factors η_N and η_{Nj} have

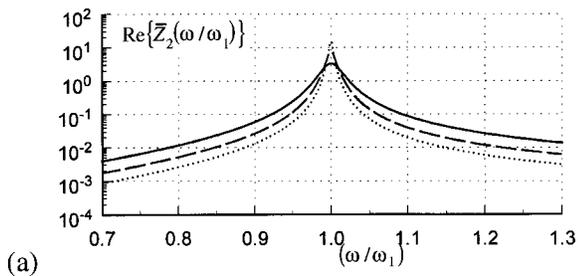


(a)

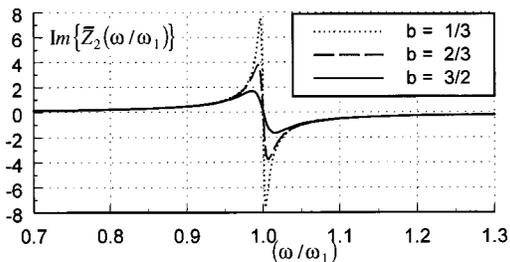


(b)

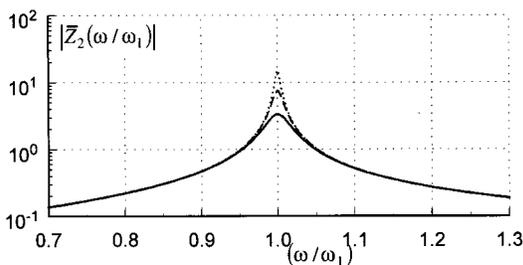
FIG. 5. The normalized quadratic resonance frequency $(\omega_j/\omega_1)^2$, as a function of j . (a) $\gamma = 0.4$ for $N = 6, 11,$ and 18 [cf. Eq. (23a)]. (b) $\gamma_0 = 0.02$ for $N = 6, 11,$ and 18 [cf. Eq. (23b)].



(a)



(b)

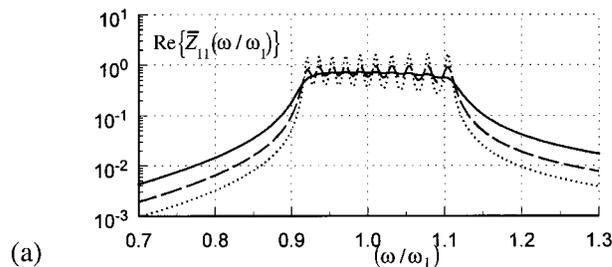


(c)

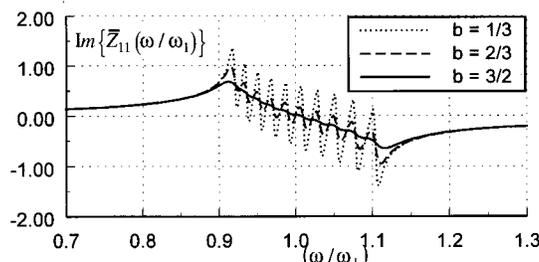
FIG. 6. Normalized joint impedance $\bar{Z}_2(\omega/\omega_1)$, as a function of ω/ω_1 , of a set of a single sprung mass. The sprung mass is defined by $a=0.1$ and $\eta=0.02$ [cf. Table I].

just been investigated in the preceding section [cf. Eq. (9) and Figs. 3 and 4]. The displays of the computations of the normalized joint impedance $\bar{Z}_N(\omega/\omega_1)$ are presented in Figs. 6–14 in terms of $\text{Re}\{\bar{Z}_N(\omega/\omega_1)\}$, $\text{Im}\{\bar{Z}_N(\omega/\omega_1)\}$ and/or $|\bar{Z}_N(\omega/\omega_1)|$, as functions of the normalized frequency ω/ω_1 ; e.g., these three normalized quantities for a set composed of a single-sprung mass are depicted in Fig. 6(a)–(c), respectively. The parametric values pertaining to each figure are identified and tabulated in Table I. In Figs. 6–13, the curves are assigned in reference to the overlap factor b , and the assignments are indicated on the figures. In Fig. 14, on the other hand, the curves are assigned in reference to the number $(N-1)$ of sprung masses in the device. By and large, the characteristics in the behavior of the normalized joint impedance $\bar{Z}_N(\omega/\omega_1)$ clearly emerge in the data presented in Figs. 6–14. A few of the characteristics are nonetheless emphasized and enumerated as a guide to the reader:

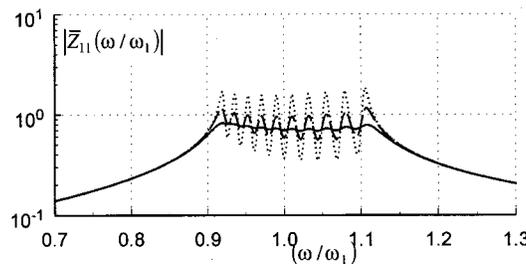
(i) Under careful scrutiny it is observed that Fig. 7, which employs the *on the average* loss factor, yields the less symmetric (uniform) undulations and, naturally, Fig. 8, which employs the *local* loss factors, yields undulations that are fairly symmetric (uniform). [The symmetry is about the center frequency ω_1 .] It



(a)



(b)

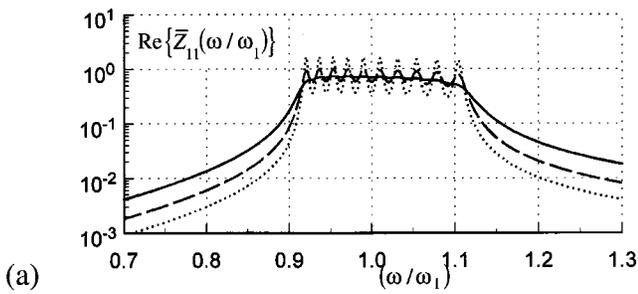


(c)

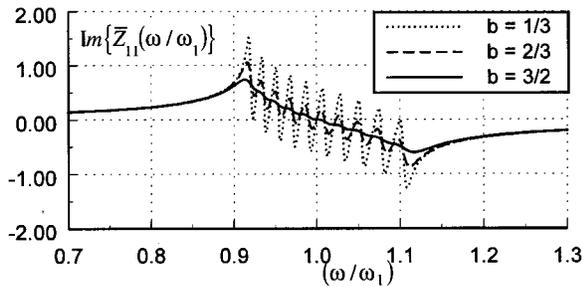
FIG. 7. Normalized joint impedance $\bar{Z}_{11}(\omega/\omega_1)$, as a function of ω/ω_1 , of a set of ten sprung masses. The set is further defined by $a=0.1$ and $\gamma=0.4$ and by “on the average loss factors” [cf. Eq. (9a) and Table I].

is noted, however, that in the range displayed in these figures, the difference between the two forms of loss factors, the *on the average* and *locally* based loss factors, is at best a factor of 2. From a noise control view a design that calls for the control of loss factors to better than a factor of (2) is impractical. In that sense then, the comparison for differences between Figs. 7 and 8 is merely of an academic interest.⁵

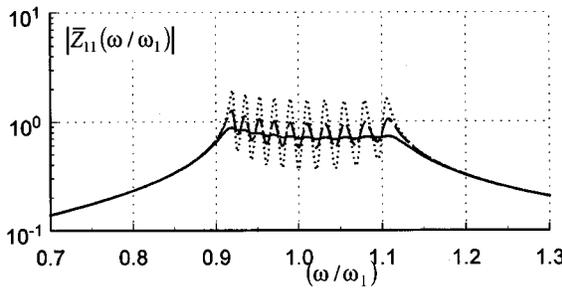
(ii) Further examination of Figs. 7 and 8 reveals that the excursions in the undulations are the more pronounced the less is the value of b . The undulations reveal the presence of individual contributions by the sprung masses to a dynamic quantity. The presence of undulations often invokes mean-value representation.¹² Thus, Figs. 7 and 8 suggest that the undulated values of the normalized joint impedance may not only possess mean values, but that these mean values converge onto common values that coalesce with those that are just suppressed as the overlap factor b reaches, from below, the neutral value of $\frac{3}{2}$ [cf. Eq. (10)]. If this observation is sustained by further evidence, this convergence may define the manner by which the undulations can be meaningfully averaged.¹²



(a)



(b)



(c)

FIG. 8. Normalized joint impedance $\bar{Z}_{11}(\omega/\omega_1)$, as a function of ω/ω_1 , of a set of ten sprung masses. The set is further defined by $a=0.1$ and $\gamma=0.4$ and by “local loss factors” [cf. Eq. (9b) and Table I].

(iii) In addition, when the undulations are subdued to suppression, $b \gtrsim \frac{3}{2}$, the signatures of the individual contributors are no longer apparent in the display of a dynamic quantity, e.g., in Figs. 7 and 8. Before bringing in this further and fresh evidence, a gross variation on the theme is introduced. In this variation Fig. 8(c) is repeated in Fig. 9, except that the overlap factor b is changed from the three standard values of $\frac{1}{3}$, $\frac{2}{3}$, and $\frac{3}{2}$ to the four values of $\frac{3}{2}$, 2, 4, and 8, respectively. An erosion of the level of $|\bar{Z}_N(\omega/\omega_1)|$ when the overlap condition is overly satisfied is clearly discernible in Fig. 9. Indeed, in the light of Fig. 9, one may deduce that the solid line curves in Figs. 7 and 8, which pertain to an overlap factor $b = \frac{3}{2}$, already bear tinges of erosion and hence the marginal deviations of these curves from strictly a common mean-value curve.¹² The erosion is initially most pronounced at the edges of the frequency band. Further increases in the loss factors, through increases in b , result in further erosion. The erosion is initially slight, when $b \gtrsim \frac{3}{2}$, but it magnifies with excessive increases in b . The overlap

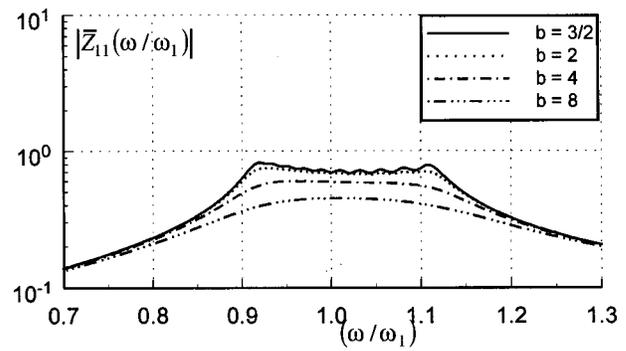
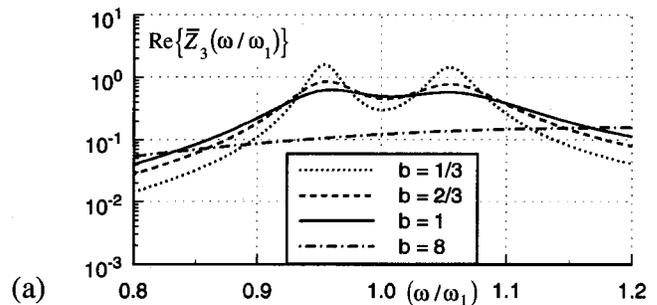
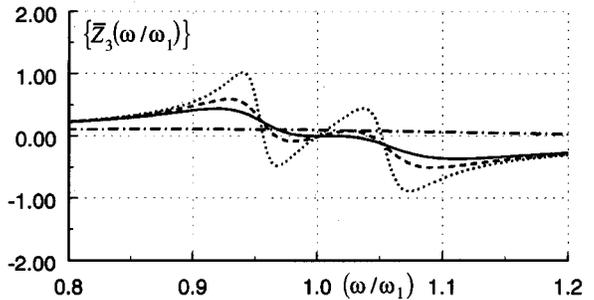


FIG. 9. Absolute value of the normalized joint impedance $\bar{Z}_{11}(\omega/\omega_1)$, as a function of ω/ω_1 , of a set of ten sprung masses. The set is further defined by $a=0.1$ and $\gamma=0.4$ and by “local loss factor” [cf. Eq. (9b) and Table I].

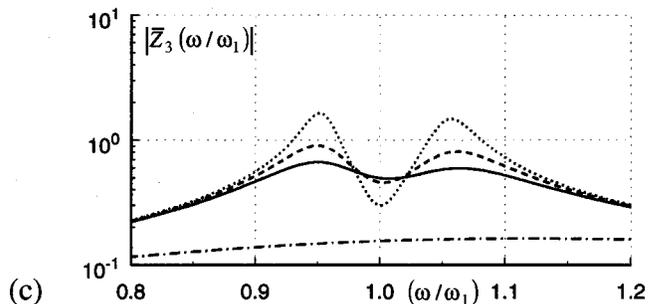
factor b is excessive when it is an order of magnitude or more; $b \gtrsim (8)$ [cf. Figs. 8 and 9]. To qualify and quantify the statement just presented, Fig. 8 is repeated in Figs. 10–13. In the first pair of figures, i.e., Figs. 10 and 11, the numerical increment (γ_N) is de-



(a)



(b)



(c)

FIG. 10. Normalized joint impedance $\bar{Z}_3(\omega/\omega_1)$, as a function of ω/ω_1 , of a set of two sprung masses. The set is further defined by $a=0.1$ and $\gamma=0.4$ and by “local loss factors” [cf. Eq. (9b) and Table I].

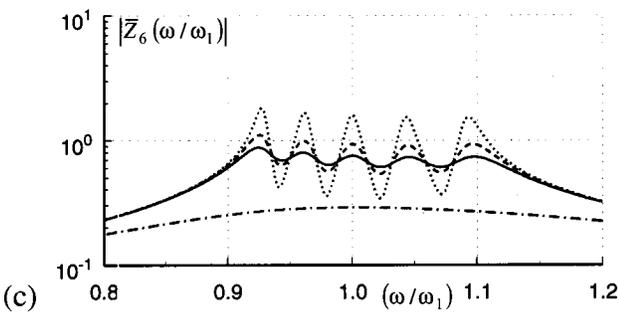
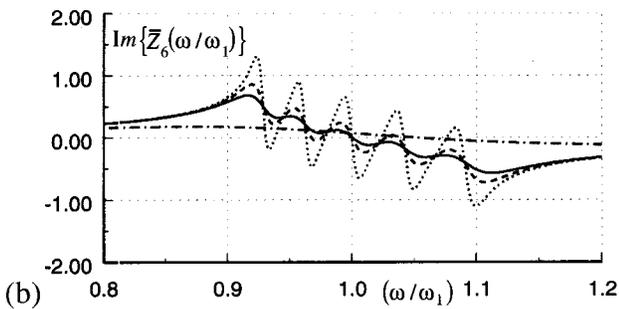
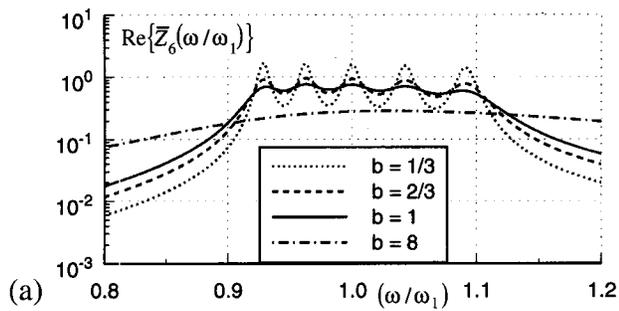


FIG. 11. Normalized joint impedance $\bar{Z}_6(\omega/\omega_1)$, as a function of ω/ω_1 , of a set of five sprung masses. The set is further defined by $a=0.1$ and $\gamma=0.4$ and by "local loss factors" [cf. Eq. (9b) and Table I].

finied by Eq. (17a) and in the second pair by Eq. (17b). In the two pairs, instead of setting the overlap factor $b = \frac{1}{3}, \frac{2}{3},$ and $\frac{3}{2}$ as in Fig. 8, the overlap factor b is set equal to $\frac{1}{3}, \frac{2}{3}, 1,$ and 8 . Also, in Figs. 10 and 12, N is equal to 3 and in Figs. 11 and 13, N is equal to 6; in Fig. 8, $N=11$.

- (iv) Finally, in Fig. 14, the values of the levels of the normalized joint impedance $|\bar{Z}_N(\omega/\omega_1)|$ are presented as a function of ω/ω_1 for a number of N 's, $N=2, 3, 4, 6, 8, 11, 14,$ and 18 . In Fig. 14 the overlap factor b is equal to the neutral value of $\frac{3}{2}$ [cf. Eq. (10)]. In Fig. 14(a), γ_N takes the definition in Eq. (17a) and in Fig. 14(b), γ_N takes the definition in Eq. (17b). Again, γ and γ_0 are selected as 0.4 and 0.02, respectively, so that N_0 as defined in Eq. (17c), is 11. Obvious in Fig. 14(a) is the dramatic decrease in the mean levels of the joint impedance, which is accompanied by a dramatic increase in the frequency bandwidth, as N increases from 2 to 3. [This phenomenon, which is exemplified in Fig. 14(a) is commensurate

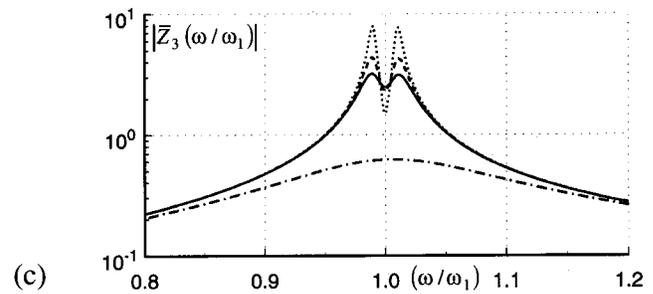
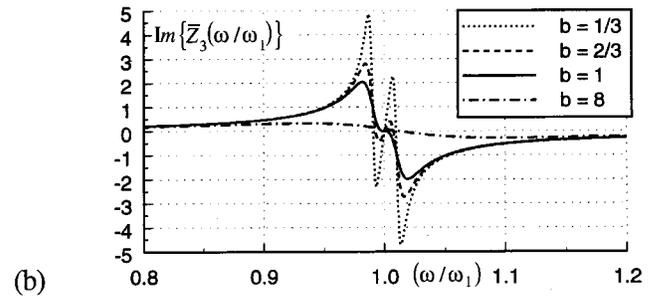
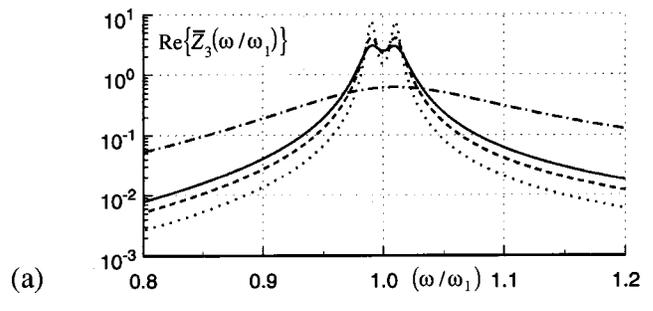


FIG. 12. Repeats Fig. 10 except that $\gamma_0=0.02$ replaces $\gamma=0.4$ (cf. Table I).

with a dramatic change in the loss factor, as N changes from 2 to 3. This kind of a change in the loss factor is depicted, for example, in Fig. 3(a).] The increase of the frequency bandwidths and the accompanied increase in the mean levels of $|\bar{Z}_N(\omega/\omega_1)|$ within these frequency bands, as N is increased beyond 3 is clearly exhibited in Fig. 14(a). Then, the saturation of both, the bandwidths and the mean levels, with further increases in N , is also clearly exhibited in Fig. 14(a). On the other hand, the continuous broadening of the frequency bandwidths and the accompanied diminishing in the mean levels of $|\bar{Z}_N(\omega/\omega_1)|$ within those frequency bands, as N increases, emerges loud and clear in Fig. 14(b). Nonetheless, despite these fundamental differences between the trends in the curves in Fig. 14(a) and (b), the curve in Fig. 14(a) for $N=11=N_0$ and the corresponding curve in Fig. 14(b), are identical as dictated by Eq. (17c).

A brief recapulation: A set of sprung masses at a point constitutes a device that is intended to control the response of a master structure to which the device may be attached at a point. The potential viability of the device is assessed in

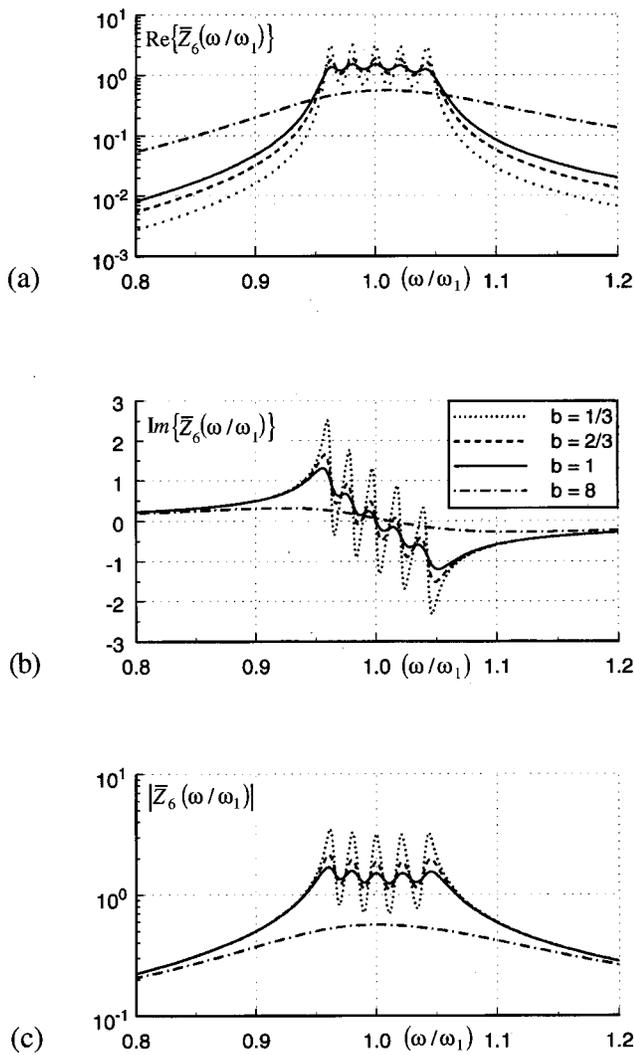


FIG. 13. Repeats Fig. 11 except that $\gamma_0=0.02$ replaces $\gamma=0.4$ (cf. Table I).

terms of its joint impedance $Z_N(\omega)$. The joint impedance is equal to the sum of the individual impedances of the sprung masses. The joint impedance is dependent on the number $(N-1)$ of sprung masses in the set; on the mass ratio (a) , which accounts for the weight of the device relative to the effective weight of a master structure in the vicinity of the point attachment; on the loss factor associated with the individual sprung masses (this association is based either *on the average* loss factor η_N or *local* loss factor η_{N_j} ; and last, but not least, on the distribution of the resonance frequencies of the sprung masses, which is prescribed in terms of the universal parameter γ_N . The loss factors are related to the distributions of resonance frequencies in terms of the overlap factor b . When b is less than $\frac{3}{2}$ undulations exist in the joint impedance as a function of the normalized frequency ω/ω_1 , revealing, thereby, the individual presence of the sprung masses. As b increases to $\frac{3}{2}$ the undulations are suppressed; these undulation-free values of the joint impedance serve as the mean values of the undulated values of the joint impedance. When b increases beyond $\frac{3}{2}$, erosions occurs in the values of the joint impedance; the erosion is the more severe the higher the value of the overlap factor b . Two distinct resonance frequency distributions are defined; each pertains

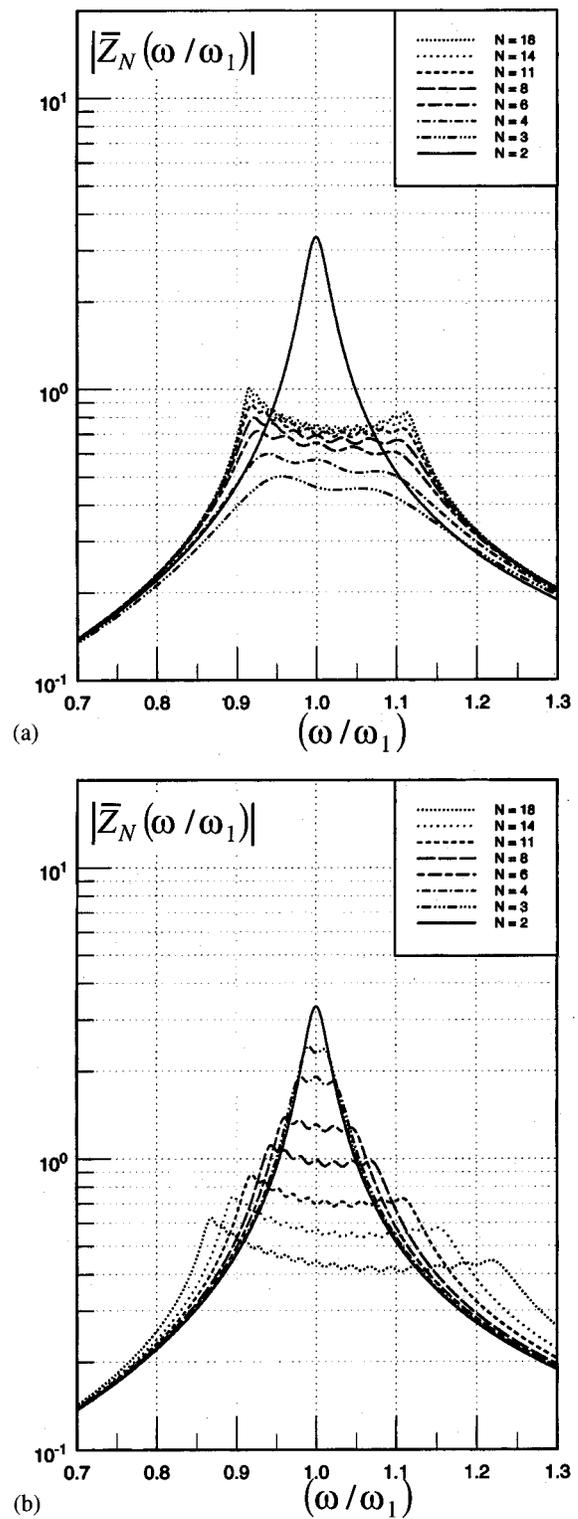


FIG. 14. (a) Absolute value of the normalized joint impedance $\bar{Z}_N(\omega/\omega_1)$, as a function of ω/ω_1 , for a number of sets of sprung masses; $N=2, 3, 4, 6, 8, 11, 14,$ and 18 . A set is further defined by $a=0.1$ and $\gamma=0.4$ and by ‘‘local loss factors’’ [cf. Eq. (9b) and Table I]. (b) Repeats (a) except that $\gamma_0=0.02$ replaces $\gamma=0.4$ [cf. Table I].

to a particular form of the universal parameter γ_N . The dual forms for γ_N define two resonance frequency distributions that raise two contrasting behaviors for the joint impedance as a function of ω/ω_1 . This duality in the distributions is degenerate for a set containing (N_0-1) -sprung masses; e.g.,

TABLE I. A parametric identification for Figs. 6–14.

Figure	N	a	η	γ	γ_0	Re	Im	$ \bar{Z}_N $	η_N	η_{Nj}
						$\{\bar{Z}\}$	$\{\bar{Z}_N\}$			
6	2	0.1	0.02	a	b	c
7	11	0.1	0.02	0.4	...	a	b	c	✓	...
8	11	0.1	0.02	0.4	...	a	b	c	...	✓
9	11	0.1	0.02	0.4	✓	...	✓
10	3	0.1	0.02	0.4	...	a	b	c	...	✓
11	6	0.1	0.02	0.4	...	a	b	c	...	✓
12	3	0.1	0.02	...	0.02	a	b	c	...	✓
13	6	0.1	0.02	...	0.02	a	b	c	...	✓
14(a)	2,3,4,6,8, 11,14,18	0.1	0.02	0.4	✓	...	✓
14(b)	2,3,4,6,8, 11,14,18	0.1	0.02	...	0.02	✓	...	✓

the joint impedance $Z_{N_0}(\omega)$ of the set is the same in the two resonance frequency distributions. The definitions of the prescribing parameters and the parametric dependences and their influences on the characteristic behaviors of the joint impedance of a device composed of a multiplicity of sprung masses is investigated by computer experiments. The data from these experiments are displayed graphically in Figs. 6–14. Finally, the potential value of a set of sprung masses, as a device for wideband noise control, is assessed and discussed in a companion paper.¹⁴ That companion paper utilizes and expands the analytical and experimental material just presented herein.

¹C. Soize, “Probabilistic structural modeling in linear dynamic analysis of complex mechanical systems,” *Rech. Aerosp.* **1986-3**, 23–48 (1986); “A model and numerical method in the medium frequency range for vibroacoustic predictions using the theory of structural fuzzy,” *J. Acoust. Soc.*

Am. **94**, 849–865 (1993); R. Ohayon and C. Soize, *Structural Acoustics and Vibrations* (Academic, San Diego, 1998).

²A. D. Pierce, V. W. Sparrow, and D. A. Russell, “Fundamental structural acoustics idealizations for structures with fuzzy internals,” *Trans. ASME, J. Vib. Acoust.* **117**, 1–10 (1995); A. D. Pierce, “Fuzzies but simpler analytic models for physical acoustics and structural acoustics,” *J. Acoust. Soc. Am.* **103**, 2954 (1988).

³M. Strasberg and D. Feit, “Vibration damping of large structures induced by attached small resonant substructures,” *J. Acoust. Soc. Am.* **99**, 335–344 (1996).

⁴M. Strasberg, “Continuous structures as ‘fuzzy’ substructures,” *J. Acoust. Soc. Am.* **100**, 3456–3459 (1996).

⁵M. Brennan, “Wideband vibration neutralizer,” *Noise Control Eng. J.* **45**, 201–207 (1997).

⁶G. Maidanik, “Power dissipation in a sprung mass attached to a master structure,” *J. Acoust. Soc. Am.* **98**, 3527–3533 (1995).

⁷K. U. Ingard, *Fundamentals of Waves and Oscillations* (Cambridge U.P., Cambridge, 1988).

⁸G. Maidanik and K. J. Becker, “Noise control of a master harmonic oscillator coupled to a set of satellite harmonic oscillators,” *J. Acoust. Soc. Am.* **104**, 2628–2637 (1998).

⁹M. Strasberg, “Is the dissipation induced by ‘fuzzy’ substructures real or apparent,” *J. Acoust. Soc. Am.* **102**, 3130 (1997).

¹⁰Maidanik and K. J. Becker, “Various loss factors of a master harmonic oscillator that is coupled to a number of satellite harmonic oscillators,” *J. Acoust. Soc. Am.* **103**, 3184–3195 (1998).

¹¹R. H. Lyon, *Statistical Energy Analysis of Dynamic Systems: Theory and Applications* (MIT, Cambridge, 1975); R. H. Lyon and R. G. DeJong, *Theory and Application of Statistical Energy Analysis* (Butterworth-Heinemann, Boston, 1995).

¹²E. Skudrzyk, “The mean-value method of predicting the dynamic response of complex vibrators,” *J. Acoust. Soc. Am.* **67**, 1105–1135 (1980).

¹³R. J. Nagem, I. Veljkovic, and G. Sandri, “Vibration damping by a continuous distribution of undamped oscillators,” *J. Sound Vib.* **207**, 429–434 (1997).

¹⁴G. Maidanik and K. J. Becker, “Criteria for designing multiple-sprung masses for a wideband noise control,” *J. Acoust. Soc. Am.* **106**, 3119–3127 (1999).

Criteria for designing multiple-sprung masses for wideband noise control

G. Maidanik and K. J. Becker

Naval Surface Warfare Center, Carderock Division (DTMB), 9500 MacArthur Boulevard, West Bethesda, Maryland 20817-5000

(Received 3 September 1998; revised 25 May 1999; accepted 4 August 1999)

In a companion paper the characterization of a multiple-sprung masses design for wideband noise control is presented. The characterization is largely conducted in terms of the *point* impedance of a set of sprung masses. The sprung masses in the set are collectively acting at a *point*. In that companion paper the dependencies of the joint impedance on the number of sprung masses, the modal overlap factors, the total mass, and the resonance frequency distribution are of particular interest. In the present paper this characterization of a set of sprung masses is utilized to define a number of design criteria that determine the potential viability of the set as a noise control device. In the final analysis, the device must be assessed in an *in situ* setting. In that setting the device is tested in terms of the *overall gain*. The determination of the overall gain requires, in addition to the joint impedance of the device, the impedance of the hosting master structure at the *point* of attachment. (The device is meant to control the noise in the master structure.) It is argued that although the overall gain is the final judge, the potential *viability* of a device may be a useful critique of the intended performance of the device. In this vein, a *viability* and a *criterion of promise* are defined to assist with the design processes of this noise control device. In assessing the *viability* and in satisfying the *criterion of promise*, the desired characteristics in a set of sprung masses can be judicially selected. © 1999 Acoustical Society of America. [S0001-4966(99)00912-1]

PACS numbers: 43.20.Wd, 43.40.Tm, 43.50.Gf [DEC]

INTRODUCTION

In a companion paper the characteristics of a set of sprung masses, as a device for achieving a wideband noise control, is presented both analytically and experimentally.¹ The experiments are conducted on computer models and the data are illustrated graphically. The characterization is based largely on the behavior of the joint impedance of a set of sprung masses at a *point*.¹ The dependence of the joint impedance upon the number of sprung masses, the modal overlap factors, the total mass, and the resonance frequency distribution, all of which prescribe the set, are of particular interest. To enrich the characterization, two distinct resonance frequency distributions are employed and contrasted. The increase of the frequency bandwidth, which is central to both papers, entails an amplification to saturation in level of the joint impedance in one distribution and a continual diminution in level in the other.¹ One may suspect that when the frequency bandwidth of the device is increased the amplification or the diminution in the level of the joint impedance, within its viable frequency band, may cause the noise control effectiveness of the device to rise or to fall, respectively. In order to assess this potential noise control effectiveness, a noise control criterion is proposed in terms of the *overall gain*. As already mentioned in Ref. 1, given that the joint impedance $Z_N(\omega)$ of a set of $(N-1)$ -sprung masses may be designed with specific characteristics, the overall gain $\Gamma_{1N}(\omega)$ may be defined in the form

$$\Gamma_{1N}(\omega) = |1 + [Z_N(\omega)/Z_1(\omega)]|^2, \quad (1)$$

where $Z_1(\omega)$ is a test *point* impedance for a master structure in reference to the *point* of attachment.^{2,3} A number of com-

puter experiments are performed to examine the overall gain $\Gamma_{1N}(\omega)$ that some selected sets of sprung masses may be able to muster. The characteristics of the joint impedance $Z_N(\omega)$ in these experiments are drawn from material presented in Ref. 1. The significance of the utilization of the material presented in Ref. 1 is emphasized by calling directly upon equations and figures in Ref. 1. These equations and figures are recognized by assigning to them, in this paper, the numeral I; e.g., Eq. (6a) and Fig. 6(b) of Ref. 1 are stated herein merely as Eq. (I.6a) and Fig. I.6(b), respectively. On the other hand, the test impedance $Z_1(\omega)$, which is assigned to the master structure at a *point* in these experiments, is to be one of the two distinct forms

$$\{Z_1(\omega)/(\omega_1 M_1)\} = \begin{cases} i(\omega/\omega_1), & (2a) \\ 1 = (4/\sqrt{3})\{(mhc_l)/(\omega_1 M_1)\}, & (2b) \end{cases}$$

where in Eq. (2a) the point impedance $Z_1(\omega)$ is that of the mass M_1 and in Eq. (2b) the point impedance $Z_1(\omega)$ is that of an unbounded plate and in that plate m , h , and c_l are the mass per unit area, the thickness, and the longitudinal speed, respectively.³ Typically, these two forms are sketched in Fig. 1(a) and (b), respectively. A case in which the master structure itself is a sprung mass is presented extensively under a separate cover.⁴ Here it is briefly mentioned that this case is accounted for by

$$\{Z_1(\omega)/(\omega_1 M_1)\} = i(\omega/\omega_1)[(1 - (\omega_1/\omega)^2)(1 + i\eta_1)], \quad (2c)$$

where ω_1 and η_1 are the resonance frequency and the loss factor of the master-sprung mass, respectively [cf. Eq. (I.1)]. A sketch of this dynamic system is given in Fig. 1(c). In contrast to using Eqs. (2a) and (2b), only a cursory use is

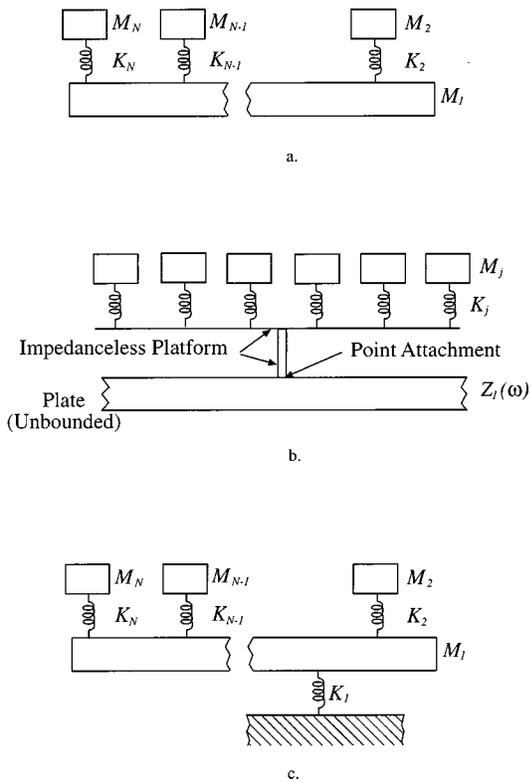


FIG. 1. (a) A master structure as a mass (cf. Fig. I.1). (b) A master structure as an unbounded plate. The device is attached at a "point" through an impedanceless platform. (c) A master structure possessing at a point the impedance of an oscillator (a sprung mass).

made of Eq. (2c) in a subsequent computation and display.

It has been argued that to assess the effectiveness of a noise control device, constituted by a set of sprung masses acting at a *point*, one is obliged to invoke the overall gain. Nonetheless, it may be useful to consider the viability of this set, not exclusively by its *in situ* performance, but merely by its potential promise as a noise control device. In this less imperative case, the *viability* may be assessed without the immediate examination of the overall gain; a specific behavior of the set may, in itself, be a useful design process. In this vein, a *parameter of promise* and a *criterion of promise* are defined to be useful *measures* of the potential noise control that a set of sprung masses at a point may harbor. The *criterion of promise* is akin to the "gain bandwidth product" introduced and defined in Ref. 2.

I. OVERALL GAIN OF A SET OF SPRUNG MASSES ATTACHED TO A MASTER STRUCTURE

The overall gain $\Gamma_{1N}(\omega)$ for a set of sprung masses is stated Eq. (I.19) and, again, in Eq. (1). Utilizing Eqs. (1) and (2), the overall gain may be stated in the form

$$\Gamma_{1N}(\omega/\omega_1) = |1 + [\bar{Z}_N(\omega/\omega_1)/\bar{Z}_1(\omega/\omega_1)]|^2, \quad (3)$$

where the normalized point impedance $\bar{Z}_1(\omega/\omega_1)$ in the master structure may be cast in the forms

$$\bar{Z}_1(\omega/\omega_1) = \begin{cases} i(\omega/\omega_1), & (4a) \\ 1, & (4b) \\ i(\omega/\omega_1)[1 - (\omega_1/\omega)^2(1 + i\eta_1)]. & (4c) \end{cases}$$

It is recognized that an overall gain $\Gamma_{1N}(\omega/\omega_1)$ that is unity is gainless, less than unity is ineffective, and greater than unity is effective in gaining a beneficial noise reduction. Clearly, given an adequate frequency bandwidth, the more the overall gain exceeds unity in that frequency band the better is the *in situ* performance of the referenced noise control device. Indeed, the employment of a multiplicity of sprung masses is an attempt to widen the frequency bandwidth and sustain, on balance, an effective overall gain within that wider frequency band. Figures 2(a), (b), 3(a), and (b) depict the overall gain, as a function of ω/ω_1 , derived by converting data presented in Fig. I.14(a) and Fig. I.14(b), à *l'â* Eq. (3) under the guidance of Eqs. (4a) and (4b), respectively. The identical solid line curves in Figs. 2(a) and 3(a) and in Figs. 2(b) and 3(b) pertain to a set consisting of a single-sprung mass that is attached at a point on a master structure. The test impedances at this attachment point are as stated in Eqs. (4a) and (4b), respectively. These identical curves in Figs. 2(a) and 3(a) exhibit a narrow ridge in the overall gain. That ridge peaks at $\omega/\omega_1 = 1$ and is accompanied by a nadir at a slightly higher frequency. At the peak, the overall gain is effective; at the nadir it is ineffective. Analogously, in Figs. 2(b) and 3(b), the ridge, in the identical solid line curves, is also narrow and it peaks at $\omega/\omega_1 = 1$. In variance, however, at and in the vicinity of $\omega/\omega_1 = 1$, the peak stands alone. At the peak the overall gain is effective; outside the narrow band, which is centered on the peak, the overall gain bottoms on unity. Thus, for a set consisting of a single-sprung mass, Figs. 2 and 3 readily exhibit an effective overall gain. However, this effectiveness is confined to a narrow frequency band. Indeed, the narrowness of the frequency band in which the effective overall gain lies, i.e., the solid line curves in Figs. 2 and 3, is the reason for proposing a device incorporating a set composed of a multiplicity of sprung masses.^{2,5-8} A query arises in this connection: Is such a proposal viable in the sense that an extended bandwidth can be attained with an overall gain that is effective in that wider frequency band? To prepare the answer to this question, the data presented in Figs. 2 and 3 are more extensively studied. In this vein, it is observed that in Fig. 2 a pronounced decrease in the effectiveness of the overall gain, accompanied by a pronounced increase in the frequency bandwidth, occurs immediately as N increases from 2 to 3.¹ As N increases beyond 3, a recovery commences in the effectiveness and the frequency bandwidth increases. However, both processes approach saturation as N increases further.¹ On the other hand, in Fig. 3 the decrease in effectiveness and the increase in frequency bandwidth, of the overall gain, occur gradually and continuously with increase in N .^{1,2} More specific observations reveal that with an increase in N beyond 3, Fig. 2(a) exhibits more prominent frequency bandwidths, a shift in the peaks to lower frequencies and an enhancement in the effectiveness. (One is reminded that in Figs. 2-4, as in Fig. I.14, the increase in N is identified by assigned curve designations; the higher the N , the less complex is the curve designation.) It is also observed that as N increases, the ineffectiveness in the overall gain becomes more severe and the nadirs are shifted to lower frequencies. However, the prominence in the bandwidths, the

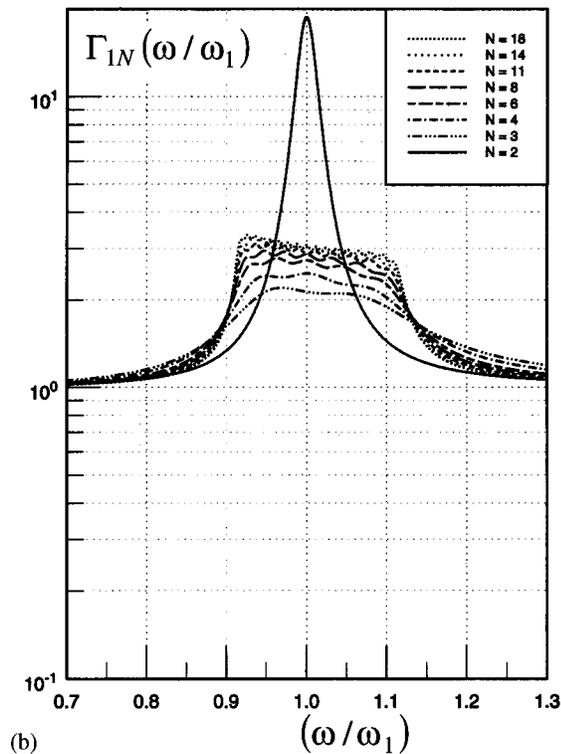
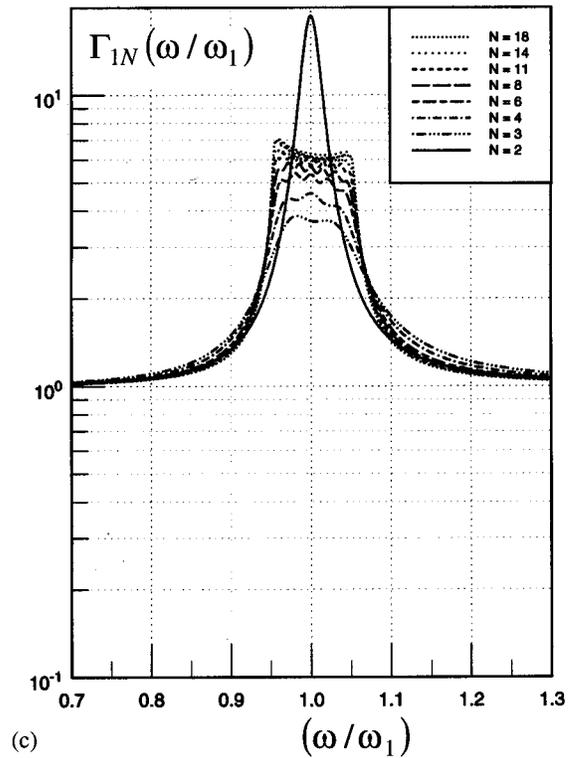
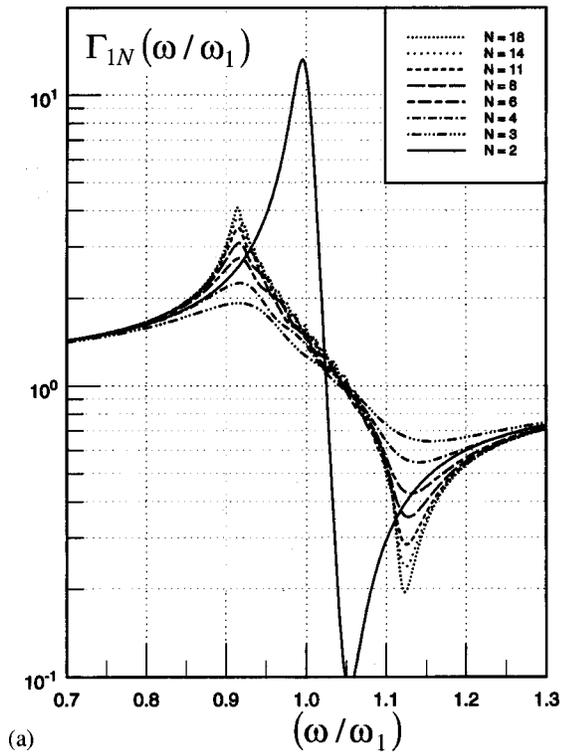


FIG. 2. The overall gain $\Gamma_{1N}(\omega/\omega_1)$ as functions of ω/ω_1 [cf. Eq. (3)]. (a) Computed $\Gamma_{1N}(\omega/\omega_1)$ from Fig. I.14(a) and $\bar{Z}_1(\omega/\omega_1) = i(\omega/\omega_1)$ [cf. Eq. (4a)]. (b) Computed $\Gamma_{1N}(\omega/\omega_1)$ from Fig. I.14(a) and $\bar{Z}_1(\omega/\omega_1) = 1$ [cf. Eq. (4b)]. (c) Repeats (b) except that γ is halved, from 0.4 to 0.2. The value of $N_0 = 6$ [cf. Eq. (I.17c)].

frequency shifts of the peaks and nadirs, and the enhancement in the effectiveness and the ineffectiveness all saturate in Fig. 2(a) with a further increase in N .¹ This saturation renders the changes, with increase in N , largely insignificant once N exceeds the critical number N_c . Therefore, $(N_c - 1)$ is a *useful upper limit* on the number of sprung masses, in a device, for attaining a potentially effective gain. Adding complications, such as increasing the number of sprung masses, without definitive benefits is not a good design pro-

cedure. Similarly, with increases in N , Fig. 3(a) exhibits a steady broadening of the frequency bandwidths, a shift in the frequencies of the peaks and the nadirs, and a reduction in the effectiveness and the ineffectiveness. The continuity of these changes, with increases in N , cannot be beneficially maintained beyond the critical number N_c ; when N increases beyond N_c , the reduced effectiveness in the overall gain is rendered unacceptable. It emerges, then, that except for specific purposes, in general, the overall gain presented in Figs.

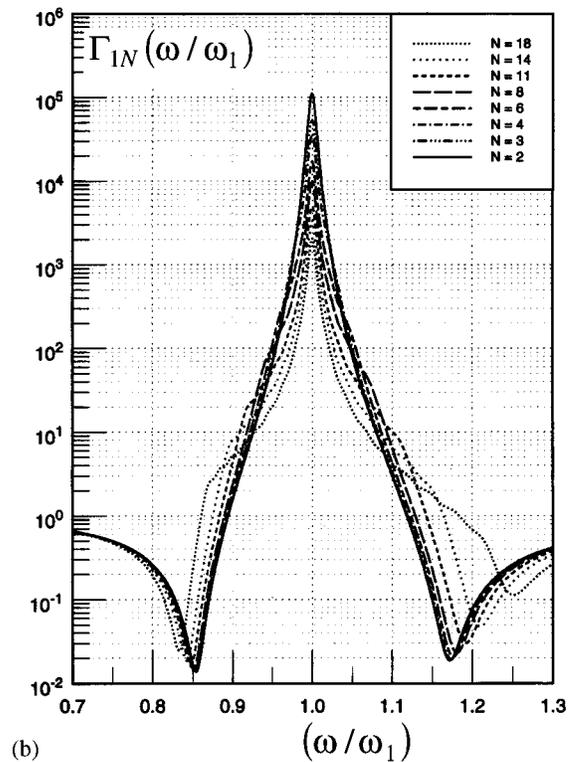
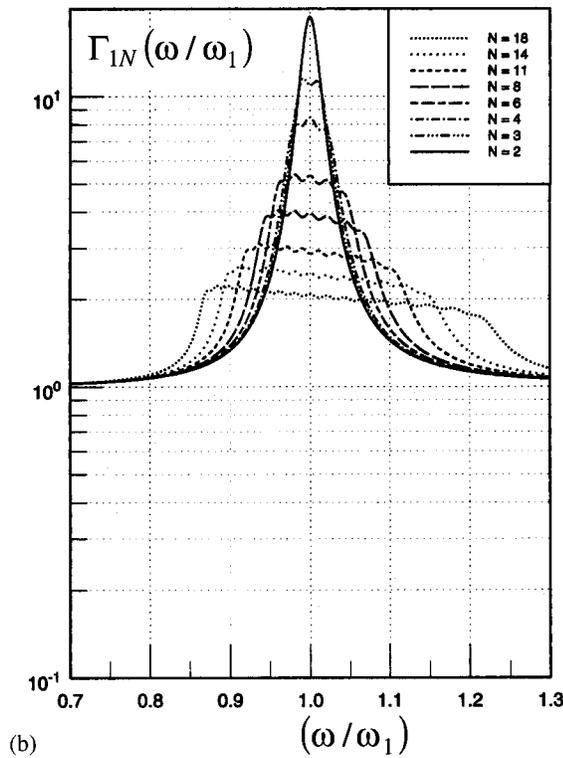
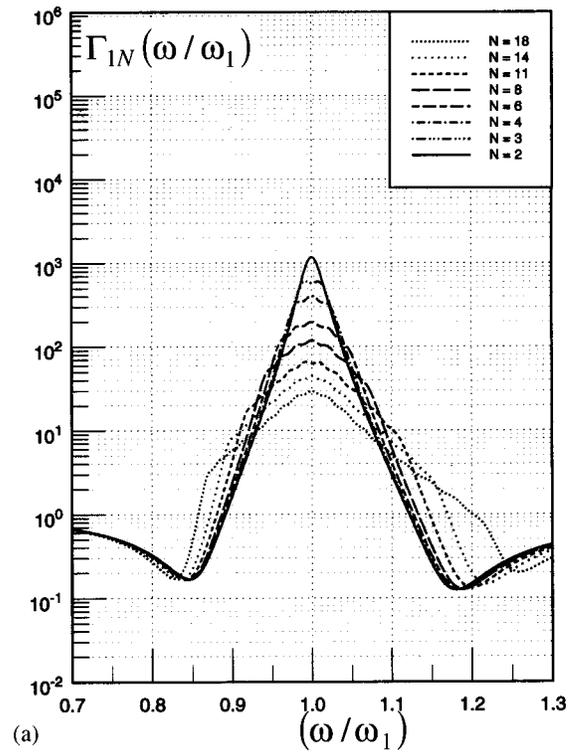
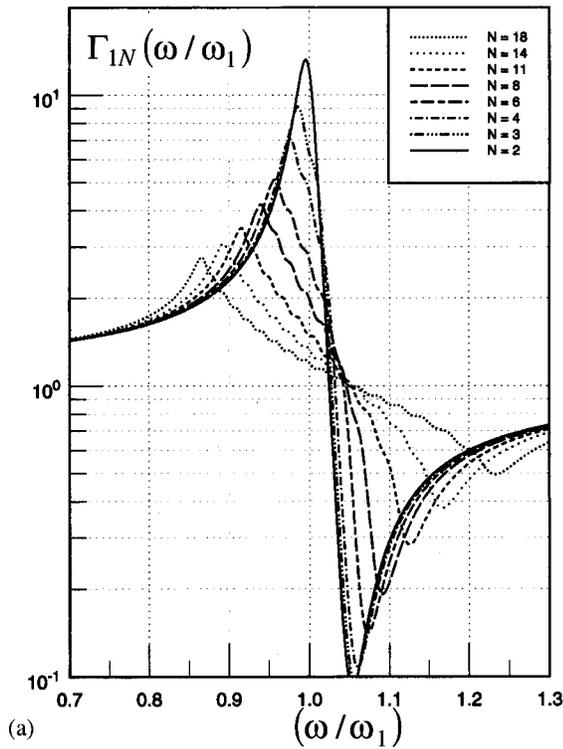


FIG. 3. The overall gain $\Gamma_{1N}(\omega/\omega_1)$ as a function of ω/ω_1 [cf. Eq. (3)]. (a) Computed $\Gamma_{1N}(\omega/\omega_1)$ from Fig. I.14(b) and $\bar{Z}_1(\omega/\omega_1) = i(\omega/\omega_1)$ [cf. Eq. (4a)]. (b) Computed $\Gamma_{1N}(\omega/\omega_1)$ from Fig. I.14(b) and $\bar{Z}_1(\omega/\omega_1) = 1$ [cf. Eq. (4b)].

FIG. 4. The overall gain $\Gamma_{1N}(\omega/\omega_1)$ as a function of ω/ω_1 [cf. Eq. (3)] and note change of ordinate scales from those in Figs. 2 and 3]. (a) Computed $\Gamma_{1N}(\omega/\omega_1)$ from Fig. I.14(b) and $\bar{Z}_1(\omega/\omega_1)$ as specified in Eq. (4c) with $\eta_1 = 10^{-1}$. (b) Computed $\Gamma_{1N}(\omega/\omega_1)$ from Fig. I.14(b) and $\bar{Z}_1(\omega/\omega_1)$ as specified in Eq. (4c) with $\eta_1 = 10^{-2}$.

2(a) and 3(a) for ($N_c \geq N \geq 3$) cannot be considered a major improvement over that for a set of a single-sprung mass for which $N=2$. Thus, although ostensibly the values of the levels of the normalized joint impedance, as displayed in Fig. I.14(a) and I.14(b), hold a promise for an effective overall gain, that promise is largely evaporated in Figs. 2(a) and

3(a). The normalized test impedance that is assigned to Figs. 2(a) and 3(a) is that stated in Eq. (4a); namely, $\bar{Z}_1(\omega/\omega_1) = i(\omega/\omega_1)$. How crucial is that assignment to this question is, in part, given in presenting Figs. 2(b) and

3(b) in which the normalized test impedance $\bar{Z}_1(\omega/\omega_1)$ is assigned by Eq. (4b). Figures 2(b) and 3(b) show that, as N increases, the frequency bandwidths become more pronounced and wider, respectively. It is noted, however, that in both cases the bandwidths remain centered about $\omega/\omega_1 = 1$. These figures also show that the overall gain levels bottom on unity and, therefore, the overall gain is never ineffective. The saturation of the frequency bandwidth and levels in Fig. 2(b) again defines a critical number N_c ; a larger number of sprung masses than $(N_c - 1)$ contributes little to the overall gain in both frequency bandwidth and level. It appears that the saturated frequency bandwidth and the saturated level in that frequency band relate inversely to each other. This statement is vividly illustrated in Fig. 2(c). In this figure the value of the numerical increment γ is set equal to 0.2, while in Fig. 2(b) it is equal to the standard value of 0.4.¹ In Fig. 2(c), as compared with Fig. 2(b), a reduction in the saturated frequency bandwidth is traded for an increase in the saturated level. On the other hand, in Fig. 3(b), one finds that the steady broadening of the frequency bandwidths is accompanied by a reduction in the levels of the overall gain in that frequency band. Again, this inverse relationship, between a frequency bandwidth and a corresponding level, defines the critical number N_c . When N increases beyond N_c , the resulting reduction in level (effectiveness) renders the level unacceptable. Moreover, and as an aside, since γ_0 in Fig. 3(b) is retained at the standard value of 0.02, and the standard value of 0.4 for γ is halved to 0.2 in Fig. 2(c), Eq. (I.17c) dictates that the corresponding curves, in these two figures, for $N=6=N_0$ are identical.¹ It emerges that provided the demand for a higher effectiveness can be relaxed in favor of a wider bandwidth, and in contrast to Figs. 2(a) and 3(a), the overall gains presented in Figs. 2(b) and 3(b) for $(N_c \geq N \geq 3)$ indicate an improvement over the overall gain pertaining to a set of a single-sprung mass for which $N=2$. In all these figures, the latter is depicted by the solid line curves. Thus, the promise that may be held by the absolute values of the normalized joint impedance, as presented in Figs. I.14(a) and I.14(b), materializes, but only in part, in Figs. 2(b) and 3(b). In these figures, the normalized test impedance $\bar{Z}_1(\omega/\omega_1)$ on the master structure is as stated in Eq. (4b), namely, $\bar{Z}_1(\omega/\omega_1) = 1$.

Again, and more insistently, one may ask: How crucial is the role of $\bar{Z}_1(\omega/\omega_1)$ in the determination of the noise control benefits that can be accrued from a device incorporating a multiplicity of sprung masses? In part, the answer to this question has already been advanced. The noise control benefits accrued in Figs. 2(b) and 3(b) are superior to those in Figs. 2(a) and 3(a). To enrich the answer, yet two more figures, Fig. 4(a) and (b), are displayed. These figures use data presented in Fig. I.14(b) together with the normalized test impedance stated in Eq. (4c) to compute the overall gain $\Gamma_{1N}(\omega/\omega_1)$ as a function of the normalized frequency (ω/ω_1) . In Fig. 4(a) the loss factor η_1 is set equal to 10^{-1} and in Fig. 4(b) to 10^{-2} . Figure 4(a) and (b) exhibits more extreme regions of effectiveness and ineffectiveness than the corresponding regions in Fig. 3(b). In fact, as already shown in Fig. 3(b) there are no regions of ineffectiveness. Moreover, the regions in Fig. 4(b) are more extreme than in Fig.

4(a). Therefore, the loss factor η_1 of the master sprung mass, as well as the fact that the master structure is a sprung mass, significantly influences the levels of these extremes.⁴ The attempt to *effectively* broaden the frequency bandwidth by a set that incorporates a multiplicity of sprung masses becomes, in Fig. 4, even more inconclusive than in the preceding Figs. 2 and 3. Nonetheless, in spite of these inconclusive, if not confusing, observations one can, with caution, develop a few criteria to assist with the design of a viable noise control device prior to the specification of its insitu attachment. In this design process, the characteristics of the noise control device only are specified. The influence of the test impedance of the master structure on these criteria is merely neglected, notwithstanding that in the final analysis the *in situ* performance of the device is all that really counts and, for this purpose, a knowledge of this test impedance is an essential ingredient in the final design.²

II. CRITERIA OF PROMISE FOR A NOISE CONTROL DEVICE INCORPORATING A SET OF SPRUNG MASSES

The noise control viability of a device composed of a single-sprung mass may be assessed in terms of its normalized impedance $\bar{Z}_2(1)$ at resonance and its normalized bandwidth $(\Delta\omega/\omega)_2$ about this resonance. Using Eqs. (I.3a) and (I.25) both of these quantities are readily stated in the forms

$$\bar{Z}_2(1) = (a/b\eta), \quad (\Delta\omega/\omega)_2 = (b\eta), \quad (5)$$

where η , b , and a are defined in Eqs. (I.9), (I.10), and (I.18), respectively.¹ The *viability* for this single-sprung mass device may be stated in the form

$$\bar{Z}_2(1) \gg 1, \quad (b\eta) \ll (a). \quad (6)$$

The device incorporating a single-sprung mass, as discussed earlier, suffers from a lack of adequate frequency bandwidth. This may be accounted for by defining a *parameter of promise* $C_2(b)$ defined by the product of the normalized impedance at resonance and the corresponding frequency bandwidth. From Eqs. (5) and (6) one obtains

$$C_2(b) = \bar{Z}_2(1)(\Delta\omega/\omega)_2 = (a). \quad (7)$$

Thus, the *parameter of promise* $C_2(b)$ is proportional to the mass ratio a . To increase the promise of this device, the ratio a of the mass of the sprung mass to that of the effective mass of the master structure needs to be increased, not a happy design compromise. Also, as discussed earlier, the narrow bandwidth of a device composed of a single-sprung mass motivates one to consider a device incorporating a multiplicity of sprung masses. If such a device is, indeed, designed with the intention of endowing it with a wider frequency bandwidth, that intention can be qualified in the form

$$(\Delta\omega/\omega_1)_{N \geq 3} = [(\omega_2 - \omega_N)/\omega_1] \approx [(N-2)\gamma_N] > (b\eta), \quad N \geq 3, \quad (8)$$

where use is made of Eqs. (I.6) and (I.14). The *viability* for this device is cast in terms of the average normalized level of the joint impedance $\langle |\bar{Z}_N(\omega/\omega_1)| \rangle$, where the angular brackets indicate mean-value averaging over the frequency band

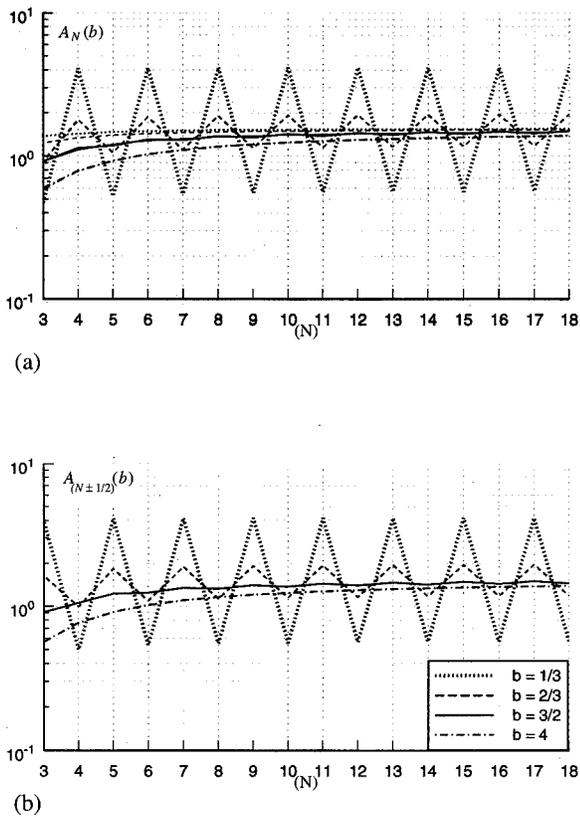


FIG. 5. (a) The quantity $A_N(b)$ as a function of N [cf. Eq. (11a)]. The values of the overlap factors b are as assigned [cf. (b)]. Overlaid is $\bar{A}_N(b)$ as stated in Eqs. (11) and (12) with b as assigned and the lighter curves. (b) The quantity $A_{(N\pm 1/2)}(b)$ as a function of N [cf. Eq. (11b)]. The values of the overlap factors b are as assigned.

defined in Eq. (8). The viability demands that $\langle |\bar{Z}_N(\omega/\omega_1)| \rangle$ exceeds unity in that frequency band; namely,

$$\langle |\bar{Z}_N(\omega/\omega_1)| \rangle \geq 1, \quad N \geq 3 \quad (9a)$$

[cf. Eq. (6)]. Only under this condition can one necessarily ensure the potential viability of the overall gain for this device. To evaluate the average quantity $\langle |\bar{Z}_N(\omega/\omega_1)| \rangle$, a device is in order.

Following the suggestion made in Eq. (5), one obtains from Eqs. (I.9), (I.16) and (I.25) that

$$\bar{Z}_N(1) = i[a/(N-1)\gamma_N] \sum_{j=2}^N \{ (ib)(\omega_j/\omega_1)^2 + [(N+2)-2j] \}^{-1}, \quad N \geq 3, \quad (10a)$$

which can be reasonably approximated in the form

$$\bar{Z}_N(1) \approx [a/(N-1)\gamma_N] A_N(b), \quad N \geq 3, \quad (10b)$$

where

$$A_N(b) = \sum_{j=2}^N i \{ (ib) + [N+2-2j] \}^{-1}, \quad N \geq 3. \quad (11a)$$

The quantity $A_N(b)$ is real and, provided the overlap factor b is about $\frac{3}{2}$, it is of the order of unity. This statement, with regards to $A_N(b)$, is illustrated in Fig. 5(a). In this figure, $A_N(b)$ is displayed as a function of N for $b = \frac{1}{3}, \frac{2}{3}, \frac{3}{2}$, and 4.

The undulations in $A_N(b)$ when b is less than $\frac{3}{2}$ is clearly discernible in Fig. 5(a). The undulations are between the maxima at a resonance for the even N 's and the minima at an antiresonance for the odd N 's when the overlap condition is not met. When this condition is met, i.e., when b equals or exceeds $\frac{3}{2}$, the undulations, as usual, are largely suppressed. The erosion that sets into $A_N(b)$ when the conditions of modal overlap is overly satisfied, $b > \frac{3}{2}$ is also clearly discernible in Fig. 5(a).¹ Thus, the *measure* $A_N(b)$ as a function of N behaves with respect to the overlap factor b as does the *measure* $\bar{Z}_N(\omega/\omega_1)$ as a function of ω/ω_1 (cf. Figs. I.7–I.13). {A *measure* is a dynamic quantity or a factor thereof, e.g., $\bar{Z}_N(1)$ is a *measure* and so is $A_N(b)$ [cf. Eq. (10)].} To impose a mean-value averaging on the *viability*, stated in Eq. (9a), one introduces an undulation-free construct of Eq. (10b). This construct is of the form

$$\langle |\bar{Z}_N(\omega/\omega_1)| \rangle \approx [a/(N-1)\gamma_N] \bar{A}_N(b) \geq 1, \quad N \geq 3, \quad (9b)$$

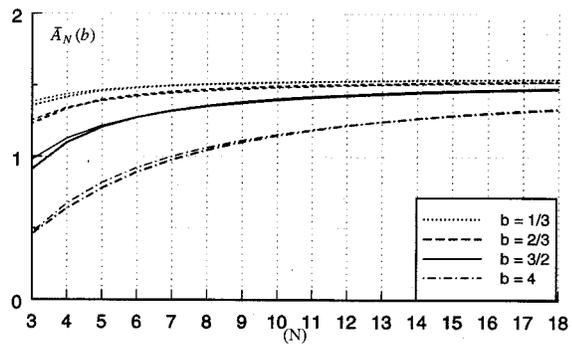
where

$$\bar{A}_N(b) = [A_N(b)A_{(N\pm 1/2)}(b)]^{1/2}, \quad N > 3, \quad (12)$$

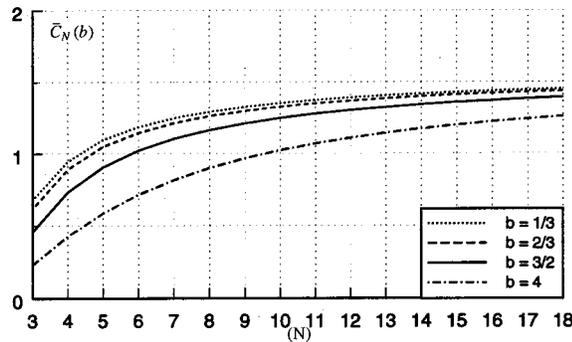
with $A_N(b)$ as stated in Eq. (11a) and with its counter factor $A_{(N\pm 1/2)}(b)$ stated in the form

$$A_{(N\pm 1/2)}(b) = (i/2) \sum_{j=2}^N \{ (ib) + (N+1-2j) \}^{-1} + \{ (ib) + (N+3-2j) \}^{-1}, \quad N \geq 3. \quad (11b)$$

Similar to the *measure* $A_N(b)$, the *measure* $A_{(N\pm 1/2)}(b)$ is real and, provided the overlap factor b is about $\frac{3}{2}$, it is of the order of unity. The *measure* $A_{(N\pm 1/2)}(b)$ is computed and displayed, as a function of N for $b = \frac{1}{3}, \frac{2}{3}, \frac{3}{2}$, and 4, in Fig. 5(b). Again similar to $A_N(b)$ in Fig. 5(a), the undulations in $A_{(N\pm 1/2)}(b)$, when b is less than $\frac{3}{2}$, are clearly discernible in Fig. 5(b). The undulations in $A_{(N\pm 1/2)}(b)$ are half of the two minima at the antiresonances on either side of the resonance at $\omega/\omega_1 = 1$ for the even N 's, and are half of the two maxima at the resonances on either side of the antiresonance at $\omega/\omega_1 = 1$ for the odd N 's. Therefore, the undulations in $A_N(b)$ and in $A_{(N\pm 1/2)}(b)$ are similar except for a displacement of unity with respect to N [cf. Fig. 5(a) and (b)]. The quantity $\bar{A}_N(b)$, stated in Eq. (12), is thus real and is free of undulations in the sense that it is the geometric average of the product of $A_N(b)$ and $A_{(N\pm 1/2)}(b)$.⁹ The undulation-free *measure* $\bar{A}_N(b)$ is computed as a function of N for $b = \frac{1}{3}, \frac{2}{3}, \frac{3}{2}$, and 4; the results of these computations are displayed in Fig. 6(a). It is noted that the undulations that beset Fig. 5(a) and (b) are suppressed, as intended, in Fig. 6(a). Moreover, the erosion that sets in when the conditions of modal overlap are overly satisfied; i.e., when b exceeds $\frac{3}{2}$, is clearly exhibited in Fig. 6(a) (cf. Fig. I.9). Figure 6(a) shows that $\bar{A}_N(b)$ saturates as N increases. Is $\bar{A}_N(b)$ a proper construct of $A_N(b)$? A proper construct is one for which the steady monotonic values largely coincide with the mean values of the undulations in the original *measure*, and the undulation-free values in the original *measure* largely coincide with those in the construct.⁹ Thus, the constructed *measure* $\bar{A}_N(b)$, stated in Eq. (12) and displayed in Fig. 6(a), is a



(a)



(b)

FIG. 6. (a) The undulation-free quantity $\bar{A}_N(b)$ as a function of N [cf. Eqs. (11) and (12)]. The values of the overlap factors b are as assigned. Overlaid is $\bar{A}_N(b)$ as stated in Eq. (14) with b as assigned and the lighter curves. (b) *Criterion of promise* $\bar{C}_N(b)$ as a function of N [cf. Eqs. (12) and (17)]. The values of the overlap factor b are as assigned.

proper construct of the original *measure* $A_N(b)$, stated in Eq. (11a) and displayed in Fig. 5(a). Briefly stated, $\bar{A}_N(b)$ is a proper mean-values counterpart of $A_N(b)$. This propriety is shown in Fig. 5(a) on which Fig. 6(a) is superposed. This superposition is displayed in terms of curves that are appropriately assigned as indicated on the figure. A question arises: Are all mean values the same, i.e., is $\bar{A}_N(b)$ independent of (b)? The answer to this question is of particular interest to this paper. In part, the answer to this question has been cavalierly given already. In the range of moderate parametric values, the preceding figures indicate that, although for large overlap factors ($b > \frac{3}{2}$) the values of $A_N(b)$ and $\bar{A}_N(b)$ are the same, both these values are eroded in the sense that they lie below those pertaining to the neutral overlap factor ($b \approx \frac{3}{2}$). The erosion is particularly severe for the smaller values of N and the larger values of b . On the other hand, for small but different overlap factors ($b \leq \frac{3}{2}$), the mean values of the undulations in $A_N(b)$, as well as the corresponding values of the undulation-free construct $\bar{A}_N(b)$, are closer to each other and this closeness is the cozier, the larger the values of N . The quantification of these observations can now be conducted. Imposing that N is large compared with b and unity, one may asymptotically evaluate $\bar{A}_N(b)$.¹⁰ From Eqs. (11) and (12) one obtains

$$\bar{A}_N(b) \rightarrow (\pi/2), \quad N \gg b, \quad N \gg 3. \quad (13)$$

The validity of Eq. (13) is dependent on N and b . However, once the condition of validation is satisfied, the evaluated $\bar{A}_N(b)$ is independent of N and b .^{6,7,11} Thus, the convergence of $\bar{A}_N(b)$ with increase in N , which is central in Fig. 6(a), is explained. Also explained is the role played by the overlap factor b . When b equals or exceeds $\frac{3}{2}$, $A_N(b) \approx \bar{A}_N(b)$, and when b is less than $\frac{3}{2}$ the mean values of $A_N(b)$ equal the values of $\bar{A}_N(b)$. One is reminded, however, that when b is the smaller, the undulated strokes of $A_N(b)$ are the larger. [$\bar{A}_N(b)$ is a *proper* mean-values counterpart of $A_N(b)$.] The propriety remains intact even if b approaches zero. Although the strokes of the undulations in $A_N(b)$ become unwieldy as $b \rightarrow 0$, their mean values, nonetheless, coincide with the undulation-free values of $\bar{A}_N(b)$. In this case $\bar{A}_N(b) \approx \pi/2$ even if N is not much greater than 3. The saturation of $\bar{A}_N(b)$ and its independence of b for large values of N , i.e., $N \gg 3$ and $N \gg b$, are recognized attributes, but they are not always stated and presented in the manner adopted herein^{6,11-13} [cf. Eqs. (9)–(12) and Fig. 6(a)]. One may attempt to find a fit for $\bar{A}_N(b)$ based on Eq. (13); this attempt yields the approximate expression

$$\bar{A}_N(b) \approx (\pi/2) \exp[-g(N)f(b)], \quad N \gg 3, \quad (14)$$

where

$$g(N) = (2N-1)[2(N-1)N]^{-1} \rightarrow (1/N), \quad N \gg 3, \quad (15a)$$

$$f(b) = (\frac{1}{2})^{1/2}[(b) + (10)^{-1}]. \quad (15b)$$

The *approximate measure* $\bar{A}_N(b)$, as expressed in Eqs. (14) and (15), is overlaid on Fig. 6(a), in terms of the lighter curves. This figure depicts the *exact measure* $\bar{A}_N(b)$ as expressed in Eqs. (11) and (12), in terms of the darker curves. The match between the two *measures* of $\bar{A}_N(b)$ is deemed close enough and, therefore, Eqs. (14) and (15) may be used to decipher the dependence of $\bar{A}_N(b)$ on both N and b .

Following the procedure that leads from Eqs. (5) and (6) to Eq. (7) for a single-sprung mass, one may define a *parameter of promise* $C_N(b)$ for a set of multiplicity of sprung masses. The definition and evaluation of the *parameter of promise* is cast in terms of the product of the average normalized joint impedance $\langle |\bar{Z}_N(\omega/\omega_1)| \rangle$ and the bandwidth $(\Delta\omega/\omega)_{N \gg 3}$.² In this paper these quantities are evaluated in Eqs. (8)–(10). Based on these approximate evaluations, the product yields the *parameter of promise* in the form

$$C_N(b) = \langle |\bar{Z}_N(\omega/\omega_1)| \rangle (\Delta\omega/\omega_1)_{N \gg 3} \approx [a/(N-1)](N-2)\bar{A}_N(b), \quad N \gg 3, \quad (16)$$

which is independent of γ_N and, therefore, of the distribution of the resonance frequencies. A *criterion of promise* $\bar{C}_N(b)$ may then be defined as the ratio of the *parameter of promise* $C_N(b)$ of a noise control device incorporating $(N-1)$ -sprung masses to the *parameter of promise* $C_2(b)$ of a noise control device incorporating but a single-sprung

mass.² From Eqs. (7) and (16) the *criterion of promise* is derived

$$\bar{C}_N(b) = [C_N(b)/C_2(b)] \approx (N-1)^{-1}(N-2)\bar{A}_N(b), \quad N \geq 3. \quad (17)$$

The larger $\bar{C}_N(b)$, the greater the promise that the device may potentially hold. [A similar quantity, defined as the “gain bandwidth product,” was presented in Ref. 2. In this reference only the second distribution of resonance frequencies was employed. Since the *parameter of promise* $C_N(b)$ is independent of γ_N , the *criterion of promise* is also independent of γ_N . Therefore, $\bar{C}_N(b)$ in Ref. 2 qualifies as a universal *measure*.] The *criterion of promise* $\bar{C}_N(b)$ is computed as a function of N for $b = \frac{1}{3}, \frac{2}{3}, \frac{3}{2}$, and 4; the results of these computations are displayed in Fig. 6(b). It is noted that the undulations that beset Fig. 5 are suppressed, as intended, in Fig. 6. Indeed, Eq. (17) states that $\bar{C}_N(b)$ is proportional to $\bar{A}_N(b)$. It is also proportional to the factor $(N-1)^{-1}(N-2)$. This factor is deficient at the low values of (N) , e.g., $[(N-1)^{-1}(N-2)]_{N=3} = \frac{1}{2}$, rapidly approaching unity as N is increased. Therefore, $\bar{C}_N(b)$ is slightly deficient with respect to $\bar{A}_N(b)$ in the lower ranges of N , and is rapidly approaching $\bar{A}_N(b)$ as N is increased. As $\bar{A}_N(b)$ in Fig. 6(a), $\bar{C}_N(b)$ in Fig. 6(b) saturates with the increase of N . Indeed, the behavior of $\bar{C}_N(b)$ closely resembles that of $\bar{A}_N(b)$. In particular, the saturation of $\bar{C}_N(b)$ defines a critical number N_c for N . Most of the promised benefits are accrued for a device incorporating $(N_c - 1)$ -sprung masses. A demand for a larger number of sprung masses may become an excessive design sophistication.

How do the *viability*, as stated in Eq. (9), and the *criterion of promise*, as stated in Eq. (17), pertain to the design of devices that incorporate sets of sprung masses that are characterized and typified in Figs. I.7–I.14 and Figs. 2–6. According to Eqs. (I.17) and Eqs. (9) and (17), these criteria require that

$$\gamma \left. \vphantom{\begin{matrix} \gamma \\ \gamma_0 \end{matrix}} \right\} \leq \bar{A}_N(b) \left\{ \begin{matrix} (2a), \\ [a/(N-1)], \end{matrix} \right. \quad (18a)$$

$$\gamma_0 \left. \vphantom{\begin{matrix} \gamma \\ \gamma_0 \end{matrix}} \right\} \leq \bar{A}_N(b) \left\{ \begin{matrix} (2a), \\ [a/(N-1)], \end{matrix} \right. \quad (18b)$$

$$\bar{C}_N(b) \rightarrow \bar{A}_N(b); \quad N \geq N_c = 6, \quad (19)$$

respectively. Since, as depicted in Fig. 6(a), $\bar{A}_N(b)$ is substantially equal to unity for $b \leq \frac{3}{2}$, the choice in Fig. I.14 of $a = 10^{-1}$, $\eta = 0.02$, and $\gamma = 0.4$ or $\gamma_0 = 0.02$ satisfies only marginally the inequality stated in Eq. (18). On the other hand, as Fig. 6(a) and (b) shows, the *criterion of promise* $\bar{C}_N(b)$, for a $b = \frac{3}{2}$ and under the condition stated in Eq. (19), exceeds unity by a factor of better than $\sqrt{2}$. Again, although this is not dramatic, it may prove significant in certain situations. [It should be remarked that a rigorous evaluation of the normalized bandwidth $(\Delta\omega/\omega)_2$ equates it more closely to $(\pi/2)(b\eta)$ than to $(b\eta)$. The latter value is that assigned to it in Eq. (5). Under this upgraded evaluation the *criterion of promise* $\bar{C}_N(b)$ approaches at best the value of unity and not the value of $\sqrt{2}$ just quoted. Again, hair splitting with respect to loss factors up to better than a factor of 2 is largely of academic interest.^{1,14]} Thus, the failure of the data in Fig.

I.14, via Eqs. (3) and (4), to achieve dramatically effective wideband-overall gains in Figs. 2(a), (b), 3(a), (b), 4(a), and (b) is explained. Although shorter on bandwidth, the more effective overall gain in Fig. 2(c) is also explained.

III. A DIFFUSED SPECULATION AND A GENERAL CONCLUSION

One may now address the speculation that the mean values of the undulations and the values of the undulation-free *measures* of a set of sprung masses are largely equal and are, therefore, independent of the loss factors. {The loss factors of the individual sprung masses are here determined by the distribution of the resonance frequencies and the overlap factor b . Thus, for a given distribution of the resonance frequencies, the loss factors are solely determined by b [cf. Eqs. (I.9) and (I.10)].} This declaration of independence is found to be only obliquely valid. The validity of this statement holds without qualification only if the number $(N-1)$ of sprung masses in the set approaches infinity so that

$$g(N)f(b) \leq 10^{-1}, \quad N \geq 3, \quad (20a)$$

where $g(N)$ and $f(b)$ are stated in Eq. (15). The statement is only partially valid, to within 10%, if

$$g(N)f(b) \leq 10^{-1}, \quad N \geq 3. \quad (20b)$$

When the condition stated in Eq. (20b) is violated, erosion, both in the mean values and in the values of the measures, progressively sets in. The erosion sets in when

$$g(N)f(b) \geq 10^{-1}, \quad N \geq 3, \quad (20c)$$

and the greater the inequality stated in Eq. (20c), the more severe the erosion. The erosion weakens both the *viability* and the *criterion of promise* and this is the reason that designers deem the overly damped sprung masses, e.g., when $b = 8$, to be detrimental to the performance of the device incorporating them [cf. Eqs. (9) and (17)–(19)]. On the other hand, the detrimental influence of the undulations on the performance of the device is obvious. The strokes of the undulations are larger the smaller b is, e.g., the undulations are severe when $b \leq \frac{1}{3}$. Therefore, an overlap factor b of $\frac{3}{2}$ is a design goal. To ensure against undulations on the one hand, and erosion on the other, Eqs. (15) and (20) suggest that for a 10% of maximum effectiveness,

$$g(N)f(b) \geq 10^{-1}, \quad [g(N_c)]^{-1} \approx 11 \approx N_c, \quad (21a)$$

and, for a 20% of maximum effectiveness,

$$g(N)f(3/2) \geq 2 \times 10^{-1}, \quad [g(N_c)]^{-1} \approx 5 \approx N_c, \quad (21b)$$

where the data presented in Fig. 6 is utilized and N_c is the critical number of (N) as defined and discussed herein.

Finally, it is emphasized and reiterated that the *in situ* viability of the devices can be ascertained only in terms of the overall gain $\Gamma_{1N}(\omega/\omega_1)$, stated in Eq. (3). The *viability*, as stated in Eq. (9), and the *criterion of promise*, as stated in Eq. (17), are merely necessary conditions. They are, by no means, sufficient conditions to ensure the effectiveness of the overall gain of a noise control device incorporating a multiplicity of sprung masses. The viability of the overall gain is crucially dependent on the normalized point impedance that

the noise control device perceives in the host structure—the master structure—to which it is attached. Moreover, to ascertain the viability of the noise control device, the characteristics of the device need to be described in a manner that facilitates the assessment of the true achievement that can be mustered by the device, e.g., the complex values of the joint impedance perceived by the master structure as compared with the absolute values of this quantity² (cf. Figs. 2–4).

¹G. Maidanik and K. J. Becker, “Characterization of multiple-sprung masses for wideband noise control,” *J. Acoust. Soc. Am.* **106**, 3109–3118 (1999).

²M. J. Brennan, “Wideband vibration neutralizer,” *Noise Control Eng. J.* **45**, 201–207 (1997).

³K. U. Ingard, *Fundamentals of Waves and Oscillations* (Cambridge U.P., Cambridge, 1988).

⁴G. Maidanik and K. J. Becker, “Noise control of a master harmonic oscillator that is coupled to a set of satellite harmonic oscillators,” *J. Acoust. Soc. Am.* **104**, 2628–2637 (1998).

⁵C. Soize, “Probabilistic structural modeling in linear dynamic analysis of complex mechanical systems,” *Rech. Aerosp.* **1986-3**, 23–48 (1986); “A model and numerical method in the medium frequency range for vibroacoustic predictions using the theory of structural fuzzy,” *J. Acoust. Soc. Am.* **94**, 849–865 (1993); R. Ohayon and C. Soize, *Structural Acoustics and Vibrations* (Academic, San Diego, 1998).

⁶A. D. Pierce, V. W. Sparrow, and D. A. Russell, “Fundamental structural acoustics idealizations for structures with fuzzy internals,” *Trans. ASME, J. Vib. Acoust.* **117**, 1–10 (1995); A. D. Pierce, “Fuzzies but simpler analytic models for physical acoustics and structural acoustics,” *J. Acoust. Soc. Am.* **103**, 2954 (1988).

⁷M. Strasberg and D. Feit, “Vibration damping of large structures induced by attached small resonant substructures,” *J. Acoust. Soc. Am.* **99**, 335–344 (1996).

⁸M. Strasberg, “Continuous structures as ‘fuzzy’ substructures,” *J. Acoust. Soc. Am.* **100**, 3456–3459 (1996).

⁹E. Skudrzyk, “The mean-value method of predicting the dynamic response of complex vibrators,” *J. Acoust. Soc. Am.* **67**, 1105–1135 (1980).

¹⁰I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965) (translation edited by A. Jeffrey, p. 36).

¹¹R. J. Nagem, I. Veljkovic, and G. Sandri, “Vibration damping by a continuous distribution of undamped oscillators,” *J. Sound Vib.* **207**, 429–434 (1997).

¹²M. Strasberg, “Is the dissipation induced by ‘fuzzy’ substructures real or apparent,” *J. Acoust. Soc. Am.* **102**, 3130 (1997).

¹³G. Maidanik and K. J. Becker, “Various loss factors of a master harmonic oscillator that is coupled to a number of satellite harmonic oscillators,” *J. Acoust. Soc. Am.* **103**, 3184–3195 (1998).

¹⁴R. H. Lyon, *Statistical Energy Analysis of Dynamic Systems: Theory and Applications* (MIT, Cambridge, 1975); R. H. Lyon and R. G. DeJong, *Theory and Application of Statistical Energy Analysis* (Butterworth-Heinemann, Boston, 1995).

A class of expansion functions for finite elastic plates in structural acoustics

Richard V. Craster

Department of Mathematics, Imperial College of Science, Technology and Medicine, London SW7 2BZ, United Kingdom

Stefan G. Llewellyn Smith

Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0411

(Received 4 February 1999; accepted for publication 26 August 1999)

Problems in structural acoustics involving finite plates can be formulated using integral equation methods. The unknown function within the integral equation must satisfy the plate edge conditions, and hence appropriate expansion functions must be used. The expansion functions developed here are aimed at treating a wide class of problems. Once such functions are found, the solution process and numerical implementation are relatively straightforward. The speed of convergence to “exact” comparison solutions is fast even in the singular limit of high frequencies and wide plates. A set of expansion functions with the required properties is constructed and some illustrative problems are treated. © 1999 Acoustical Society of America. [S0001-4966(99)03512-2]

PACS numbers: 43.20.Tb, 43.40.Dx, 43.40.Rj [CBB]

INTRODUCTION

Many papers have been written on integral equation methods for treating finite length rigid plates in acoustics, or analogous problems in elastodynamics or electromagnetism.^{1,2} In these cases the integral equations are readily and efficiently solved using expansion functions, Chebyshev polynomials, for the unknown. The vital ingredient captured by the expansion functions is that they correctly incorporate the edge condition at the plate ends. Convergence is usually rapid even for higher frequency problems. Unfortunately, the usual Chebyshev polynomial expansion functions are not well suited to treating, say, elastic plate boundary conditions; the edge behavior is altered in these cases. Our aim here is to develop a class of expansion functions, analogous to the rigid plate ones, that are equally applicable to elastic plate problems.

There is recurrent interest in sound generation or scattering by elastic plates in structural acoustics. Many structures consist of plates welded together or attached in other ways, and vibrational plate waves are potential major sources of acoustic noise. These waves may be coupled into the surrounding fluid via interactions with the plate edges, for example. Thus analytical and numerical techniques have often been used to try to describe the general physical effects involved, and to solve model problems for specific geometries.³ Problems involving finite length plates with various attachment conditions are unfortunately not amenable to exact solution, although asymptotic results for wide plates and light or heavy fluid loading can be found. Our aim is to develop efficient numerical approaches capable of solving these problems in regimes not necessarily amenable to asymptotic analysis, and which can also be generalized to more complex geometries.

In this paper we shall concentrate on a pedagogic two-dimensional example. However, the basis expansion func-

tions we develop can be utilized in more complex geometries. From a numerical perspective one has to deal with a relatively high order boundary condition on the elastic plate, such as the classical plate equation,⁴

$$B \frac{\partial^4 \eta}{\partial x^4} - m \omega^2 \eta = -p(x, 0), \quad (1)$$

for the plate displacement $\eta(x)$, where $p(x, 0)$ is the pressure on the plate. This equation is adopted here and involves the fourth derivative of η . The plate is assumed to separate fluid from a vacuum. At any joint or plate edge we must apply two edge conditions to the displacement. For instance, clamped plates have $\eta = \eta' = 0$ at the edges. The appropriate conditions must be built into any numerical scheme either implicitly or explicitly. Moreover, the edge conditions are important, as analytical studies show a marked dependence upon them.

A variety of numerical methods such as finite element schemes, modal methods, and integral equation approaches have been utilized by other authors. Finite elements are versatile. However, one has to discretize the whole domain, with the result that infinite domains are awkward, and calculations become increasingly unwieldy as frequencies increase. Modal methods have advantages for the simply supported $\eta = \eta' = 0$ edge conditions, as Fourier series solutions can be developed; however, this approach cannot be used for more general edge conditions. For more general edge conditions, analytic approximations utilizing the *in vacuo* eigenfunctions can be adopted;⁵ we briefly discuss eigenfunction methods in Sec. IV. There are also numerical approaches using the modified Wiener–Hopf technique;⁶ these are perhaps less flexible than the numerical methods based directly upon solving the integral equations. Nonetheless they are formally exact if one continues the iterations indefinitely, although this is at the cost of considerable effort.

Integral equation methods have considerable advantages for relatively simple geometries, and this is the approach we develop here. Previous authors, for instance Mattei,⁷ have adopted this technique, but have used different methods that have difficulties explicitly incorporating the edge conditions. Typically the edge conditions are treated separately, and since there have been few comparisons with exact or asymptotic methods, it is unclear how successful such numerical methods have been.

Here we treat the integral equations utilizing expansion functions that automatically take into account the edge conditions. As a result, the edge conditions are implicitly satisfied and no extra equations are required. This approach quickly leads to an accurate and efficient numerical scheme: typically only a few expansion functions are required to ensure accurate solutions. For plates with clamped edge conditions this approach has been detailed elsewhere,⁸ and has been compared to asymptotic solutions near resonant frequencies and to asymptotic solutions for light and heavy fluid loading. Compliant plate effects are also easily treated and several types of forcings are considered: incident plane waves, line forces and moments, and sources in the fluid. Our aim previously⁸ was to introduce and develop this approach and to show that it can be particularly useful in regimes not amenable to asymptotic analyses and hence convenient for quite general incident fields. The aim here is to expand upon the method and consider a wider class of edge condition.

We consider time harmonic vibrations of frequency ω , and all physical variables are assumed to have an $e^{-i\omega t}$ dependence. This is considered understood and is henceforth suppressed. Two-dimensional problems are considered with an inviscid, compressible fluid lying in $x_3 > 0$ and $-\infty < x_1 < \infty$, and a vacuum lying in $x_3 < 0$. The fluid pressure $p(x_1, x_3)$ satisfies

$$(\nabla^2 + k_0^2)p(x_1, x_3) = f(x_1, x_3), \quad (2)$$

where $f(x_1, x_3)$ corresponds to a distribution of fluid sources, and k_0 , the acoustic wave number, is related to the sound speed c_0 via $k_0 = \omega/c_0$. In what follows, the source distribution is zero except for Green's functions. The displacement in the x_3 -direction within the fluid, $\eta(x_1, x_3)$, is related to the fluid pressure via

$$\rho\omega^2\eta(x_1, x_3) = \frac{\partial p(x_1, x_3)}{\partial x_3}. \quad (3)$$

The plane $x_3 = 0$ is taken to consist of a thin elastic plate in the finite region $|x_1| < a$ in which (1) holds, and to consist of a rigid plate elsewhere. The geometry is shown in Fig. 1.

The parameters B and m are the bending stiffness and mass per unit area of the plate, respectively. These parameters are related to the properties of the elastic plate via $B = Eh^3/12(1 - \nu^2)$ and $m = \rho_s h$, with E , h , ν , and ρ_s the Young's modulus, plate thickness, Poisson ratio, and mass density of the elastic material, respectively. In order to minimize the number of parameters that occur later, we introduce the *in vacuo* flexural wave number $k_p \equiv (\omega^2 m/B)^{1/4}$. Incorporating a small loss factor will lead to attenuation of the plate waves; we shall not consider loss factors here. We in-

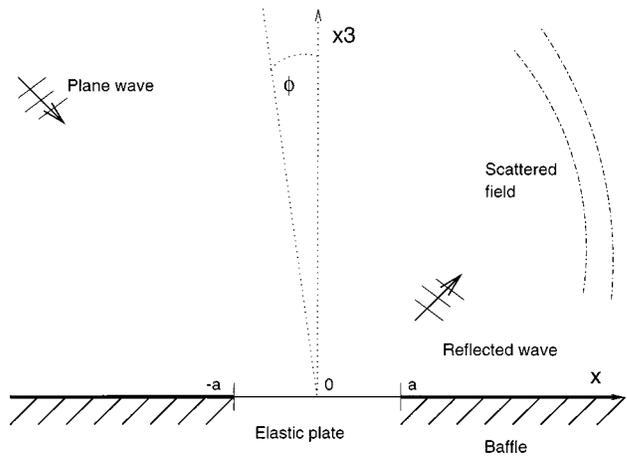


FIG. 1. Geometry of the problem. Note that $x = x_1$, and that the two sets of Green's function variables in Sec. II are (x_1, x_3) and (q_1, q_3) .

roduce the nondimensional quantities⁹ M and ϵ . The "Mach" number M is defined to be the ratio of the fluid sound speed to that of the *in vacuo* plate waves, $M \equiv k_0/k_p$. A frequency independent measure of fluid loading is provided by the parameter $\epsilon \equiv (B\rho^2/m^3 c_0^2)^{1/2}$. In essence, when the system is lossless, there are three parameters that can be varied: M , ϵ , and $k_0 a$, the last of these being the ratio of a typical length scale associated with the fluid disturbance to a typical length scale associated with the finite defect. Typically ϵ is small: for example, $\epsilon \approx 0.134$ for steel plates in water, while M , which is frequency dependent, can range through all values. The fluid loading will be termed⁹ "light" when $M \sim O(1)$, but is not in the immediate neighborhood of $M = 1$, and "heavy" when $M \ll \epsilon$. In both cases ϵ is taken to be small, i.e., $\epsilon \ll 1$.

We previously treated the clamped edge conditions,⁸ but there are other possibilities. Hinged plates satisfy the edge conditions $\eta = \eta' = 0$ at their edges. The natural expansion functions in that case can be taken as the Fourier modes, $\cos \pi(n-1/2)x$, $n = 1, 2, \dots$ and $\sin \pi nx$, $n = 1, 2, \dots$. For the special case $\eta = \eta' = 0$, we have used this approach for comparative purposes (see Sec. III), and to verify that our expansion functions are accurate in this case. We should emphasize that the Fourier modes are probably easier to use for this special case, but they cannot be used for the more general edge conditions (4) that we propose to tackle.

Leppington *et al.*¹⁰ pointed out after looking at experimental data that a more realistic set of edge conditions might be

$$\eta'' \pm \delta \eta' = 0 \quad \text{at } x = \pm 1, \quad (4)$$

together with $\eta = 0$ at $x = \pm 1$, where δ is a positive number and the plate edges are at $x = \pm 1$. The case $\delta = 0$ corresponds to the hinged case, while the limit of large δ recovers the clamped case previously considered. The more general edge condition may be interpreted as a hinged edge with a restoring couple and may be incorporated into our approach by changing the previous expansion functions.⁸

The physical problem is split into pieces that are even and odd in x , and the superscripts $(e), (o)$ are used to denote

these subproblems. The even expansion functions are taken to be

$$\psi_n^{(e)}(x) = \begin{cases} \cos[n \cos^{-1}\{(1-x^2)(1-ax^2)\}] & \text{for } n \text{ odd,} \\ (1-x^2)\cos[(n-1)\cos^{-1}(1-x^2)^2] & \text{for } n \text{ even,} \end{cases} \quad (5)$$

in $0 \leq x \leq 1$. For negative x these are extended as even functions of x . The quantity a is given by

$$a = \frac{1 + \delta}{5 + \delta}. \quad (6)$$

The odd functions become

$$\psi_n^{(o)}(x) = \begin{cases} i \sin[(n+1)\cos^{-1}\{(1-x^2)(1-ax^2)\}] & \text{for } n \text{ odd,} \\ i(1-x^2)\sin[n\cos^{-1}(1-x^2)^2] & \text{for } n \text{ even} \end{cases} \quad (7)$$

in $0 \leq x \leq 1$. These are extended as odd functions of x for x negative. This choice of expansion functions is adopted as the expansion functions satisfy the edge conditions exactly; it is worth noting that these expansion functions are related to Chebyshev polynomials with the appropriate edge behavior in their argument. The expansion functions for n even are required, as the terms with n odd exclude the terms of the form $(1-x^2)^3$ near the edge, which also satisfy the edge conditions. As we shall see, these expansion functions extend those usually used for rigid plates, allowing one to extend previous analyses to elastic plates.

I. COMPLETENESS OF THE EXPANSION FUNCTIONS

One issue surrounding these expansion functions is whether they have any underlying mathematical basis. It is important to verify that the plate displacement can actually be represented by the expansion functions, and that the latter are hence complete. The expansion functions typically used in elasticity for crack problems² may easily be shown to be complete. These functions are (to within a normalization factor),

$$\phi_n(x) = \begin{cases} \cos(n \sin^{-1} x) & \text{for } n \text{ odd,} \\ \sin(n \sin^{-1} x) & \text{for } n \text{ even,} \end{cases} \quad (8)$$

on the interval $(-1, 1)$. Both the even and odd functions of x can be conveniently considered together. The change of variable $x = \sin \theta$ maps the x -interval $(-1, 1)$ onto the θ -interval $(-\pi/2, \pi/2)$. On this interval, the expansion functions are

$$\phi_n = \begin{cases} \cos n \theta & \text{for } n \text{ odd,} \\ \sin n \theta & \text{for } n \text{ even.} \end{cases} \quad (9)$$

This set of functions is complete on the interval $(-\pi/2, \pi/2)$, since it corresponds to the usual Fourier sine expansion on the interval.

The set given by (5)–(7) requires more care.⁸ Any function f on $(-1, 1)$ may be decomposed into its odd and even parts f_o and f_e , respectively, both defined on $(0, 1)$ and with

$f_e'(0) = f_o(0) = 0$. The change of variable $\cos \theta = (1-x^2)(1-ax^2)$ is a one-to-one and onto mapping of the x -interval $(0, 1)$ onto the θ -interval $(0, \pi/2)$. On this new interval, the transformed even expansion functions are again $\cos n\theta$ for n odd. However, the expansion functions are not complete on this interval: a complete set of trigonometric functions for the expansion of even functions on this interval requires even n as well. One cannot just add the even cosines as extra expansion functions though, since this would remove the correct edge condition in the original variable x and hence introduce Gibbs' effects that would undermine the whole aim of the expansion functions. The additional functions $\psi_e' = (1-x^2)\cos[n\cos^{-1}(1-x^2)^2]$, which explicitly introduce terms of the form $(1-x^2)^3$ near $x = \pm 1$, satisfy the appropriate edge conditions, and project onto all the cosine functions, in particular onto the even ones. This was shown for the clamped case previously.⁸ Hence the set of expansion functions for the even part of f is complete. The issue of whether these expansion functions are orthogonal with respect to a particular weight function is irrelevant since orthogonality properties are never used.

The argument for the odd part of f is analogous. Therefore the set of expansion functions (5)–(7) is complete, and in addition satisfies the appropriate edge conditions.

II. FORMULATION AND SOLUTION OF INTEGRAL EQUATIONS

We consider the plane $x_3 = 0$ with the elastic plate lying on $|x_1| < a$ and a rigid baffle on $|x_1| > a$. As can be shown from a Green's function approach, the scattered pressure field at a point (q_1, q_3) is given by

$$p^{sc}(\mathbf{q}) = -\rho\omega^2 \int_{-a}^a \eta^{sc}(x_1, 0) p^G(x_1, 0; \mathbf{q}) dx_1, \quad (10)$$

where $p^G(\mathbf{x}, \mathbf{q})$ is the Green's function for (2) which has vanishing x_3 -derivative in $x_3 = 0$. The plate displacement $\eta^{sc}(x_1, 0)$ is unknown in (10) and our aim is to identify this function in the most efficient manner. Once this is identified the problem is effectively solved, since pressure fields and far-field behavior follow directly from (10). Hence we concentrate upon whether we have identified η correctly, and our comparisons with other techniques are based upon this quantity.

The Green's function is the inverse Fourier transform

$$p^G(\mathbf{x}; \mathbf{q}) = \frac{i}{4\pi} \int_C \frac{1}{\gamma_0} [e^{i\gamma_0|x_3-q_3|} + e^{i\gamma_0(x_3+q_3)}] e^{ik(x_1-q_1)} dk, \quad (11)$$

which also has a representation in terms of Hankel functions. The path C runs along the real axis suitably indented at the branch points $\pm k_0$. The function $\gamma_0 = (k_0^2 - k^2)^{1/2}$ has a positive imaginary part. An advantage of persisting with a transform based representation is that for more complex geometries the Green's functions emerge in a similar manner; the approach that is required follows that presented here.

As it stands, (10) is not in the form where we can solve for the unknown displacement, since the left-hand side is also unknown. To remedy this, we manipulate (10) so that

the left-hand side becomes the plate equation, and we are left with an integral equation to solve. The manipulation is most easily performed by replacing $p^G(x_1, 0; \mathbf{q})$ in (10) with its transform (11) and applying the operator D_q defined by

$$D_q = B \partial_{q_1}^4 - m \omega^2, \quad (12)$$

where the notation $\partial_{q_i} = \partial / \partial q_i$ has been adopted. Thus we obtain

$$\begin{aligned} p^{sc}(q_1, 0) + D_q \eta^{sc}(q_1, 0) \\ = \frac{B}{2\pi} \int_{-a}^a \eta^{sc}(x_1, 0) dx_1 \\ \times \int_{-\infty}^{\infty} \left[k^4 - \frac{m\omega^2}{B} - \frac{i\omega^2 \rho}{\gamma_0 B} \right] e^{ik(x_1 - q_1)} dk \end{aligned} \quad (13)$$

for $|q_1| < a$.

Two types of forcing can be adopted here: either incident waves or local plate excitation. We treat the incident wave case in detail, in which case the left-hand side in the above equation is equal to $-(p^{in} + D_q \eta^{in})$. For plane wave incidence, the incident and reflected pressure wave together are

$$p^{in}(q_1, q_3) = A [e^{i(kq_1 + \gamma_0 q_3)} + e^{i(kq_1 - \gamma_0 q_3)}], \quad (14)$$

where

$$k = k_0 \sin \theta_i \quad (15)$$

is the incoming wave number. This corresponds to the field produced by a pressure wave incident upon a defect-free rigid plate. The incident displacement field therefore vanishes on the elastic plate, leading to

$$-[p^{in} + D_q \eta^{in}](q_1, 0) = -2A e^{ikq_1} \quad (16)$$

for $|q_1| < a$.

Each applied incident field is split into two subproblems, one that is even in x , and one that is odd in x . The unknown displacement along the plate is expanded in terms of the expansion functions (5)–(7). The appropriate expression, in which the factor $4a^4/B$ is inserted for convenience, is

$$\eta^{sc}(x_1, 0) = \frac{4a^4}{B} \sum_{n=1}^{\infty} (a_n^{(e)} \psi_n^{(e)}(x_1/a) + a_n^{(o)} \psi_n^{(o)}(x_1/a)). \quad (17)$$

The integral equation (13) is split into even and odd subproblems. The even subproblem is then solved by multiplying (13) by $\psi_m^{(e)}(q_1/a)$ and integrating from $-a$ to a , as well as expanding the scattered displacement on the plate, leading to

$$\begin{aligned} -\pi A \int_{-1}^1 \cos(kaq_1) \psi_m^{(e)}(q_1) dq_1 \\ = \sum_{n=1}^{\infty} a_n^{(e)} \int_{-\infty}^{\infty} \int_{-1}^1 \int_{-1}^1 \psi_m^{(e)}(q_1) \psi_n^{(e)}(x_1) e^{il(x_1 - q_1)} \\ \times \left[l^4 - (ak_p)^4 - \frac{i(ak_p)^6 \epsilon}{ak_0 [(ak_0)^2 - l^2]^{1/2}} \right] dx_1 dq_1 dl. \end{aligned} \quad (18)$$

The procedure for the odd subproblem is identical. The governing equation (18) may be rewritten as an infinite set of linear equations

$$b_m^{(e)}(ka) = \sum_{n=1}^{\infty} K_{mn}^{(e)} a_n^{(e)}, \quad (19)$$

for $m = 1, \dots, \infty$, where the left-hand side terms, which depend on the incident wave number, are given by

$$b_m^{(e)}(ka) = -\pi A \int_{-1}^1 \cos(kaq_1) \psi_m^{(e)}(q_1) dq_1 \quad (20)$$

for incident plane waves; the counterpart for the odd subproblem follows in a similar manner. The right-hand side factors are given by

$$K_{nm}^{(e)} = I_{mn}^{(e)(1)} - (ak_p)^4 I_{mn}^{(e)(2)} - i(ak_p)^6 \epsilon I_{mn}^{(e)(3)} / ak_0. \quad (21)$$

Truncating this set of equations, and its counterpart for the odd subproblem, at some finite order N will give an approximate, but arbitrarily accurate (depending on the order of the truncation) solution to the original problem. Noting the symmetry with respect to m and n means that even for relatively large N , one need not evaluate an undue number of the K_{mn} .

We consider each term in K_{mn} in succession, dropping the superscripts. The corresponding pieces of the triple integral are

$$\begin{aligned} I_{mn}^{(1)} = 2\pi \int_{-1}^1 \psi_m''(x_1) \psi_n''(x_1) dx_1 \\ + 2\pi \delta [\psi_m'(1) \psi_n'(1) + \psi_m'(-1) \psi_n'(-1)]. \end{aligned} \quad (22)$$

The last two terms of this expression are zero in the limiting cases of clamped or hinged edge conditions. The second integral leads to

$$I_{mn}^{(2)} = 2\pi \int_{-1}^1 \psi_m(x_1) \psi_n(x_1) dx_1, \quad (23)$$

and the third integral is

$$I_{mn}^{(3)} = \pi \int_{-1}^1 \int_{-1}^1 H_0^{(1)}(k_0 a |x_1 - q_1|) \psi_m(q_1) \psi_n(x_1) dx_1 dq_1. \quad (24)$$

These are all relatively simple to evaluate numerically. The above could also be deduced directly from the boundary condition, but it is perhaps more natural to proceed from the integral equation. This approach is more readily adjusted to deal with other geometries.

As noted earlier, once we have the displacement we can use (10) to deduce the pressure fields. The far-field behavior of the scattered field is obtained by expanding the double integral in (10) [note the Green's function in (10) is itself an integral] for large $|\mathbf{q}|$. Taking the far-field variable as

$$\mathbf{q} = r(-\sin \phi, \cos \phi), \quad (25)$$

and using a steepest-descent approach, gives

$$p^{sc}(r, \phi) \sim \left(\frac{2}{\pi r k_0} \right)^{1/2} G(\phi) e^{i(k_0 r - \pi/4)} \quad (26)$$

for $k_0 r \gg 1$. In the even subproblem this is

$$G^{(e)}(\phi) = -\frac{2i(k_p a)^6}{(k_0 a)} \epsilon \sum_{n=1}^{\infty} a_n^{(e)} \times \int_{-1}^1 \psi_n^{(e)}(x_1) \cos(ak_0 x_1 \sin \phi) dx_1. \quad (27)$$

A similar expression holds for the odd subproblem (replacing the cosine with a sine). The total directivity $G(\phi)$ is the sum of the even and odd expressions. The coefficients $a_n^{(e)}$ depend on the type of forcing adopted, and are the solutions of (19). The directivity is hence specified entirely by the a_n ,

$$\begin{aligned} \eta^{sc}(x_1, 0) = & \frac{2A}{B(k_p^4 - k^4)} \left\{ e^{ikax} + [\delta(ka \cos k_p a \sin ka - k_p a \sin k_p a \cos ka) - \cos k_p a \cos ka \{(k_p a)^2 - (ka)^2\}] \right. \\ & \times \frac{\cosh k_p ax}{k_p a D(k_p a)} + [\delta(ka \cosh k_p a \sin ka - k_p a \sinh k_p a \cos ka) - \cosh k_p a \cos ka \{(k_p a)^2 + (ka)^2\}] \\ & \times \frac{\cos k_p ax}{k_p a D(k_p a)} + [\delta(k_p a \cos k_p a \sin ka - ka \sin k_p a \cos ka) + \sin k_p a \sin ka \{(k_p a)^2 - (ka)^2\}] \frac{i \sinh k_p ax}{k_p a F(k_p a)} \\ & \left. + [\delta(ka \sinh k_p a \cos ka - k_p a \cosh k_p a \sin ka) - \sinh k_p a \sin ka \{(k_p a)^2 + (ka)^2\}] \frac{i \sin k_p ax}{k_p a F(k_p a)} \right\}, \quad (28) \end{aligned}$$

where the denominators, which are related to resonances of the system, are

$$D(k_p a) = \delta(\cosh k_p a \sin k_p a + \sinh k_p a \cos k_p a) + 2k_p a \cosh k_p a \cos k_p a, \quad (29)$$

$$F(k_p a) = \delta(\sinh k_p a \cos k_p a - \cosh k_p a \sin k_p a) - 2k_p a \sinh k_p a \sin k_p a. \quad (30)$$

This, and the corresponding far-field directivity, are reproduced by the numerics. As noted earlier, the case $\delta=0$, corresponding to edge conditions $\eta = \eta'' = 0$, is amenable to Fourier analysis (although $\delta \neq 0$ is not). Hence we are independently comparing the performance of the expansion functions. Figure 2 shows the absolute error in the real part of the plate displacement for $k_0 a = 25$ and $\delta=0$, which corresponds to a wide plate, as well as the error for the solution calculated by a sum in Fourier modes. The present procedure converges quickly to the exact answer more quickly than the Fourier expansion. This is a consequence of the well-known properties of the Chebyshev polynomial approximation. For this specific edge condition, $\eta = \eta'' = 0$, we are not necessarily suggesting that Fourier modes are of little value—they are very simple to deal with numerically—but merely aim to demonstrate that the expansion functions we propose are accurate in this special case.

The formula (28) fails near zeros of $D(k_p a)$ or $F(k_p a)$. In this limit one can proceed to get asymptotic results using eigenfunction methods, or via a wide strip approximation and the Wiener–Hopf technique. A simpler approximation, at least for $M > 1$, is to replace the plate equation (1) by

as is the plate displacement. However, the latter quantity is simpler to compare to asymptotic approximations that we will develop and we shall hence concentrate on it.

III. ASYMPTOTIC RESULTS FOR LIMITING CASES

To demonstrate the accuracy of the above numerical scheme we compare it with various light fluid loading results. In the absence of fluid loading one can solve the plate equation, to get, for instance in the case of an incoming wave, the plate displacement

$$B \left(\frac{\partial^2}{\partial x^2} - k_l^2 \right) \left(\frac{\partial^2}{\partial x^2} - \mu^2 \right) \eta(x) = -p^{\text{inc}}. \quad (31)$$

The complex wave numbers k_l, μ are perturbations away from k_p and ik_p , respectively. For $\epsilon \ll 1$, the k_l are the leaky zeros corresponding to the leaky waves in the physical domain; they are in close proximity to the *in vacuo* wave numbers $\pm k_p$. The other wave numbers are at $\pm \mu$; these are in the proximity of $\pm ik_p$ and lead to rapidly decaying modes. If $M > 1$ these are given by the approximations¹¹

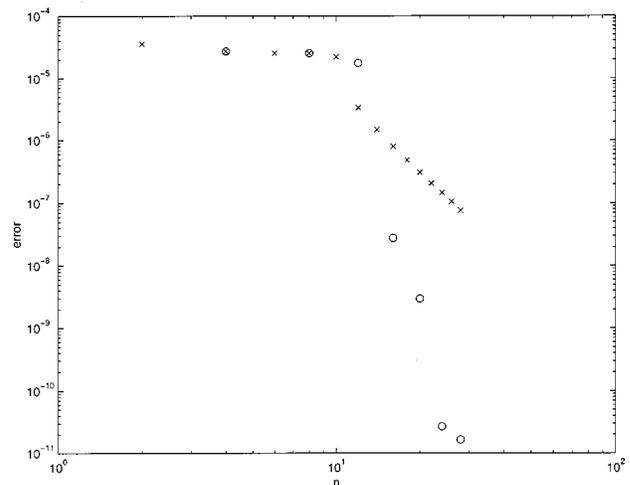


FIG. 2. Absolute error in the infinity norm of the real part of the plate displacement vs n , the order of truncation, for $k_0 a = 25$, $M = 1.5$, $\theta_i = 0$, $\epsilon = 0$, and $\delta = 0$. The circles show the error for the present scheme. The crosses show the error for the expansion in Fourier modes. Both circles and crosses are plotted for each value of n at which more resolution is achieved.

$$\mu \sim k_p \left(i - \frac{\epsilon}{4M(M^2+1)^{1/2}} \right), \quad (32)$$

$$k_l \sim k_p \left(1 + \frac{i\epsilon}{4M(M^2-1)^{1/2}} \right); \quad (33)$$

otherwise they may be found numerically as the appropriate zeros of $k^4 - k_p^4 + \epsilon k_p^6 / ik_0 \gamma_0(k) = 0$. This equation also has two real zeros, but these do not significantly affect the results in this light fluid loading limit. Thus this approximation assumes that the acoustic coupling is completely captured within the modified wave numbers. As shown⁸ via a com-

parison with more rigorous methods, this is not quite true (particularly for $M < 1$), but for practical purposes it does provide accurate solutions.

For brevity we will look at a wide strip approximation. For a wide strip we can deal with each plate edge independently and then look at the effect of the diffracted by one edge upon the other and vice-versa. This allows the natural separation of the problem into a sequence of semi-infinite problems which can be solved exactly using the Wiener-Hopf technique. However, it is much easier to use (31), and this has been justified previously.⁸ For normally incident plane waves one finds that

$$\eta(x,0) \sim \frac{2}{B(k_p^4(1+i\epsilon k_p^2/k_0^2))} \left(1 + \frac{(\mu^2 + i\mu\delta)\cos k_l x_1 - 2(k_l^2 \cos^2 k_l a + \delta k_l \sin k_l a) \cos(\mu x_1) e^{i\mu a}}{(k_l^2 \cos k_l a + \delta k_l \sin k_l a) - (\mu^2 + i\delta\mu)\cos k_l a} \right). \quad (34)$$

More explicit results showing the dependence upon ϵ , for $M > 1$, are obtained by substituting the zeros (33) into (34). At resonance, the displacement (and far-field directivity) is $O(1/\epsilon)$ rather than $O(1)$. For instance, in the simplest case, $\delta=0$, the resulting displacement is

$$\eta(x,0) \sim \frac{2i}{Bk_p^4} (-1)^{n+1} \cos\left(\left(n + \frac{1}{2}\right)\pi \frac{x}{a}\right) \frac{2M(M^2-1)^{1/2}}{\epsilon k_p a}, \quad (35)$$

for $k_p a = (n+1/2)\pi$. The plate displacements found near a resonance for $\epsilon=0.134$, $k_0 a = 7.209$, $M=1.5$, and $\delta=1$ are shown, normalized by a factor B/a^4 in Fig. 3.

Strictly, the approximation holds for $\epsilon \ll 1$ (and for $M > 1$). However it still appears to give good results for relatively large ϵ , such as 0.134 with the accuracy improving as ϵ decreases. In any case, for our purposes, it provides a useful confidence check upon the numerical results.

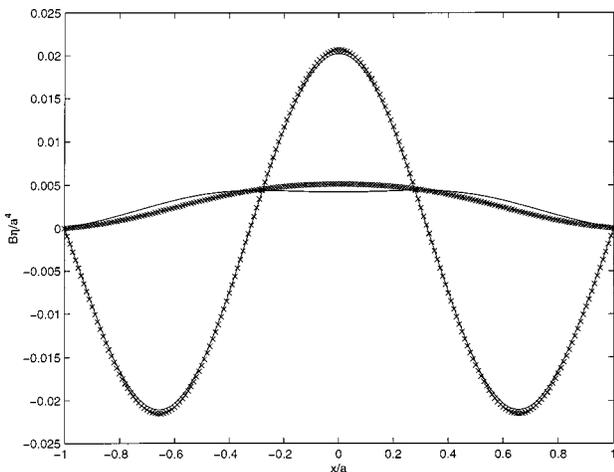


FIG. 3. Real and imaginary parts of the plate displacement $B\eta^{sc}(x,0)/a^4$: solid lines are the numerical solution (the imaginary part is the larger one), crosses are the analytical solution (28) vs x/a for $k_0 a = 7.209$, $M=1.5$, $\theta_i=0$, $\epsilon=0.134$, and $\delta=1$.

The plate displacements found for $\delta=0$, $\epsilon=0.134$, $k_0 a = 11.781$, and $M=1.5$ are shown, normalized by a factor B/a^4 in Fig. 4. The approximation (35) is also shown.

Once the plate displacements are found, the directivities follow from

$$G(\phi) = \frac{\rho\omega^2}{2i} \int_{-a}^a \eta(x,0) e^{ik_0 x \sin \phi} dx, \quad (36)$$

and there is close agreement between the numerical and asymptotic results. All numerical work is also checked utilizing power balance and reciprocity relations.⁸

IV. EIGENFUNCTION EXPANSIONS

The set of expansion functions we have utilized is not the only one that could be used. Indeed one could use any complete set that satisfied the edge conditions. Clearly, an alternative candidate set consists of the eigenfunctions asso-

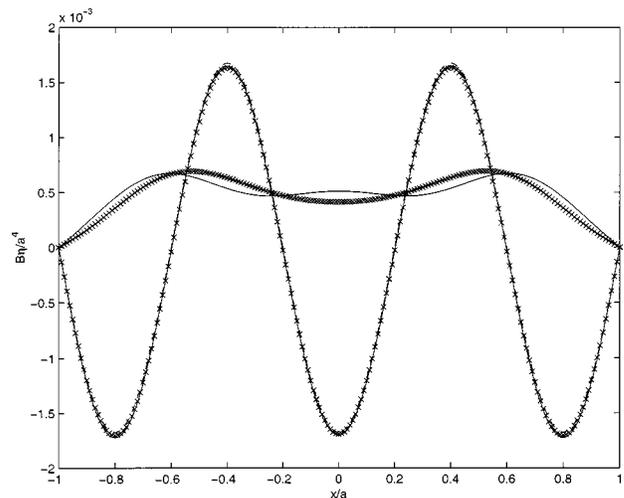


FIG. 4. Real and imaginary parts of the plate displacement $B\eta^{sc}(x,0)/a^4$ (solid lines and crosses as in Fig. 3 with the imaginary part again having larger magnitude) vs x for $k_0 a = 11.781$, $M=1.5$, $\theta_i=0$, $\epsilon=0.134$, and $\delta=0$. The dashed line shows the imaginary part of (35).

ciated with the *in vacuo* plate displacements. For example, for the special case of clamped edge conditions with even loadings, we could try

$$\eta^{sc}(x,0) = \frac{4a^4}{B} \sum_{n=1}^{\infty} a_n (\cos \lambda_n x \sinh \lambda_n + \sin \lambda_n \cosh \lambda_n x), \quad (37)$$

where the eigenvalues λ_n are the consecutive zeros of the transcendental equation

$$\sinh \lambda_n \cos \lambda_n + \sin \lambda_n \cosh \lambda_n = 0. \quad (38)$$

These are easily found using the approximate zeros $\lambda_n = \pi(n - 1/4)$, $n = 1, 2, \dots$ as the initial values in a Newton–Raphson scheme. These expansion functions satisfy the edge conditions exactly, and moreover they are orthogonal. There is no reason why these cannot be used as expansion functions in an identical manner to that which we have adopted in earlier sections, and indeed for comparative purposes we have done that. We have found these less flexible than the set of expansion functions we have developed, primarily because the expressions become more unwieldy, particularly for general δ and general loadings [(37) is just for the clamped case with even loading]. They are not easily carried across to three-dimensional problems, whereas the simpler sine and cosine functions are easy to program in. In addition, the eigenfunctions will not have the desirable features of Chebyshev approximation of the expansion functions we have developed. As our aim is to provide simple and flexible expansion functions we prefer not to use the *in vacuo* eigenfunctions. In addition, for heavy fluid loading these are not convenient expansion functions, as in that case, in rescaled coordinates, the structural inertia [the k_p^4 terms in the plate equation (1)] vanishes to leading order. Thus it is felt that a set of expansion functions satisfying the edge conditions exactly, but not tied in too closely with one piece of the boundary condition, is more flexible.

V. CONCLUSION

A fast, efficient, and flexible numerical scheme capable of dealing with a variety of plate edge conditions is pre-

sented; this is an extension of a technique often used for rigid plates. We focus on determining the plate displacement, as all other details can ultimately be determined from this quantity. Clearly it is important to verify that this class of expansion functions produces accurate results, particularly near resonance. To verify accuracy, the numerical results are compared with asymptotic results for light fluid loading, both near to, and far from, resonance. The results are also compared with a Fourier analysis of the special case of simply supported edges. Further numerical results for the clamped edge condition can be found in Ref. 8. The scheme for the more general edge conditions we have discussed here has similar accuracy and versatility.

A useful approximation (34) for light fluid loading, based upon varying the *in vacuo* wave number is also highlighted. Good agreement was found in all cases.

ACKNOWLEDGMENTS

S.G.L.S. was funded by a Lindemann Trust Fellowship administered by the English Speaking Union during the duration of this work, which was carried out at the Scripps Institution of Oceanography. R.V.C. acknowledges the support of an EPSRC Advanced Fellowship.

¹J. D. Achenbach and Z. L. Li, *Wave Motion* **8**, 225–234 (1986).

²F. L. Neerhoff and J. H. M. T. van der Hijden, *J. Sound Vib.* **93**, 523–536 (1984).

³D. G. Crighton, A. P. Dowling, J. E. Ffowcs Williams, M. Heckl, and F. G. Leppington, *Modern Methods in Analytical Acoustics* (Springer, New York, 1992).

⁴M. C. Junger and D. Feit, *Sound, Structures and their Interaction*, 2nd ed. (Acoustical Society of America, Westwood, 1986).

⁵F. G. Leppington, *Q. J. Mech. Appl. Math.* **29**, 527–546 (1976).

⁶J. F. M. Scott, *Philos. Trans. R. Soc. London, Ser. A* **338**, 145–167 (1992).

⁷P. O. Mattei, *J. Sound Vib.* **179**, 63–77 (1995).

⁸S. G. Llewellyn Smith and R. V. Craster, *Wave Motion* **30**, 17–41 (1999).

⁹D. G. Crighton and D. Innes, *J. Sound Vib.* **91**, 293–314 (1983).

¹⁰F. G. Leppington, E. G. Broadbent, and K. H. Heron, *Proc. R. Soc. London, Ser. A* **393**, 67–84 (1984).

¹¹D. G. Crighton, *J. Sound Vib.* **63**, 225–235 (1979).

Nonlinear propagation of laser-generated sound pulses in a water and granular medium

K. A. Naugolnykh

Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado/NOAA, Environmental Technology Laboratory, 325 Broadway, Boulder, Colorado 80303

S. V. Egerev, I. B. Esipov, and K. A. Matveev

N. Andreev Acoustics Institute, 4 Schverniki Street, Moscow, 117036, Russia

(Received 13 August 1998; revised 31 May 1999; accepted 29 August 1999)

Nonlinear propagation of finite-amplitude acoustic pulse in water and through a sample of water-saturated granular medium is considered. To generate high-intensity acoustic pulses laser generation of sound was used. The region of fluid perturbed by the laser acts as a volume-distributed source. In a fluid with weak light attenuation, a cylindrical source could be formed by a narrow laser beam. The nonlinear distortion of the cylindrical finite-amplitude wave in water is investigated. The measured rate of distortion corresponds to that calculated in the approximation of nonlinear acoustics. In a strongly light-absorbing medium, a wide (compared to the typical sound wavelength) laser beam produces a circular planar source. Such a source produces acoustical pulses of amplitude up to 3 MPa and duration about 1 μ s in different fluids. This source was used to investigate the propagation of high-intensity wide frequency band sound signals in a sample of water-saturated cobalt–manganese crust (CMC). Specific acoustical features of the crust such as nonlinear sound pulse distortion and the frequency dependence of attenuation, varying with the amplitude, are considered. Theoretical interpretation of the results is given. © 1999 Acoustical Society of America. [S0001-4966(99)04512-9]

PACS numbers: 43.25.Cb [MAB]

INTRODUCTION

Nonlinear propagation of finite-amplitude acoustic pulse in water and through a sample of water-saturated cobalt–manganese crust (CMC) is considered. The nonlinear evolution of cylindrical finite-amplitude sound wave in water and plane wave in crust are examined. The crust is a kind of ocean sediment that covers a vast bottom area in deep-sea regions of the Pacific Ocean. Porosity of the crust deposit is close to 50%–60%, and its structure looks like tightly packed spherical granules several tenths of a millimeter in diameter. The crust has a complex structure and can be considered as micro-inhomogeneous material that contains structure elements such as grains, grain joints, pores, cracks, etc. with a characteristic spatial scale significantly exceeding the interatomic spacing, but significantly less than the acoustic wavelength.^{1–6} The nonlinearity of micro-inhomogeneous material is not due to unharmonicity of the interatomic potential but due to nonlinear interaction of the mechanical elements of its structure. In such a medium the effect of anomalous nonlinearity can be observed connected with the fact that the medium responds differently to equivalent compressions and tensions. This effect can be described by the bilinear stress-strain relation.^{1,7,8} One of the interesting characteristics of granular materials is the frequency dependence of sound wave attenuation.⁹ Hamilton¹⁰ noticed on the basis of experimental evidence that attenuation in such materials scales with the first power of frequency from seismic to high ultrasonic frequencies while Kibblewhite supports a different view,⁹ concluding that attenuation in fluid-saturated sediments does not accurately follow the linear law. There are

different theoretical models of sound wave propagation in a structured medium.^{11–14} Biot¹¹ considered a fluid-saturated porous medium which consists of a solid elastic matrix, or skeletal frame, containing a compressible, viscous fluid. The frequency dependence of the sound attenuation for this model is in proportion to f^2 at low frequencies and at higher frequencies it is in proportion to $f^{1/2}$. Recently Buckingham¹⁴ has developed an internally consistent theory of acoustic wave propagation in a saturated, unconsolidated, marine sediment which is considered as a two-phase medium, consisting of solid particles and seawater, but possessing no rigid frame. In accord with this theory the attenuation coefficient of a sound wave is proportional to the first power of frequency.

To investigate crust acoustic properties high-intensity broad-frequency-band sound signals are needed. To produce such signals it is convenient to apply laser generation of sound. The physical mechanisms of this effect are diverse, but at moderate light intensity the effect of thermal expansion of the disturbed volume is predominant. The rapid heating of a separate region of the medium due to light absorption causes thermal expansion, which produces density perturbation, generating sound.^{15–17} At higher intensities of laser radiation, rapid evaporation of a fluid is developed. The resulting vapor ejection causes recoil pressure momentum being applied to the fluid, generating acoustic waves.¹⁸ The region of fluid perturbed by the laser acts as a volume-distributed source of sound—optoacoustical source. The features of this source depend on the spatial distribution of the optically induced acoustical sources, which are determined mainly by the light beam size and the light penetration dis-

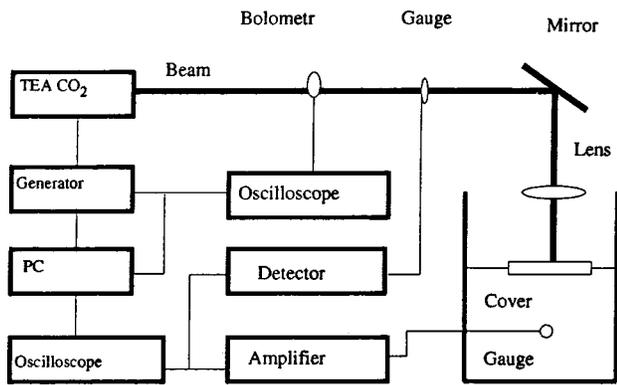


FIG. 1. Diagram of the experimental setup for laser generation of sound in fluids.

tance in comparison with the typical sound-pulse space scale $\lambda = c\tau$, where c is the sound velocity and τ is the sound pulse duration. In a strongly light-absorbing medium, a wide (compared to λ) laser beam produces a circular planar source; in a medium with weak light attenuation, a cylindrical source could be formed by a narrow laser beam.

In the present paper some features of the laser sound source are considered, for both planar and cylindrical sources produced in different fluids. The nonlinear distortion of the cylindrical finite-amplitude wave in water is examined. The measured rate of distortion corresponds to that calculated in the approximation of nonlinear acoustics. Then high-intensity, wide-frequency-band sound signals produced by the planar array are used for acoustic diagnostics of a sample of water-saturated cobalt–manganese crust (CMC).

The propagation of high-intensity sound pulses in the water-saturated samples of the crust reveals its nonlinear features. A wide frequency band of laser-generated signals allows measurement of the frequency dependence of acoustical properties, such as the sound attenuation and its variation with amplitude.

I. EXPERIMENT

A. Generation of sound. Planar sound source

To produce high-intensity broad-frequency-band sound pulses it is convenient to use laser generation of sound. The laser generation of the sound pulses in water, oil, and gasoline was investigated in the experiment represented schematically in Fig. 1. The experimental setup provided laser generation and registration of wide-frequency-band high-intensity pulses in different fluids. Sound pulses were generated in a cuvette $30 \times 30 \times 30 \text{ cm}^3$, which was filled with a fluid under investigation. On the liquid surface different acoustical tops made of Ge, Zn, Se and polyester can be positioned. The lower surface of the cover could be flat or concave, forming an acoustic lens. The thickness of the cover was about 1 cm. The cover provides a rigid boundary condition for sound generation in the adjacent fluid layer. The total energy of incident CO₂-laser pulse was measured by a pass-wire differential bolometer. The laser pulse duration was about $1 \mu\text{s}$ and its energy was varied in a range from several joules up to several dozen joules. The laser

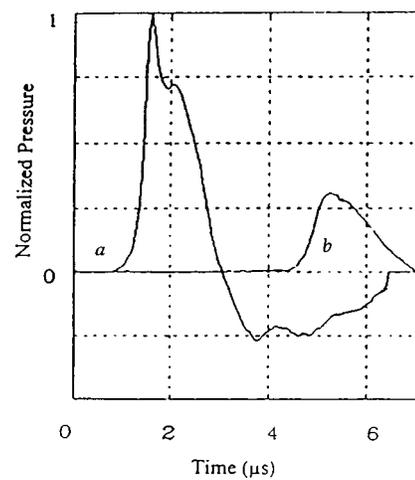


FIG. 2. Normalized oscillograms of the sound pulses; incident pulse (a), and pulse that transmitted through CMC sample (b). Laser pulse energy density is 0.08 J/cm^2 . Thermoelastic mechanism of sound generation.

radiation was focused on the fluid surface by means of a ZnSe lens and copper mirrors. The light spot diameter on the fluid surface was 2.0 cm.

The sound pulses propagating in fluid were registered by a specially designed wideband receiving system using polyvinylidene fluoride (PVDF) probes. The typical sensitivity of the probes was about 0.7 mV/Pa , and the bandwidth of the frequencies registered was in a range of $1 \text{ kHz} - 7 \text{ MHz}$. Gauges were placed in the near zone of the sound source in front of and behind the CMC sample along the laser beam axis. To avoid distortion of signals the gauge was adjusted so that its sensitive surface was parallel to the water surface which determines the wavefront geometry. The accuracy of adjustment depends upon the wavelength and the gauge size. The required accuracy of adjustment (in our case, 1.5 degrees) was obtained with the help of an auxiliary He–Ne laser, which produced a reference beam. The electrical signal produced by the gauge was transformed into a digital one by digital storage oscilloscope. The sampling period was 10 ns, duration of realization was $20 \mu\text{s}$, and the number of amplitude quantization levels was 256. The data stored in the oscilloscope were transferred to the PC for storing and processing.

The process of optical sound generation is varied with the increase of energy released in the fluid due to laser pulse absorption. At low-energy inputs (in comparison with specific heat of vaporization) the sound in the fluid was produced as a result of thermal expansion of a heated fluid volume. A typical sound pulse generated by thermoelastic mechanism in water is presented in Fig. 2(a). The pulse amplitude p is about 3 MPa and its duration is close to that of the optical pulse and is about $\tau \approx 2 \mu\text{s}$. At larger energy inputs the subsurface heating caused by the laser is accompanied by pronounced evaporation from the surface and thereby leads to recoil pressure momentum applied to the surface at the area corresponding to the laser beam cross section which forms the surface sound source. Both the thermoelastic and vaporization mechanisms contribute to the acoustical response in this regime; the relative contribution of mechanism depends upon the energy input.¹⁹ This mecha-

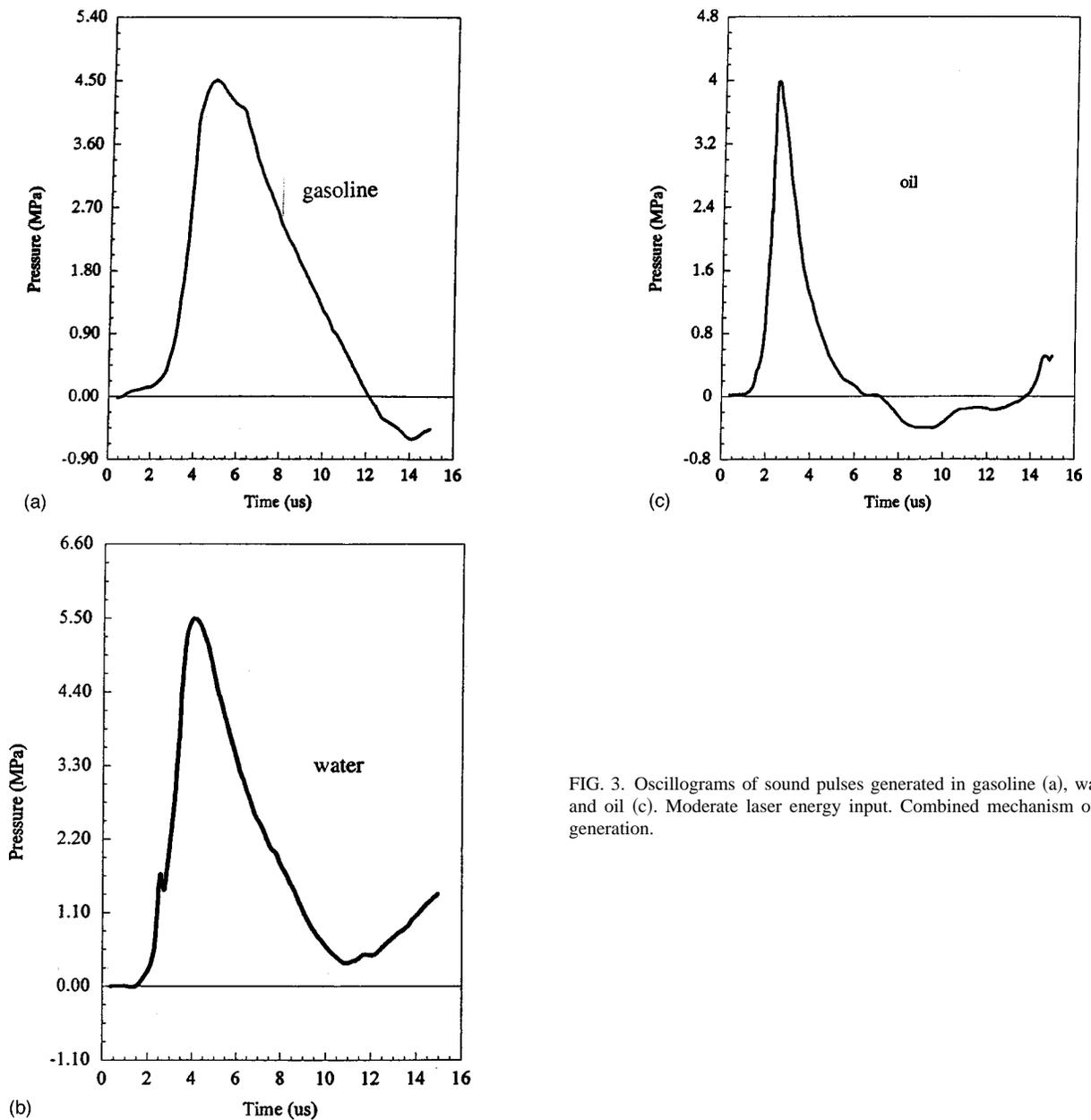


FIG. 3. Oscillograms of sound pulses generated in gasoline (a), water (b), and oil (c). Moderate laser energy input. Combined mechanism of sound generation.

nism of sound generation in different fluids was also investigated. Some results are presented in Fig. 3 where the oscillograms of acoustical pulses generated in gasoline, water, and oil at moderate input energy density are given. Acoustical pulses were registered in the focal region of a concave cover. Due to the difference in thermodynamic properties of fluids (boiling temperature, evaporation heat) the sound pulses produced in these fluids by the same laser pulse differ. These experiments demonstrate that laser generation of sound can be used to produce high-intensity wide-frequency-band acoustical signals in different fluids.

B. Generation of sound. Cylindrical sound source

In a weakly light-absorbing medium, where the light propagation length is long in comparison to the beam cross-section radius, the cylindrical volume is perturbed as a result of laser radiation action. Optoacoustical conversion in this regime could be influenced by the presence of light-

absorbing particles in a fluid. Real liquids often contain impurities which give rise to additional contributions to laser sound generation. At small energy input the linear thermo-optical effect dominates in the process of sound generation. The presence of the light-absorbing impurities in a liquid changes the pattern of sound generation. The efficiency of the process increases if the energy input exceeds the threshold of vaporization of liquid surrounding the absorbing particles. As a result of liquid vaporization, bubbles appear and their expansion leads to effective sound generation. At high energy input, light-absorbing particles initiate the optical breakdown of the liquid.²⁰ In our experiment we investigated laser sound generation at moderate energy input when both thermal expansion and vaporization mechanisms contributed to sound generation, giving rise to generation of intense sound pulses. The diagram of the experiment for measuring the radiation produced by the cylindrical source in water is presented in Fig. 4.

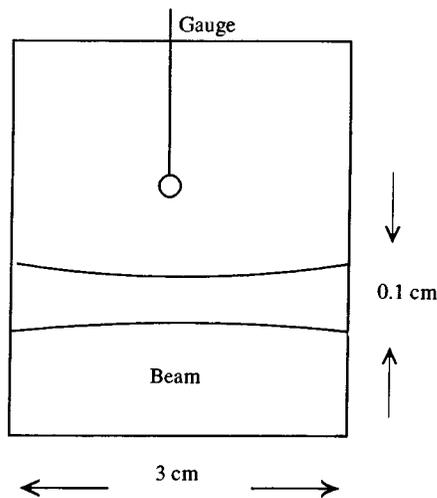


FIG. 4. Diagram of the experimental setup for laser generation of cylindrical waves in water.

A XeCl laser pulse with a wavelength of $\lambda=308$ nm, intensity of 2 J/cm^2 , and a total duration of $\tau \approx 12$ ns was focused in a water tank of 3-cm width to produce the cylindrical beam of about $r_0=0.1$ -cm radius. The acoustic pulse of amplitude about 2 MPa at a distance r_0 was generated. The acoustic pulse profile was recorded by a wide-frequency-band piezoelectric transducer, which was placed at distances in a range of 1–3 cm from the beam axis. The nonlinear effects in the sound wave lead to steepening of the front of the high-intensity sound pulse. The steepness of the pulse front, measured experimentally as a function of distance from the beam axis, is presented in Fig. 5 by circles. To consider nonlinear wave distortion theoretically one has to take into account both nonlinear and divergency effects.²¹ In the approximation of nonlinear acoustics this process can be described by the equation for dimensionless pressure perturbation u :^{22,23}

$$u_x - uu_\xi = 0. \quad (1)$$

Here

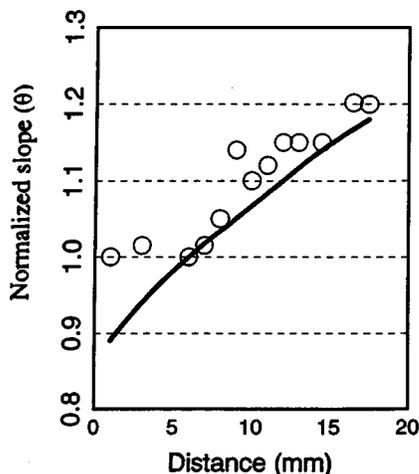


FIG. 5. Steepening pulse front as a function of distance traveled by the sound wave: theory (solid curve) and experiment (circles). The slope θ is normalized to its value at a distance of $r=0.6$ cm.

$$u = \frac{p}{p_0} \sqrt{\frac{r}{r_0}}, \quad (2)$$

where p is the pressure perturbation at distance r , $p = P - P_0$, P and P_0 are the pressure in the fluid and its equilibrium value, respectively, $\xi = \omega y$ is the normalized time, and $\omega = 2\pi/\tau$ is the characteristic frequency, τ being the time scale. Then

$$y = t - \frac{r}{c_0} \quad (3)$$

is the accompanying coordinate,

$$x = 2\sigma_0 \left[\sqrt{\frac{r}{r_0} - 1} \right] \quad (4)$$

is the normalized space coordinate, and

$$\sigma_0 = \varepsilon M \frac{\omega r_0}{c_0}, \quad M = \frac{p_0}{\rho_0 c_0^2}, \quad \varepsilon = \frac{n+1}{2}. \quad (5)$$

Here M is the Mach number of the sound perturbation, c_0 and ρ_0 are the sound speed and the density of the liquid, respectively, p_0 and τ are the initial amplitude of the sound pulse and its duration at distance r_0 , ε is the nonlinear parameter of the liquid, and n is the adiabatic index of the liquid. Here $n=7.15$ for the water, so $\varepsilon \approx 4$.

Let us consider the propagation of perturbation, generated at $r=r_0$, the value of r_0 is determined by the laser beam cross-section size. The corresponding solution of Eq. (1) for the simple wave can be presented as follows:

$$u = f(\xi + ux), \quad (6)$$

where $f(\xi)$ is the function determined by the boundary condition at $x=0$. This solution can be presented in the following form:

$$\xi = -ux + F(u), \quad (7)$$

where $F(u)$ is the function, reciprocal to $f(\xi)$. The first term on the right-hand side of the Eq. (7) describes the nonlinear steepening of the wave while the second presents the initial pulse profile $F(u)$. Schematically this solution is presented in Fig. 6, where the shape of nonlinearly distorted sound pulse [Fig. 6(b)] is given as a result of summation of the initial profile [Fig. 6(a)] and the straight line [Fig. 6(c)] in accord with Eq. (7). The pulse front steepness can be characterized by the ratio $S = u_m / \xi_m$, where u_m is the normalized pulse amplitude and ξ_m is the pulse front duration. Its initial value is $\xi_0 = F(u_m)$, so the initial value of slope is $S_0 = u_m / \xi_0$. Due to nonlinear distortion with the increasing of distance traveled by the wave the front duration, ξ_m decreases according to

$$\xi_m = F(u_m) - u_m x = \xi_0 \left(1 - \frac{u_m}{\xi_0} x \right), \quad (8)$$

and the slope grows in accord with

$$S(r) = \frac{u_m}{\xi_m} = \frac{u_m}{\xi_0 (1 - (u_m / \xi_0) x)}. \quad (9)$$

Substituting Eqs. (4) and (5) into (9), one obtains

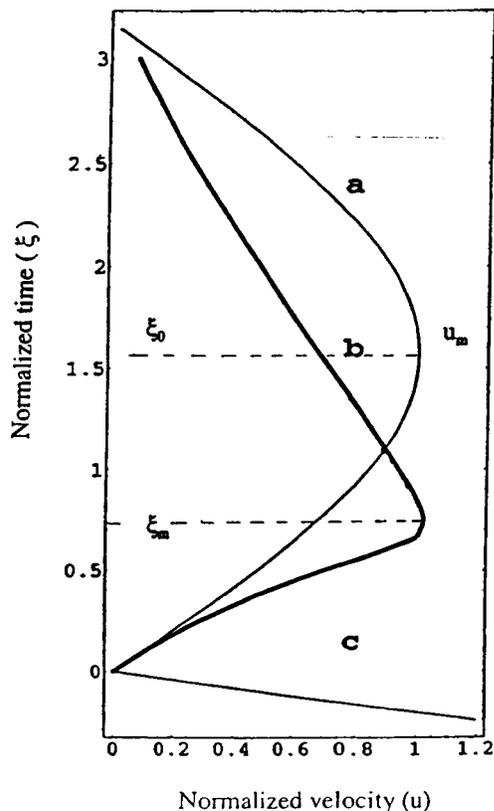


FIG. 6. Diagram of nonlinear distortion of sound pulse. Initial profile (a) and distorted profile (b), obtained by summing up initial profile and the straight line (c) in accord with Eq. (7).

$$S(r) = \frac{u_m}{\xi_0} \left(1 - \frac{u_m}{\xi_0} 2\varepsilon \frac{p_0}{\rho c_0^2} \frac{\omega r_0}{c_0} \left[\sqrt{\frac{r}{r_0}} - 1 \right] \right)^{-1}. \quad (10)$$

The rate of the slope increasing is in proportion to the sound pulse amplitude and the square root of the distance traveled by the wave.^{21,23} To compare with the experimental data the slope S was normalized to its value at $r_1 = 0.6$ cm, as it was made at the presentation of experimental data. The theoretical dependence of the normalized slope

$$\theta = \frac{S(r)}{S_1}, \quad (11)$$

where $S_1 = S(r_1)$, $r_1 = 0.6$ cm, upon the distance obtained at $p_0 = 2$ MPa, $r_0 = 0.1$ cm, $\tau = 10^{-7}$ s, and $\varepsilon = 4$, is presented in Fig. 6 by the solid curve that corresponds fairly well to the experimental data.

C. Nonlinear propagation of a sound pulse through a sample of cobalt-manganese crust (CMC)

The laser setup of sound generation was used to investigate the propagation of high-intensity sound signals through the sample of the crust. The cobalt-manganese crust (CMC) sample with a transverse dimension 4.0 cm and 1.5-cm thickness was placed at 5.0-cm distance from the sound generation region (Fig. 7). A number of CMC samples were taken from different places at the North Pacific bottom. Measured elasticity threshold of them at compression (stress limit) spread from 0.6×10^6 N/m² to 7.6×10^6 N/m² and with strain limit spread in the range of $(0.05-0.63) \times 10^6$ N/m².

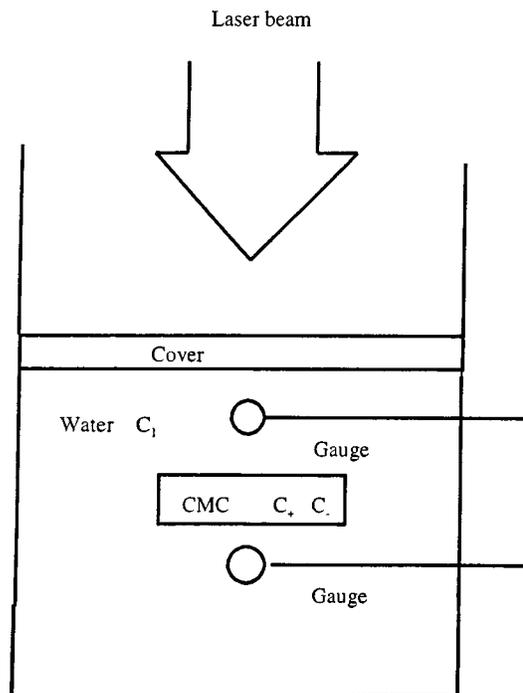


FIG. 7. Diagram of the experimental setup for acoustic diagnostic of a sample of CMC; c_+ and c_- are the sound velocities at the phase of compression and dilatation of the sound pulse.

The measured CMC density is equal to 1.5×10^3 Kg/m³ and sound velocity of the low intensity sound is equal to 1.5×10^3 m/s, so the impedance of water-saturated CMC is close to that of water, which is why the main part of the incident pulse energy was transmitted through the sample and the sound reflection effect is negligible.

In Fig. 2 typical oscillograms of both the incident sound pulse [Fig. 2(a)] and the pulse that passed through the sample [Fig. 2(b)] are presented. One can notice that as a result of sound attenuation in the CMC, the amplitude of the pulse decreases and the top of the pulse becomes more flat. In Fig. 8 normalized oscillograms of the sound pulse passed through

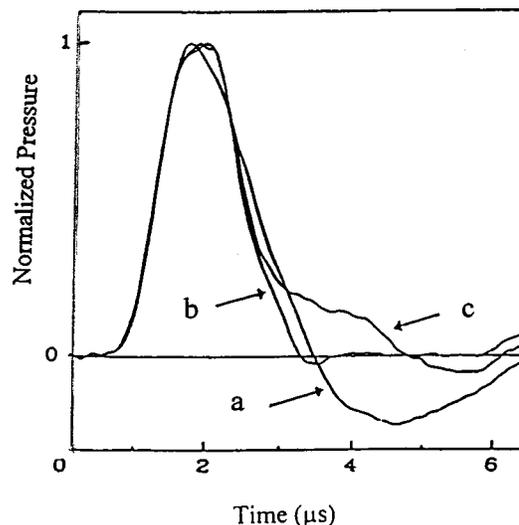


FIG. 8. Normalized oscillograms of the sound pulses transmitted through the sample of CMC. Laser radiation energy density is 0.08 J/cm² (a), 0.8 J/cm² (b), and 1.20 J/cm² (c).

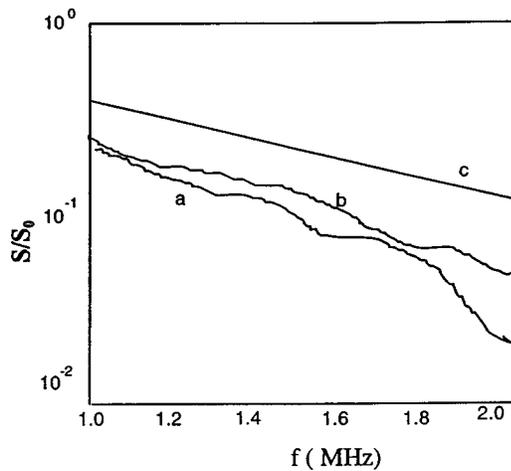


FIG. 9. Transition characteristics of the CMC sample. Low-amplitude sound pulse (a), finite-amplitude sound pulse (b), and theoretical dependence according to Eq. (25) (c).

the CMC sample are presented. It is seen that the front of the transmitted pulse is slightly retarded as the pulse amplitude increases as well as the negative phase of the pulse gradually decreases and completely vanishes. This pattern of nonlinear distortion is different from the standard nonlinear steepening and is apparently connected with the effect of modification of the complex crust structure under the action of the finite amplitude sound pulse.⁷

Interesting features of sound propagation in crust can be revealed by comparison of the spectrum of incident and passed pulses. The ratio of transmitted and incident pulse spectrum is presented in Fig. 9. It is seen that the low-amplitude sound pulse attenuation grows up linearly with the frequency in the frequency range from 1 MHz up to 1.7 MHz [Fig. 9(a)]. At higher frequencies the attenuation factor grows much faster with frequency than the typical dependence f^{1-2} for water-saturated porous media.¹⁰⁻¹⁴ Abnormal sound attenuation in the high frequency band can be connected with the sound energy capture and its localization as oscillating modes of the structure.²⁴ At high intensity of the probing pulse the granular structure is damaged, which apparently leads to the vanishing of additional mechanisms of sound attenuation and the frequency dependence of the attenuation coefficient becomes linear [Fig. 9(b)].

II. DISCUSSION

The main features of the sound pulse propagation through the granular medium CMC sample could be explained in terms of the effects of specific nonlinearity and dissipation. The elasticity of the granular medium is different under compression and expansion. This effect could be modeled by the bilinear equation of state¹

$$p = \begin{cases} c_+^2 \rho, & \rho \geq 0, \\ c_-^2 \rho, & \rho \leq 0, \end{cases} \quad (12)$$

presented in Fig. 10. Here p and ρ are the perturbation of pressure and density, and c_+ and c_- are the sound velocities in the granular medium under compression and dilatation.

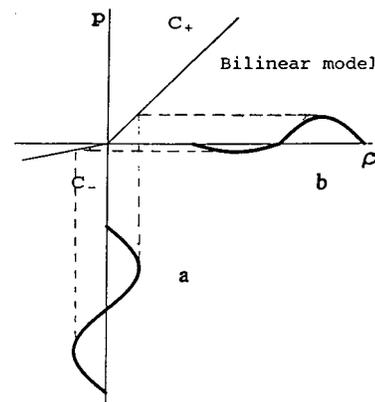


FIG. 10. Transmission of the wave through the bilinear medium, incident wave (a), and transmitted wave (b). Here c_+ and c_- are the sound velocities in the bilinear medium under compression and dilatation.

Consider the propagation of a plane wave through a layer of a granulated medium governed by Eq. (12). The instantaneous pressure p_i of the pulse transmitted at the interface of two fluids (i, k) is determined by their impedances $\rho_i c_i$, $\rho_k c_k$:

$$p_t = \frac{2\rho_k c_k}{\rho_i c_i + \rho_k c_k} p_i, \quad (13)$$

p_i being the amplitude of the incident wave. Applying this relationship to the sound signal transmission at the interface water/sample ($i=1, k=2$) and sample/water ($i=2, k=3$) separately for the compression and the dilatation phase of the sound pulse and taking into account that

$$c_+ = c_1, \quad c_- \ll c_1, \quad \rho_2 = \rho_1, \quad (14)$$

one obtains the equations for the instantaneous pressure in the transmitted wave:

$$p_{+3} = p_{+2} = p_{+1}, \quad (15)$$

$$p_{-3} = 2p_{-2} = 4(c_-/c_1)p_{-1}. \quad (16)$$

Here p_{+3} and p_{-3} correspond to the pressure in the phase of compression and dilatation, respectively. It is seen that in the bilinear approximation the amplitude of the compression pulse does not change [Eq. (15)] while the amplitude of the dilatation phase of the pulse transmitted through the layer of granulated medium decreases proportional to the ratio of the sound velocity of the sample to that of the surrounded fluid [Eq. (16)]. Qualitatively the effect of bilinear distortion that decreases the intensity of the dilatation phase in comparison to that of compression phase of the pulse is demonstrated in Fig. 11, where both the incident pulse (a) and transmitted one (b) are presented.

In addition to the bilinear distortion, the sound pulse propagated through the granular medium undergoes to essential attenuation. Recently Buckingham¹⁴ developed an internally consistent theory of acoustic wave propagation in a marine sediment considered as an unconsolidated two-phase medium. Several interesting results emerge from this theory including the linear dependence of the sound wave attenuation coefficient upon the frequency. This result corresponds to that of a substantial body of experimental data indicating

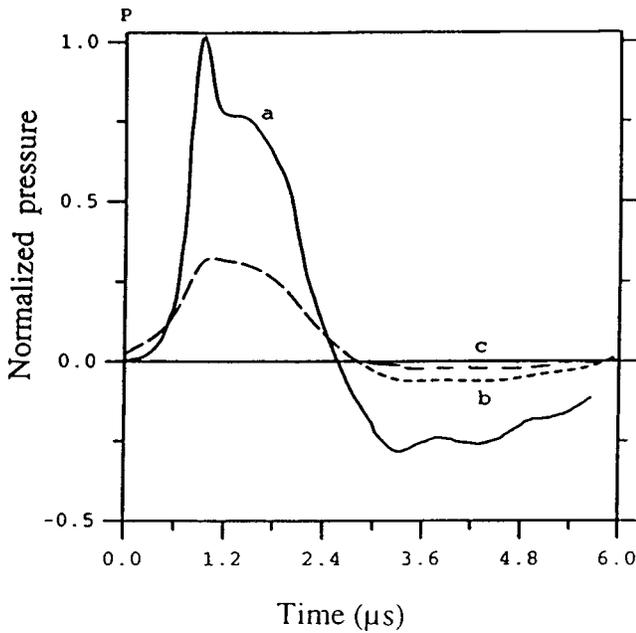


FIG. 11. Numerical modeling of the sound pulse propagation through the bilinear dissipative medium. Incident pulse (a); effect of attenuation of the transmitted pulse (b) modeled according to Eq. (17); bilinear distortion and attenuation of the transmitted pulse (c) calculated in accord with Eq. (24).

that sound attenuation in a porous medium varies as the first power of frequency over an extended frequency range. It is natural, therefore, to use this theory to interpret our experimental results. The sound attenuation in the granular unconsolidated materials is governed by the equation

$$p = p_0 \exp(-\alpha x), \quad (17)$$

where x is the distance travelled by the wave and α is the sound absorption coefficient:¹⁴

$$\alpha = \frac{\omega \beta}{c}. \quad (18)$$

Here ω is the wave frequency, c is the speed of the sound wave, and

$$\beta = \frac{\pi n \chi}{4(1 + \chi)^{3/2}} \quad (19)$$

is the magnitude of the dimensionless loss tangent, n is the parameter of the relaxation equation of state,

$$h(t) = u(t) \frac{t_0^{n-1}}{t^n}, \quad (20)$$

t is time, t_0 is the scaling constant, $u(t)$ is the step function, and η_f , and λ_f are the viscosity coefficients and

$$\chi = \frac{(\frac{4}{3})\eta_f + \lambda_f}{\rho c^2 t_0}. \quad (21)$$

Coefficient $\beta=0.015$ for water-saturated sediment of granules of diameter $d=(1-5)10^2 \mu$ and the sound speed c could be taken as equal to that of the pure water. Then

$$\alpha = 10^{-7} \omega \text{ cm}^{-1}. \quad (22)$$

The sound attenuation coefficient at the distance of $L = 1.5 \text{ cm}$ (the sample thickness) is

$$\alpha L = 10^{-7} 3 \pi f, \quad (23)$$

so the instantaneous pressure in the pulse transmitted through the sample can be expressed by the equation

$$p_t = p_{\pm 3} \exp(-10^{-7} 3 \pi f x), \quad (24)$$

where $p_{\pm 3}$ are defined by the Eqs. (15) and (16).

The spectral transfer function is

$$S_{\text{out}} = S_{\text{inp}}(-3 \pi f 10^{-7}), \quad (25)$$

where f is the frequency of the sound wave (Hz) and S_{inp} and S_{out} are the energy spectrum of incident and transmitted pulse, respectively. This equation is presented in Fig. 9(c). It is interesting to note that the experimental data on the attenuation of the finite amplitude wave [Fig. 9(b)] fits better to the theoretical prediction—the sample becomes “more granulated” when probed by the high-intensity pulse.

Taking into account both the bilinear distortion [Eqs. (14) and (15)] and the sound attenuation in the granulated medium [Eq. (17)] it is possible to model the sound pulse propagation through the CMC sample. The incident pulse form [Fig. 11(a)] is taken in accordance with the experimental data [Fig. 2(a)]. Then the application of the Fourier transform gives the spectral amplitudes of the initial sound signal. The effect of attenuation is accounted for by multiplication with the damping factor according to Eq. (17) and performing the inverse Fourier transform; the result is presented in Fig. 11(b). It is seen that frequency-dependant attenuation decreases the amplitude of the pulse and makes smooth its shape. Then the bilinear distortion of the transmitted sample pulse was modeled. The result presented corresponds to $(c_-/c_+) = 0.1$, which is in the range of values of this parameter obtained by elasticity measurement. The transmitted pulse form which is affected both by the attenuation and bilinearity in accord with Eq. (24) is shown in Fig. 11(c). It corresponds qualitatively to the experimental data obtained.

III. CONCLUSIONS

Nonlinear propagation of a finite-amplitude acoustic pulse in water and through a sample of water-saturated cobalt–manganese crust (CMC) is considered. The crust has a complex structure and, respectively, specific acoustical features. To investigate crust acoustic properties high-intensity broad-frequency-band sound signals are needed. To produce such signals it is convenient to apply laser generation of sound. The interaction of laser radiation with an absorbing fluid leads to the formation of a volume-distributed source. In a strongly light-absorbing medium, a wide (compared to the typical sound wavelength) laser beam produces a circular planar source. In a fluid with weak light attenuation, a cylindrical source could be formed by a narrow laser beam. Both regimes of laser sound generation are studied experimentally. The nonlinear distortion of cylindrical finite-amplitude sound pulses was considered. The measured rate of distortion corresponds to that calculated in the approximation of nonlinear acoustics. The propagation of laser-generated, high-intensity sound signals through a sample of water-saturated

cobalt–manganese crust (CMC) reveals the following specific acoustic features: as the probing pulse amplitude increases, (1) strong nonlinear distortion of the sound pulse is observed and the amplitude of dilatation phase of the transmitted pulse gradually decreases and completely vanishes with increasing amplitude; (2) abnormal sound attenuation in the high-frequency band diminishes and the frequency dependence of the sound attenuation coefficient tends to be linear. The measurement of the sound attenuation coefficient in a wide frequency range corresponds fairly well to Buckingham's theory of sound attenuation in unconsolidated granular media. Numerical modeling of the effects of the nonlinear distortion and sound attenuation of the sound pulse propagating through the granular medium on the base of a bilinear model of unconsolidated medium qualitatively explains the experimental results.

ACKNOWLEDGMENTS

We are grateful to Dr. J. Shaw for the help in the preparation of this paper. The work was supported in parts by the Russian Foundation of Fundamental Research, project code 95-02-06353.

- ¹K. A. Naugolnykh and L. A. Ostrovsky, *Nonlinear Wave Processes in Acoustics* (Cambridge U.P., Cambridge, 1998).
- ²L. A. Ostrovsky, "Wave processes in media with strong acoustic nonlinearity," *J. Acoust. Soc. Am.* **90**, 3332–3338 (1990).
- ³L. Bjorno, "Finite-amplitude wave propagation through water-saturated marine sediments," *Acustica* **38**(4), 195–200 (1977).
- ⁴J. M. Hovem, "The nonlinear parameter of saturated marine sediments," *J. Acoust. Soc. Am.* **66**, 1463–1467 (1979).
- ⁵I. I. Belyaeva, V. Y. Zaitzev, and L. A. Ostrovsky, "Nonlinear acoustoelastic property of granulated media," *Acoust. Zurn.* **39**, 25–32 (1993) [*Acoust. Phys.* **39** (1), 11–14 (1993)].
- ⁶V. F. Nesterenko, "Nonlinear pulse propagation in granulated media," *Prikl. Mekh. Tekh. Fiz.* **5**, 136–148 (1983).
- ⁷A. V. Nikolaev, "The seismic properties of a friable medium," *Izvestiya, Phys. Solid Earth* **2**, 23–31 (1967).
- ⁸S. A. Ambartsumyan, *Different Modulus Theory of Elasticity* (Nauka, Moscow, 1982).
- ⁹A. C. Kibblewhite, "Attenuation of sound in marine sediments: A review with emphasis on new low frequency data," *J. Acoust. Soc. Am.* **86**, 716–738 (1989).

- ¹⁰E. L. Hamilton, "Acoustic properties of sediments," in *Acoustics and the Ocean Bottom*, edited by A. Lara-Sarnz, C. Ranz Cuierra, and C. Carbo-Fite (Consejo Superior de Investigaciones Cientificas, Madrid, 1987), pp. 3–58.
- ¹¹M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range," *J. Acoust. Soc. Am.* **28**, 179–191 (1956).
- ¹²I. Yu. Belyaeva, L. A. Ostrovsky, V. Yu. Zaitsev, V. Stefan, and A. M. Sutin, "Comparison of linear and nonlinear elastic moduli for reservoir rock by use of granular medium model," *J. Acoust. Soc. Am.* **99**, 1360–1365 (1996).
- ¹³E. A. Koen Van Den Abeele, "Elastic pulsed wave propagation in media with second- or higher-order nonlinearity. I. Theoretical framework," *J. Acoust. Soc. Am.* **99**, 3334–3345 (1996).
- ¹⁴M. J. Buckingham, "Theory of acoustic attenuation, dispersion, and pulse propagation in unconsolidated granular materials including marine sediments," *J. Acoust. Soc. Am.* **102**, 2579–2596 (1998).
- ¹⁵E. F. Carome, C. E. Moeller, and N. A. Clark, "Intense ruby-laser induced acoustic impulses in liquids," *J. Acoust. Soc. Am.* **40**, 1462–1466 (1966).
- ¹⁶P. J. Westervelt and R. S. Larcen, "Laser-excited broadside array," *J. Acoust. Soc. Am.* **54**, 121–122 (1973).
- ¹⁷E. F. Kozyaev and K. A. Naugolnykh, "The thermal acoustooptical effect," *Sov. Phys. Acoust.* **22**, 206–207 (1976).
- ¹⁸K. A. Naugolnykh, O. V. Puchenkov, V. V. Zosimov, and A. E. Pashin, "Surface motion induced by the interaction of pulsed laser radiation with highly absorbing dielectric fluids," *J. Fluid Mech.* **250**, 385–421 (1993).
- ¹⁹V. N. Alekseev, S. V. Egerev, K. A. Naugolnykh, O. B. Ovchinnikov, A. E. Pashin, and O. B. Puchenkov, "Acoustic diagnostics of transient interaction processes between optical radiation and a highly absorbing dielectric fluids," *Sov. Phys. Acoust.* **33**, 561–565 (1987).
- ²⁰S. V. Egerev, K. A. Naugolnykh, A. E. Pashin, and Ya. O. Simonovsky, "Nonlinear optoacoustical pulse phenomena in complicated media," in *Nonlinear Acoustics*, edited by K. Naugolnykh and L. Ostrovsky (AIP, New York, 1994), p. 76.
- ²¹D. T. Blackstock, "On plane, spherical, and cylindrical sound waves of finite amplitude in lossless fluid," *J. Acoust. Soc. Am.* **36**, 217–219(L) (1964).
- ²²K. A. Naugolnykh, "Absorption of finite amplitude waves," in *High Intensity Ultrasonic Fields*, edited by L. Rozenberg (Plenum, New York, 1968).
- ²³K. A. Naugolnykh, S. A. Soluyan, and R. V. Khokhlov, "Cylindrical waves of finite amplitude in dissipative medium," *Moscow University Physics Bulletin* **3** (4), 65–71 (1962).
- ²⁴I. B. Esipov, V. V. Zosimov, and K. I. Matveev, "Propagation of acoustic pulse of the finite amplitudes in granulated medium," *Acoust. Phys.* **43**, 558–563 (1997).

Nonlinear longitudinal waves in inhomogeneously predeformed elastic media

Arvi Ravasoo^{a)}

The Ångström Laboratory, Solid Mechanics, Uppsala University, Box 534, SE-751 21 Uppsala, Sweden

(Received 5 March 1997; revised 15 September 1999; accepted 16 September 1999)

The quasi one-dimensional problem of nonlinear longitudinal wave propagation in the elastic medium undergoing inhomogeneous plane prestrain is investigated theoretically. The analytical solution to describe the propagation of the wave with an arbitrary smooth initial profile is derived. The influence of the magnitude of the prestrain intensity on the distortion of the wave profile is studied. The sine-wave propagation in the medium subjected to the distributed static load is considered in more detail. The dependence of the sine-wave characteristics on the physical and geometrical properties of the medium and on the parameters of the predeformed state is clarified. The possibility to enhance the efficiency of ultrasonic nondestructive testing making use of the nonlinear effects of wave propagation is discussed. The algorithm to evaluation of the parameters of plane strain on the basis of wave profile evolution data is proposed. © 1999 Acoustical Society of America. [S0001-4966(99)08112-6]

PACS numbers: 43.25.Dc [MAB]

INTRODUCTION

The problem of wave propagation in predeformed media is of great interest to many researchers because of its importance in nondestructive testing (NDT). Since the inhomogeneity of predeformation field induces additional problems in NDT, the amount of information to be determined increases. Therefore, the necessity to extract more information from the wave propagation measurement data arises.^{1,2} One possible solution in this direction is to pose the problem of acousto-diagnostics of inhomogeneously predeformed media on the basis of nonlinear wave propagation data.³

The purpose of this work is to develop in detail⁴⁻⁶ the mathematical basis of acousto-diagnostics of inhomogeneously predeformed elastic medium (structural elements) and to investigate the usefulness of nonlinear wave propagation data in NDT. We study the quasi-one-dimensional problem of nonlinear propagation of a longitudinal wave with an arbitrarily smooth initial profile in an inhomogeneously predeformed elastic medium. The predeformed state of the medium is proposed to be plane strain. We describe this state with polynomials and demonstrate the efficiency of the method of polynomial approximation of the solution to the equations of equilibrium of nonlinear theory of elasticity. This method is widely used in the linear case.⁷

The asymptotic solution of the problem of wave propagation in an inhomogeneously predeformed medium is derived. This solution enables one to determine the dependencies of the wave characteristics on the parameters of the elastic medium and the predeformed state. Nonlinear properties of the medium and nonlinear effects of wave propagation are taken into account. The evolution of nonlinear effects and the interaction of the effects of nonlinearity and inhomogeneity are analyzed for various values of intensity of the excited wave.

The possibility to utilize this solution in NDT is studied in the case of sine-wave propagation in an elastic medium with traction-free upper and loaded lower surfaces. The applied static load distributed quadratically has normal and shear components. With the view to evaluating the predeformed state, the wave process is excited on a traction-free surface and the wave-induced stress is recorded for the same surface.⁸ For this case the analytical expressions for wave characteristics as functions of the parameters of the predeformed state are derived. On the basis of these cumbersome expressions, plots of the wave characteristics versus parameters of the predeformed state are presented. These plots facilitate analysis of the sensitivity of wave characteristics to the parameters of the predeformed state. As an example for nondestructive evaluation, they are applied to the determination of the plane strain caused by uniformly distributed normal and shear loads from wave profile distortion data.

I. THEORY

A. Nonlinear equations of motion

We consider an isotropic and homogeneous elastic medium and distinguish three different states. Initially, the medium is in the undeformed natural state. At the instant $t = 0$, this medium is subjected to external forces and at a later time it is deformed. The predeformed state of the medium is proposed to be static; i.e., independent of time. At some instant $t > 0$, a longitudinal wave is excited in the predeformed medium.

I. THEORY

A. Nonlinear equations of motion

Deformation of the medium is described on the basis of the nonlinear theory of elasticity⁹ taking the physical and geometrical nonlinearity into account. Our attention is confined to small but finite two-dimensional deformations. We consider the case of the predeformed state that corresponds

^{a)}Permanent address: Department of Mechanics and Applied Mathematics, Institute of Cybernetics at Tallinn Technical University, Akadeemia tee 21, 12618 Tallinn, Estonia. Electronic mail: arvi@ioc.ee

to plane strain; i.e., one of the displacement vector components is equal to zero.

It is assumed that deformations and their spatial derivatives caused by the propagating wave are much larger in the propagating direction than in the orthogonal direction. This leads to the quasi-one-dimensional problem of longitudinal wave propagation in a two-dimensional medium undergoing inhomogeneous plane strain:

$$[1 + k_1 U_{1,1}^0(X_1, X_2) + k_2 U_{2,2}^0(X_1, X_2)] U_{1,11}(X_1, X_2, t) + [k_1 U_{1,11}^0(X_1, X_2) + k_3 U_{1,22}^0(X_1, X_2) + k_5 U_{2,12}^0(X_1, X_2)] U_{1,1}(X_1, X_2, t)$$

$$[1 + k_1 U_{I,I}^0(X_1, X_2, t) + k_2 U_{J,J}^0(X_1, X_2, t)] U_{I,II}^0(X_1, X_2, t) + [2k_3 U_{I,J}^0(X_1, X_2, t) + 2k_4 U_{J,I}^0(X_1, X_2, t)] U_{I,IJ}^0(X_1, X_2, t) + [k_7 + k_3 U_{I,I}^0(X_1, X_2, t) + k_3 U_{J,J}^0(X_1, X_2, t)] U_{I,JJ}^0(X_1, X_2, t) + [k_4 U_{I,J}^0(X_1, X_2, t) + k_3 U_{J,I}^0(X_1, X_2, t)] U_{J,II}^0(X_1, X_2, t) + [k_3 U_{I,J}^0(X_1, X_2, t) + k_4 U_{J,I}^0(X_1, X_2, t)] U_{J,JI}^0(X_1, X_2, t) + [k_6 + k_5 U_{I,I}^0(X_1, X_2, t) + k_5 U_{J,J}^0(X_1, X_2, t)] U_{J,JI}^0(X_1, X_2, t) = 0. \quad (2)$$

Two equations result from Eq. (2): $I=1, J=2$ for the first equation and $I=2, J=1$ for the second equation.

The constants in Eqs. (1) and (2)

$$\begin{aligned} k_1 &= 3 + 6k(\nu_1 + \nu_2 + \nu_3), & k_2 &= k(\lambda + 6\nu_1 + 2\nu_2), \\ k_3 &= 1 + k(\nu_2 + \frac{3}{2}\nu_3), & k_4 &= k(\mu + \nu_2 + \frac{3}{2}\nu_3), \\ k_5 &= k[\lambda + \mu + 3(2\nu_1 + \nu_2 + \frac{1}{2}\nu_3)], & & (3) \\ k_6 &= k(\lambda + \mu), & k_7 &= k\mu, & k &= (\lambda + 2\mu)^{-1}, \\ c^{-2} &= \rho_0 k \end{aligned}$$

characterize the properties of the nonlinear elastic medium (ρ_0 denotes the density, λ and μ are Lamè constants and ν_1, ν_2 , and ν_3 are elastic constants of the third order).

B. Longitudinal wave propagation

To investigate the longitudinal wave propagation in an elastic medium undergoing inhomogeneous predeformation, we excite the longitudinal wave with an arbitrary smooth initial profile in the medium; i.e., we solve the equation-of-motion (1) under the following initial and boundary conditions:

$$\begin{aligned} U_{1,t}(X_1, X_2, 0) &= U(X_1, X_2, 0) = 0, \\ U_{1,t}(0, X_2, t) &= -\varepsilon a_0 \varphi(t) H(t). \end{aligned} \quad (4)$$

Here $H(t)$ denotes the Heaviside function and a_0 is a positive constant. The arbitrary and smooth initial wave profile is determined by the function $\varphi(t)$ under the condition $|\varphi(t)|_{\max} = 1$.

We use the perturbation theory and seek a solution to Eq. (1) by making the assumption that the displacements caused by the propagating wave may be described by the series

$$+ k_1 U_{1,11}(X_1, X_2, t) U_{1,1}(X_1, X_2, t) - c^{-2} U_{1,tt}(X_1, X_2, t) = 0. \quad (1)$$

Here $U_K(X_1, X_2, t)$ and $U_K^0(X_1, X_2)$ denote displacements caused by wave motion and predeformation, respectively. The indices after the comma indicate differentiation with respect to Lagrangian rectangular coordinates X_K , $K=1, 2$ or time t .

The coefficients in Eq. (1) are functions of displacements $U_K^0(X_1, X_2)$ that are determined as the solution to the system of equilibrium equations of the medium in the predeformed state:

$$U_1(X_1, X_2, t) = \sum_{n=1}^{\infty} \varepsilon^n U_1^{(n)}(X_1, X_2, t) \quad (5)$$

and the displacements of the predeformed state by the series

$$U_K^0(X_1, X_2) = \sum_{m=1}^{\infty} \epsilon^m U_K^{0(m)}(X_1, X_2), \quad (6)$$

where ε and ϵ are positive constants that satisfy the conditions $\epsilon \ll 1$ and $\varepsilon \ll 1$. The small parameters ε and ϵ may be of different order.

We introduce the series Eqs. (5) and (6) with small parameters ε and ϵ into Eqs. (1) and (2) and obtain a system of equations that enables one to describe the wave propagation in the elastic medium undergoing static inhomogeneous predeformation.

Prior to excitation of the longitudinal wave in the medium, the properties and the predeformed state of the medium are known. This means that we solve the equations deduced from Eq. (1) as one-dimensional hyperbolic equations with constant coefficients and with known right-hand sides. The coordinate X_2 may be regarded as a parameter.

Following the perturbation procedure presented in Ref. 3, we determine the first term in the series Eq. (5)

$$\begin{aligned} U_1^{(1)}(X_1, t) &= -a_0 \left[\int_0^\zeta \varphi(\tau) d\tau + \int_0^\eta \varphi(\tau) d\tau \right], \\ \zeta &= t - c^{-1} X_1, & \eta &= t + c^{-1} X_1. \end{aligned} \quad (7)$$

This is a well-known result. It describes the propagation of waves in a homogeneous isotropic elastic medium in the positive and negative directions of the axis X_1 . We consider the wave that propagates in the positive direction of axis X_1 and neglect the negative direction by neglecting the second term.

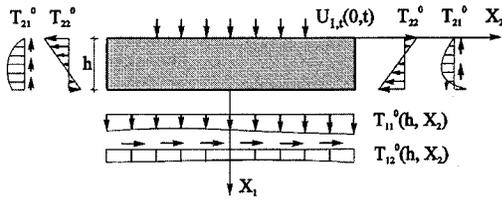


FIG. 1. Loading scheme of the medium.

In the series Eq. (5) the second and subsequent terms can be determined from corresponding equations under the initial and boundary conditions in the form of a convolution integral

$$U_1^{(n)}(X_1, X_2, t) = \sum_{j=1}^m \int_0^\xi F_j^{(n)}(\xi - \tau) P_j^{(n)}(\tau, X_1) d\tau. \quad (8)$$

This solution determines all terms in series Eq. (5) except the first. The function $F_j^{(n)}(\xi)$ depends on the initial wave profile $\varphi(t)$ and the function

$$P_j^{(n)}(\tau, X_1) = \lim_{Y \rightarrow \infty} \frac{1}{2\pi i} \int_{\alpha - iY}^{\alpha + iY} e^{\tau p} \int_0^{X_1} e^{2c^{-1}p\xi} \times \int_0^\xi e^{-2c^{-1}py} G_j^{(n)}(y, X_2) dy d\xi dp \quad (9)$$

depends on the predeformed state of the medium characterized by the function $G_j^{(n)}(y, X_2)$. Here p denotes the Laplace integral transform parameter.

As a result, we have the solution Eq. (5) to a quasi-one-dimensional problem Eq. (1) that describes the propagation of the longitudinal wave with an arbitrary smooth initial profile $\varphi(t)$ in the elastic medium undergoing inhomogeneous plane prestrain.

II. APPLICATION TO NDT

Our intention is to investigate the propagation of longitudinal waves in inhomogeneously predeformed elastic media and to discuss the possibility to utilize the results in nondestructive evaluation.

The following problem of nondestructive evolution is posed. Let the elastic medium (structural element) have finite dimensions (Fig. 1). Its thickness in Lagrangian rectangular coordinates X_K , $K=1,2$, bounded to one of its surfaces is equal to h in the direction of the axes X_1 and its length in the perpendicular direction is $2l$. On the planes $X_2 = \pm l$ it is supported. The external distributed static load acts on the surface $X_1 = h$. A longitudinal wave in this predeformed medium is excited on the surface $X_1 = 0$ and is recorded on the same surface. The aim is to evaluate the deformed state of the medium on the basis of recorded wave propagation data.

A. Polynomial description of the plane strain

The predeformed state of the medium corresponds to the plane strain. The equilibrium of the elastic medium in this predeformed state is described by the system of two Eqs. (2). We seek the solution to these equations in the form of series Eq. (6) and solve the corresponding equations in the special

case. We consider the equilibrium of the element of the medium (structural element) with the thickness h (Fig. 1). This element has the traction-free surface $X_1 = 0$ and it is subjected to the external load on the opposite surface $X_1 = h$. This load is defined by the the boundary conditions in terms of the components $T_{KL}^0(X_1, X_2)$ of the Kirchhoff pseudostress tensor

$$\begin{aligned} T_{11}^0(0, X_2) &= T_{12}^0(0, X_2) = 0, \\ T_{11}^0(h, X_2) &= \epsilon(p_0 + p_1 X_2 + p_2 X_2^2), \\ T_{12}^0(h, X_2) &= \epsilon(q_0 + q_1 X_2 + q_2 X_2^2), \end{aligned} \quad (10)$$

where p_n and q_n , $n=0,1,2$ are constants.

We assume that the solution Eq. (6) may be presented in the form of polynomials

$$U_K^0(X_1, X_2) = \epsilon P_{K1}^{6,6}(X_1, X_2) + \epsilon^2 P_{K2}^{11,11}(X_1, X_2). \quad (11)$$

Here $P_{KL}^{m,n}(X_1, X_2) = \sum_{p=0}^m \sum_{s=0}^n c_{ps} X_1^p X_2^s$, $K, L=1,2$ and c_{ps} denotes a constant.

By introducing the expression Eq. (11) into the equations of equilibrium Eq. (2), applying boundary conditions Eq. (10), and making use of the perturbation theory, the solution to the problem follows. Beforehand the boundary conditions Eq. (10) must be presented in terms of displacement vector components. The obtained solution is too cumbersome to include. We present, as an illustration, the expressions for the polynomials in solution Eq. (11) determined on the basis of the linear theory of elasticity under the most simple boundary conditions corresponding to the uniformly distributed external pressure and shear:

$$\begin{aligned} P_{11}^{6,6}(X_1, X_2) &= [(3\lambda + 2\mu)(2hX_1^3 - X_1^4)p_0 \\ &\quad + 6\lambda p_0 X_1^2 X_2^2 + (\lambda + 2\mu)p_0 X_2^4 \\ &\quad + X_1(2q_0 h^2 \lambda X_2 - 6h\lambda p_0 X_2^2)] / \\ &\quad (8h^3 \mu(\lambda + \mu)), \\ P_{21}^{6,6}(X_1, X_2) &= \{X_2^2[(\lambda + 2\mu) \\ &\quad \times (2hp_0 X_2 - 4p_0 X_1 X_2 - q_0 h^2)] \\ &\quad + X_1^2[(3\lambda + 4\mu)(q_0 h^2 + 4p_0 X_1 X_2 \\ &\quad - 6hp_0 X_2)]\} / (8h^3 \mu(\lambda + \mu)). \end{aligned} \quad (12)$$

The 156 coefficients of polynomials $P_{K2}^{11,11}(X_1, X_2)$ in solution Eq. (11) are determined from the corresponding system of linear differential equations with known r.h.s. under appropriate boundary conditions.

Now, on the basis of solution Eq. (11) and the nonlinear theory of elasticity⁹ one can calculate the stresses in the medium. Some results of calculations are presented in Figs. 2–5.

B. Sine-wave in predeformed medium

It is convenient to investigate the wave profile distortion in the predeformed medium on the basis of sine-wave propagation data and to define the initial profile of the wave by the function

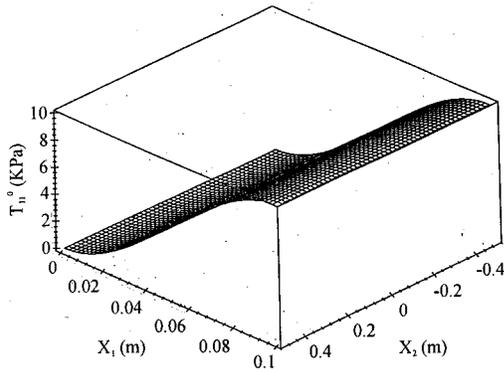


FIG. 2. Stress T_{11}^0 in the medium subjected to the uniform normal load.

$$\varphi(t) = \sin \omega t, \quad (13)$$

where ω denotes the frequency. This makes it possible to compare our results with well-known results in special cases.

Introducing function Eq. (13) into the boundary conditions Eq. (4) and making use of the procedure described above, the solution Eq. (5) with accuracy of three leading terms that describe the longitudinal sine-wave propagation in a nonlinear elastic medium undergoing inhomogeneous plane strain follows in the expected form:

$$U_{1,i}(X_1, X_2, t) = A_0 + A_1 \sin(\omega \zeta + \theta_1) + A_2 \sin(2\omega \zeta + \theta_2) + A_3 \sin(3\omega \zeta + \theta_3). \quad (14)$$

Here A_0 is the nonperiodic term, A_J and θ_J , $J=1,2,3$ denote the amplitudes and phase shifts of harmonics, respectively. The analytical expressions for functions A_0 , A_J , and θ_J are too cumbersome to be reproduced here.

It should be noted that in the special case of nonlinear sine-wave propagation in the prestress free elastic medium the solutions Eq. (14) coincide with the results of other authors.¹⁰

The basis effects responsible for the distortion of the wave profile are inhomogeneity (inhomogeneous predeformation), dispersion, and nonlinearity. We intend to use nonlinear effects of wave propagation in nondestructive evaluation of inhomogeneous properties of the medium. It is important that the recorded nonlinear effects contain maximum information about the inhomogeneity. The amount of this information is dependent on the ratio of strain intensities

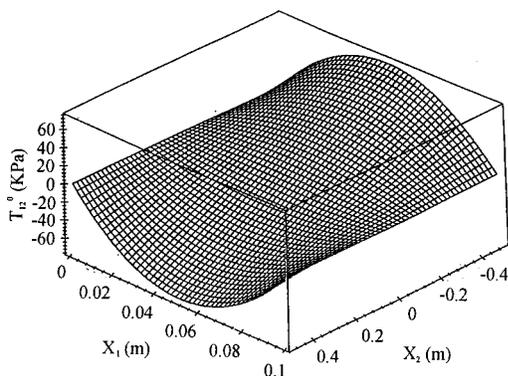


FIG. 3. Stress T_{12}^0 in the medium under uniform normal load.

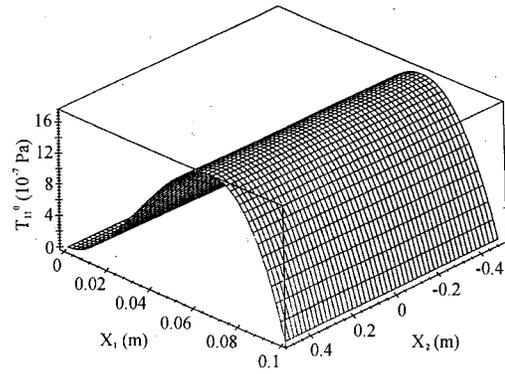


FIG. 4. Stress T_{11}^0 induced in the medium by the uniform shear load.

in the medium caused by the propagating wave and the predeformation, correspondingly. With the view to study this problem in more detail, we propose that $|U_I^{(n)}(X_1, X_2, t)| \approx |U_K^{0(n)}(X_1, X_2)|$ in Eqs. (5) and (6) and distinguish three different cases of wave motion:

Case A. The displacements of the material points of the medium caused by wave motion are much smaller than the corresponding displacements caused by predeformation, i.e., $\varepsilon \approx \varepsilon^2$ in series Eqs. (5) and (6).

Case B. The displacements caused by wave motion and predeformation are of the same order, i.e., $\varepsilon \approx \varepsilon$.

Case C. The displacements caused by wave motion are much larger than the displacements caused by predeformation, i.e., $\varepsilon^2 \approx \varepsilon$.

In all three cases we consider three first terms in solution Eq. (5). It is necessary to point out that the solution Eq. (5) consists of the following terms. The first term Eq. (7), of the solution describes the propagation of a longitudinal wave in a homogeneous isotropic elastic medium. The second and the third terms Eq. (8), correct the solution taking into account the effects caused by inhomogeneity (inhomogeneous predeformation) and nonlinearity.

Let us consider case A: The displacements caused by wave motion are much smaller than the displacements caused by predeformation. The first term Eq. (7) of solution Eq. (5) determines the amplitude of the first harmonic in solution Eq. (14). The second term corrects the solution for inhomogeneity (predeformation). This term describes the phase shift of the first harmonic also. The third term corrects the first harmonic once more and describes the evolution of

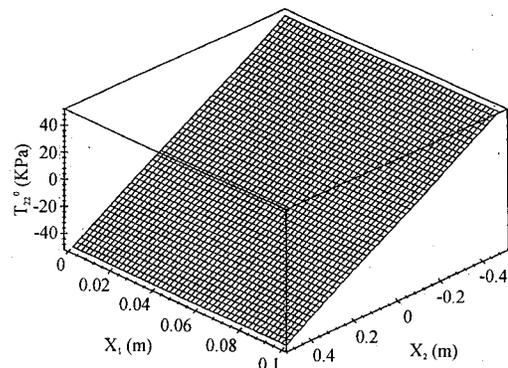


FIG. 5. Stress T_{22}^0 caused in the medium by the uniform shear load.

the second harmonic in a homogeneous medium. In this solution nonlinear effects are small and in the range of accuracy considered here they are not affected by inhomogeneity (predeformation). It is impossible to use this solution to correct for nonlinear effects in nondestructive evaluation of predeformed media.

If the displacements caused by wave motion are of the same order as the displacements caused by predeformation (case B), then the nonlinear and inhomogeneous effects are described simultaneously by the second term in solution Eq. (5). On this level the nonlinear effects are not dependent on the inhomogeneity of the medium. Interaction of nonlinearity and inhomogeneity is described by the third term in solution Eq. (5). This term defines the influence of inhomogeneity on the evolution of the second harmonic in solution Eq. (14) and describes the evolution of the third harmonic in a homogeneous medium.

In case C the large displacements caused by wave propagation are superposed on small displacements of the predeformed state. The wave motion in the medium is characterized by strong nonlinear effects. The second and the third terms in solution Eq. (5) describe the evolution of the second and the third harmonics in homogeneous elastic media, respectively. In this case the influence of inhomogeneity on the wave process is a small effect. The third term defines the influence of inhomogeneity on the first harmonic only. It does not correct the description of the second harmonic. This means that in case C the first three terms of solution Eq. (5) do not correct for nonlinear effects of wave propagation. Nonlinear effects are not affected by inhomogeneity.

We can conclude that in ultrasonic NDE of inhomogeneous materials it is important to pay attention to the intensity of the applied wave. The best correction for nonlinear effects of wave propagation may be achieved by choosing the wave intensity such that deformations of the medium caused by the wave propagation are of the same order as deformations of the predeformed state.

C. Acoustodiagnostics of the plane strain

Solution Eq. (14) describes sine-wave propagation in the medium undergoing inhomogeneous plane prestrain. The wave propagation is characterized by the frequency ω , by the nonperiodic term A_0 and by the amplitudes and phase shifts of harmonics $A_1, A_2, A_3, \theta_1, \theta_2$, and θ_3 . The amount of information in these functions depends on the ratio of deformations caused by the wave motion and by the predeformation. As we indicated these functions contain maximum information in case B; i.e., if this ratio is near unity. We consider case B henceforth. In this case the amplitudes and the phase shifts of the first two harmonics are dependent on the physical properties of the medium and on the parameters of the predeformed state. The amplitude and the phase shift of the third harmonic A_3 and θ_3 are not dependent on predeformation. The corrections appear in the next higher approximation. We neglect the third harmonic in our discussions.

The phase and group velocities of the harmonics c_p^n and c_{gr} may be calculated by the formulas^{11,12}

$$c_p^n = nc\omega / (n\omega - c\theta_{n,1}), \quad (15)$$

$$c_{gr} \equiv (\omega_{n+1} - \omega_n) / (g_{n+1} - g_n), \quad (16)$$

where g_n is the wave number of the n -th harmonic.

The unknown quantities to be determined in NDT are (i) the physical properties of the nonlinear elastic medium, (ii) the dimensions of this medium, and (iii) the parameters of the predeformed state. The medium is characterized by the elastic constants of the second and third order $\lambda, \mu, \nu_1, \nu_2, \nu_3$, by the density ρ_0 and by the dimensions h and l . The predeformed state (plane strain) is determined by the coefficients c_{ps} in polynomials $P_{KJ}^{m,n}(X_1, X_2)$ in solution Eq. (11). These numerous coefficients c_{ps} are functions of the geometrical and physical properties of the medium and the external load. The latter is determined by the boundary conditions Eq. (10). It becomes clear that it is convenient to characterize the predeformed state of the medium by the constants p_j and q_j ($j=0,1,2$) instead of the coefficients c_{ps} .

Consequently, if we try to use the solution Eq. (14) in nondestructive evaluation of 14 constants that characterize the medium, at our disposal there are four basic functions $A_1, A_2, \theta_1, \theta_2$, one constant ω and three deduced functions c_p^1, c_p^2 , and c_{gr} . We can conclude that it is not possible to determine the properties and the state of the medium on the basis of these functions only. That is why we propose to use the solution Eq. (14) provided preliminary information about the properties and the state of the medium is available.

We illustrate this approach following the idea presented in Ref. 8. The wave process is excited on the traction free surface $X_1=0$ of the medium undergoing plane inhomogeneous prestrain (Fig. 1) in accordance with the boundary condition Eq. (4) i.e., in terms of the particle velocity. The evolution of the wave is recorded on the same surface, but in terms of the stress. The nonlinear theory of elasticity⁹ describes this stress as a function of the derivative of the particle displacement with respect to the spatial coordinates. In our quasi-one-dimensional case this stress is a function of $U_{1,1}(X_1, X_2, t)$ in the form

$$U_{1,1}(X_1, X_2, t) = A_0 + A_1 \sin(\omega\zeta + \theta_1) + A_2 \sin(2\omega\zeta + \theta_2). \quad (17)$$

Here it is convenient to keep the notation introduced for the amplitudes and phase shifts in Eq. (14). Since these amplitudes A_0, A_1, A_2 and the phase shifts θ_1 and θ_2 are functions of spatial coordinates and time, they are not expressed exactly by the same expressions in Eqs. (14) and (17). But, in principle, the amplitudes and the phase shifts in Eq. (17) have the same properties as the corresponding functions in Eq. (14). They are functions of the properties of the medium and the predeformed state. They are affected by the ratio of the magnitudes of displacements caused by the predeformation and the wave motion in the same way.

Henceforth, we clarify the dependence of wave characteristics on the parameters of the predeformed state of the medium. We deduce the analytical expressions for the amplitudes A_0, A_1, A_2 and the phase shifts θ_1 and θ_2 in Eq. (17). These nonlinear expressions turned out to be too cumbersome for the analytical analyses. Instead, we pose the

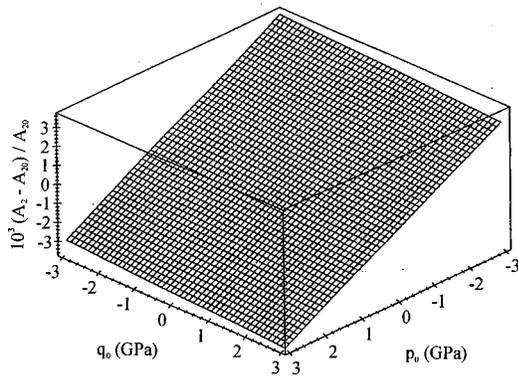


FIG. 6. Relative amplitude of the second harmonic versus uniform normal and shear loads.

numerical experiment and consider one possible special case of longitudinal wave propagation in an elastic medium undergoing plane prestrain. It is noteworthy that the results obtained characterize qualitatively the dependence of the wave characteristics on the parameters of the considered predeformed state for nonlinear elastic media of wide range of physical properties and for the different values of the wave frequency.

Following our scheme of nondestructive testing, we propose that the properties of the medium are known. They correspond to aluminum with density $\rho_0 = 2800 \text{ kg/m}^3$, constants of elasticity $\lambda = 50 \text{ GPa}$, $\mu = 27.6 \text{ GPa}$, $\nu_1 = -136 \text{ GPa}$, $\nu_2 = -197 \text{ GPa}$, $\nu_3 = -38 \text{ GPa}$, and with the dimensions $h = 0.1 \text{ m}$, $l = 1 \text{ m}$. The purpose is to evaluate the parameters of the predeformed state of the medium p_j , q_j , with $j = 0, 1, 2$.

The predeformed state of the medium is described by the solution Eq. (11). We illustrate the nature of this state by the distribution of stress in the medium. Under the applied load [see Eq. (10)] it is strongly nonlinear in space. In the more simple case when the external load corresponds to a uniform normal load equal to 10 KPa the distribution of stresses $T_{11}^0(X_1, X_2)$ and $T_{22}^0(X_1, X_2)$ is plotted in Figs. 2 and 3. The distribution of the stress $T_{11}^0(X_1, X_2)$ (Fig. 2) coincides qualitatively with the results presented in Ref. 13.

In the second case we apply a pure shear equal to 10 KPa on the surface $X_1 = h$. Once more the distribution of the stress $T_{11}^0(X_1, X_2)$ is strongly nonlinear (Fig. 4) into the depth of the medium but the distribution of the stress

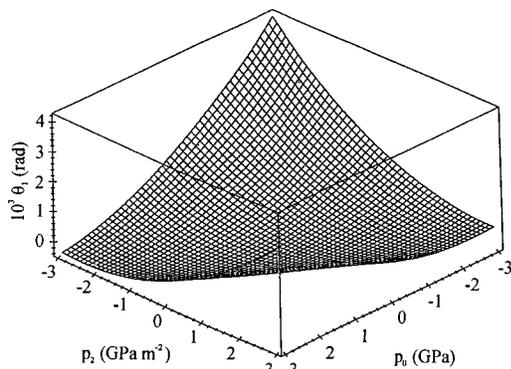


FIG. 7. Phase shift of the first harmonic versus uniformly and quadratically distributed normal loads.

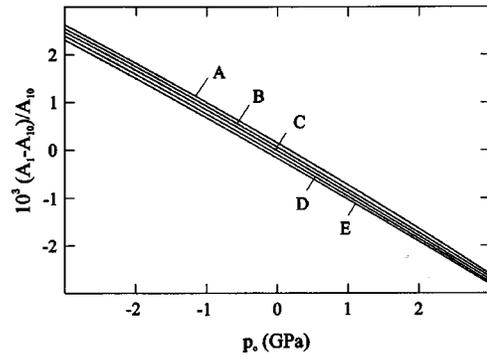


FIG. 8. Variation of the relative amplitude of the first harmonic in the medium subjected to normal and shear loads. Curves A, B, ..., E correspond to the values of the load q_0 equal to $-6, -3, \dots, 6 \text{ GPa}$, respectively.

$T_{22}^0(X_1, X_2)$ may be considered as linear (Fig. 5). A longitudinal sine-wave with the frequency $\omega = 10^6 \text{ rad s}^{-1}$ and the parameter $\varepsilon = 3 \times 10^{-4}$ is excited in the medium. Its evolution can be analyzed on the basis of Eq. (17). As mentioned above, the derivation of analytical expressions for the parameters p_j , q_j , $j = 0, 1, 2$ as functions of wave characteristics is too complicated in the nonlinear case and therefore we determine the dependence of the amplitudes and phase shifts of wave harmonics on the parameters p_j and q_j numerically.

Computational simulation verifies the fact that the amplitude of the second harmonic and the dependence of phase shift of the second harmonic on the parameters p_j , q_j are very close to linear (Fig. 6). An exception is the dependence of the phase shift of the first harmonic (Fig. 7) and the phase velocity of the first harmonic on the parameters p_0 and p_2 (the functions A_{10} , A_{20} , θ_{10} , and θ_{20} denote here the amplitudes and phase shifts of two first harmonics of the wave in the medium without predeformation).

It is essential from the point of view of nondestructive testing that the sensitivity of wave characteristics to the parameters p_j and q_j is different. For example, the variation of the amplitude of the second harmonic is strongly sensitive to the uniform normal load characterized by the parameter p_0 and less sensitive to the uniform shear load q_0 (Fig. 6). The sensitivity of the amplitude of the first harmonic to these parameters is similar (Fig. 8). The phase shifts of both harmonics are strongly sensitive to both parameters p_0 and q_0 (Fig. 9). These different sensitivities permit us to propose the

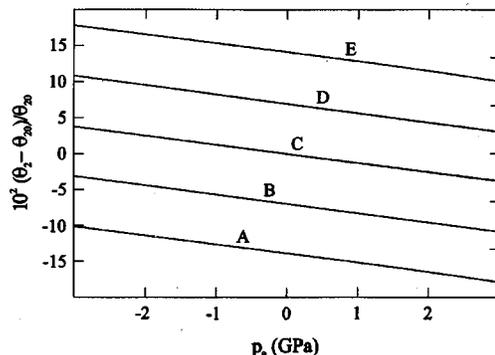


FIG. 9. Relative phase shift of the second harmonic versus uniform normal and shear loads. Curves A, B, ..., E correspond to the values of the load q_0 equal to $-6, -3, \dots, 6 \text{ GPa}$, respectively.

use of the algorithm to acoustodiagnostics for determining the predeformed state of the medium.

There are many possibilities to use the results presented above in nondestructive determination of the parameters of the predeformed state of the elastic medium. Let us confine our attention to the following case, as an illustration. The preliminary inspection of the loading scheme of the medium (structural element) permits us to propose that the medium is undergoing plane strain under uniformly distributed external load. The problem is to evaluate the predeformed state, i.e., the values of the parameters p_0 and q_0 . We determine the physical and geometrical parameters of the medium. Taking into consideration the results of computation presented above we decide to solve the problem on the basis of the first harmonic amplitude and the second harmonic phase shift measurement data. We plot the curves similar to Fig. 8 and Fig. 9 for the real medium. Then making use of the wave propagation measurement data and resorting to the curves in the analogue of Fig. 8, we determine the parameter p_0 . After that the measured value of the phase shift of the second harmonic and the known value of the parameter p_0 enables us to determine the parameter q_0 from the curves in Fig. 9.

III. DISCUSSION AND CONCLUSIONS

The topic of this paper may be regarded as the development of the ideas presented in Ref. 3. The extension of the scope of nondestructive testing to the range of evaluation of inhomogeneous predeformation brings the necessity to increase the amount of information available from the recorded data of wave motion. It is clear that it is difficult to identify even the simple inhomogeneous predeformed states of an elastic medium on the basis of wave velocity measurement data only.³ We are in need of additional information. The development of the measurement technique encourages us to propose the wave profile distortion data as the source of additional information in nondestructive evaluation of inhomogeneous predeformation fields.

The distortion of the wave profile in an inhomogeneously predeformed elastic medium is caused primarily by the effects of inhomogeneity and nonlinearity. We tried to clarify the interaction of these effects in wave motion excited in the medium with various magnitudes of predeformation. These results facilitate the understanding of the nature of the wave profile distortion and the formulation of the nondestructive evaluation problem.

We described the nonlinear propagation of the longitudinal wave with an arbitrary smooth initial profile in the medium with variable properties by the solution Eq. (5) determined on the basis of Eqs. (7)–(9). In principle, this solution may be used to describe the wave motion in media with a different nature of inhomogeneity.

In this paper we confined our attention to wave propagation in an elastic medium undergoing inhomogeneous plane strain. Making use of the software for algebraic computations under the UNIX operating system we found the exact polynomial asymptotic solution Eq. (11) to the systems of nonlinear equations of equilibrium Eq. (2) of the elastic

medium with nonlinear physical properties subjected to a distributed load on one of its surfaces (Fig. 1). The load is described by the polynomials up to the second order Eq. (10) and may be decomposed into the normal and shear components. We draw a conclusion that provides information about the character of the predeformed state (plane strain) and the physical and geometrical properties of the medium.

The sine-wave propagation in this predeformed medium is investigated in detail. The difference between the weak and the strong waves gives a profile of the evolution in the predeformed medium. Expressions are derived for wave characteristics versus physical and geometrical properties of the medium and the parameters of the predeformed state. On the basis of expressions not presented in this paper because of cumbersomeness are derived the wave characteristics versus the predeformation parameters and plots are computed. The analysis of these plots verifies the assumption that the extraction of information from the wave profile evolution data enables one to enhance the efficiency of nondestructive testing and to solve problems of nondestructive evaluation of inhomogeneous predeformed states of an elastic medium.

ACKNOWLEDGMENTS

The work related to this paper was mainly carried out when the author was at the School of Engineering of the Uppsala University; the assistance provided by the University is acknowledged. The author wishes to express his sincere thanks to Professor Bengt Lundberg for stimulating discussions, important suggestions and for support to this work. This work was also supported by grants from The Royal Swedish Academy of Sciences and from The Estonian Science Foundation (No. 1908).

¹M. Kato, T. Sato, K. Kameyama, and H. Ninoyu, "Estimation of the stress distribution in metal using nonlinear acoustoelasticity," *J. Acoust. Soc. Am.* **98**, 1496–1504 (1995).

²T. Takuoka, "Weak and short waves in one-dimensional inhomogeneous nonlinear elastic materials," *J. Acoust. Soc. Am.* **69**, 66–69 (1981).

³A. Ravasoo, "Ultrasonic nondestructive evaluation of inhomogeneous plane strain in elastic medium," *Res. Nondestruct. Eval.* **7**, 55–68 (1995).

⁴Y.-H. Pao, W. Sachse, and H. Fukuoka, "Acoustoelasticity and ultrasonic measurements of residual stresses," in *Physical Acoustics*, edited by W. P. Mason and R. N. Thurston (Academic, Orlando, 1984), Vol. 17, pp. 61–143.

⁵R. Truell, C. Elbaum, and B. B. Chick, *Ultrasonic Methods in Solid State Physics* (Academic, New York, 1969).

⁶M. C. Bhardwaj, "Fundamental developments in ultrasonics for advanced NDC," in *Nondestructive Testing of High-Performance Ceramics, Conference Proceedings* (American Ceramic Society, Westerville, OH, 1987), pp. 472–527.

⁷S. P. Timoshenko and J. N. Goodier, *Theory of Elasticity* (McGraw-Hill, New York, 1982), pp. 35–53.

⁸B. Lundberg, "On the separation of overlapping elastic waves from measured data," in *Application of Stress-Wave Theory on Piles, Second International Conference*, edited by G. Holm, H. Bredenberg, and C.-J. Grävare (Balkema, Stockholm, 1984), pp. 441–442.

⁹A. C. Eringen, *Nonlinear Theory of Continuous Media* (McGraw-Hill, New York, 1962).

¹⁰U. K. Nigul, *Echo-Signals from Elastic Objects* (Valgus, Tallinn, 1978), Vol. 1, pp. 207–208 (in Russian).

¹¹G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).

¹²K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1975).

¹³M. Filonenko-Borodich, *Theory of Elasticity* (Foreign Languages Publishing House, Moscow, 1980), p. 167.

Multifrequency plane, nonlinear, and dissipative waves at arbitrary distances

Claes M. Hedberg^{a)}

Department of Mechanical Engineering, University of Karlskrona—Ronneby, 371 79 Karlskrona, Sweden

(Received 6 November 1998; accepted for publication 16 August 1999)

A solution for multifrequency plane waves propagating through a dissipative and nonlinear medium is presented. It originates from the well-known Bessel function series ratio for a pure sinusoidal wave, introduced by Cole and Mendousse. The solution is exact. The only limitation, inherited from the single-frequency solution, is the slow convergence of the series when the nonlinearity is very large compared to the dissipation. Otherwise any frequencies, amplitudes and phases can be introduced in the original wave and the solution is valid for any propagated distance. © 1999 Acoustical Society of America. [S0001-4966(99)02412-1]

PACS numbers: 43.25.Lj, 43.25.Cb [MFH]

LIST OF SYMBOLS

$A_n^{(l)}$	$= I_n(\alpha_l/2\alpha_l\epsilon^*)(-1)^n \exp(in\gamma_l)$	V_{mfq}	the subscript mfq stresses that the solution is for multifrequency boundary condition, as opposed to solutions for single frequencies
a	amplitude of single-frequency boundary condition	V^*	$= v/v_{00}$, a specific choice of dimensionless velocity amplitude
a_l	amplitude of frequency α_l in boundary condition	v	velocity amplitude
b	the effect of viscosity and heat conduction	v_0	characteristic velocity of the medium
b_l	amplitude coefficient in boundary condition	α	integer angular frequency
c_0	the undisturbed sound velocity in the medium	α_l	integer angular frequency belonging to the frequency denoted by l in the boundary condition
c_k	frequency-dependent coefficient in U and V	β	$= (\gamma + 1)/2$, the nonlinearity of the medium
I_n	the modified Bessel function	γ	phase in boundary condition
k	the frequency parameter in the sum (or integral) that is the solution U	γ_l	phase in boundary condition belonging to frequency α_l
n	summation index	ϵ	$= b\omega/2\beta c_0 v_0 \rho_0$, a dimensionless ratio between dissipation and nonlinearity
$n^{(l)}$	summation variable belonging to frequency α_l	ϵ^*	$= b\omega_0/2\beta c_0 v_{00} \rho_0$, a specific choice of ϵ
x	propagated distance	θ	$= \omega\tau$, dimensionless time
S	$= S(\sigma)$, the factor dependent on distance σ in the separation of $U(\sigma, \theta)$	θ^*	$= \omega_0\tau$, specific choice of dimensionless time
T	$= T(\theta)$, the factor dependent on retarded time θ in the separation of $U(\sigma, \theta)$	ρ_0	undisturbed density of the medium
U	transformation variable defined by $V = 2\epsilon\partial/\partial\theta(\ln U)$	σ	$= \beta\omega v_0 x/c_0^2$, dimensionless propagation distance
U_{mfq}	the subscript mfq stresses that the solution is for multifrequency boundary condition, as opposed to solutions for single frequencies.	σ^*	$= \beta\omega_0 v_{00} x/c_0^2$, specific choice of dimensionless propagation distance
V	$= v/v_0$, dimensionless velocity amplitude	τ	$= t - x/c_0$, retarded time
V_l	the subscript l is the solution for the frequency denoted by l with frequency α_l	ω	characteristic angular frequency of the signal
		$1/\omega$	a characteristic time

INTRODUCTION

Nonlinear evolution of plane waves through a dissipative medium is described by the Burgers equation. For a single-frequency source solutions have been known for a long time. There are, as the best known examples, the Fubini solution¹ valid for propagation up to shock formation and the Fay solution,² whose validity is starting at approximately four shock distances. Both of these are Fourier series expres-

sions and neither one is exact. Both may be derived³ from the solution presented by Cole⁴ and Mendousse,⁵ which is a ratio between two Fourier series with Bessel function coefficients. This solution has one disadvantage in that the convergence of the series is poor⁶ and the solution does not work well when the nonlinearity is very large compared to the dissipation.

The propagation in a nonlinear and dissipative medium is a contest between the nonlinearity and the dissipation. Nonlinearity makes the positive amplitudes go faster than the equilibrium and the negative go slower. This results in the

^{a)}Electronic mail: claes.hedberg@ima.hk-r.se

creation of combination frequencies of the original frequencies and their higher harmonics. Thus, most often also lower frequencies appear. The higher harmonics may turn the wave into a shocked state.

In contrast, dissipation is attenuating the higher harmonics much more than lower, thus making the waves go into shock less easily.

For original signals consisting of several frequencies there has been little more than numerics to use with a few exceptions for the case of no dissipation. There is one originating from Burgers for zero dissipation,⁷ where the total wave solution can be obtained for all distances through the finding of the maximum of a function. The most distinguished result is the analytic solution by Fenlon⁸ for zero dissipation, valid up to the formation of shock, which is a generalization of the Fubini solution. The solution to the Burgers equation in the form of a ratio between two infinite integrals yields a complete solution for the given distance. The integration interval naturally has to be much smaller and the integrand ought to be examined beforehand in order to reach accurate results. In this paper a solution for multi-frequency conditions for dissipative and nonlinear propagation is introduced. It originates from the single-frequency Cole–Mendousse solution and it is a generalization of the same. Therefore the derivation of the multi-frequency solution begins with the well-known derivation of the single-frequency solution.

I. THE SOLUTION FOR ONE FREQUENCY

The nonlinear equation for plane waves in a homogeneous dissipative medium is Burgers' equation, given in dimensionless variables as

$$\frac{\partial V}{\partial \sigma} - V \frac{\partial V}{\partial \theta} - \epsilon \frac{\partial^2 V}{\partial \theta^2} = 0. \quad (1)$$

The definition of the dimensionless variables—using a characteristic velocity v_0 of the medium, b the effect of viscosity and heat conduction, $\tau = t - x/c_0$ the retarded time, ρ_0 the undisturbed density, c_0 the undisturbed sound velocity, $\beta = (\gamma + 1)/2$ the nonlinearity for a fluid and a characteristic time $1/\omega$ —are

$$V = \frac{v}{v_0}, \quad (2)$$

$$\theta = \omega \tau, \quad (3)$$

$$\sigma = \frac{\beta}{c_0^2} \omega v_0 x, \quad (4)$$

$$\epsilon = \frac{1}{2\beta} \frac{b\omega}{c_0 v_0 \rho_0}. \quad (5)$$

If the dimensionless ratio between dissipation and nonlinearity ϵ is less than 1, nonlinear theory should be used. Equation (1) can be transformed into a linear equation through

$$V = 2\epsilon \frac{\partial U / \partial \theta}{U} = 2\epsilon \frac{\partial}{\partial \theta} (\ln U), \quad (6)$$

and the heat conduction equation is obtained,

$$\frac{\partial U}{\partial \sigma} = \epsilon \frac{\partial^2 U}{\partial \theta^2}. \quad (7)$$

This transformation was found by Hopf⁹ and Cole.⁴

The solution can be found by separation of the variables:

$$U(\sigma, \theta) = S(\sigma)T(\theta). \quad (8)$$

The general solution of (7) can then be written as a sum,

$$U = \sum_{k=-\infty}^{\infty} c_k e^{-\epsilon k^2 \sigma} e^{ik\theta}, \quad (9)$$

or as an integral,

$$U(\sigma, \theta) = \int_{-\infty}^{\infty} c(k) e^{ik\theta - \epsilon k^2 \sigma} dk. \quad (10)$$

Before turning to the Fourier series form a short discussion on the integral form will be conducted. It is a general solution, but in practice there is numerical trouble encountered when solving this integral because the exponent in the infinite integral will vary greatly. As an example, let the boundary condition be a simple sine wave:

$$V(\sigma = 0, \theta) = \sin \theta. \quad (11)$$

This results through (6) and (1)¹⁰ in the solution

$$V = 2 \frac{\int_{-\infty}^{\infty} \frac{\theta' - \theta}{\sigma} \exp\left[\frac{1}{2\epsilon}(-\cos \theta') - \frac{(\theta - \theta')^2}{4\epsilon\sigma}\right] d\theta'}{\int_{-\infty}^{\infty} \exp\left[\frac{1}{2\epsilon}(-\cos \theta') - \frac{(\theta - \theta')^2}{4\epsilon\sigma}\right] d\theta'}. \quad (12)$$

This is a simple solution to Burgers' equation. The limits in the integration has to be decreased considerably as the numbers may become very large, in particular, for small ϵ . The way to proceed is to examine the exponents beforehand to be able to choose the relevant interval for the integration. The integration interval may often have to be divided into several smaller intervals. This procedure leads naturally to the saddle point method concept which only takes into account the major parts of the integrand around certain maxima. Through the saddle point method both the Fubini and Fay solutions may be derived. The Fubini through use of one saddle point¹¹ and the Fay through two saddle points yielding the time domain Khokhlov solution,¹² which is then converted into a Fourier series.¹⁰ For the simple sine wave there is only one maximum per period. For more complicated input the analytical saddle point method leads to solutions valid only locally,¹³ except for nondissipative continuous waves before the formation of shocks.

The Fourier series representation (9) is appropriate when the boundary condition is periodical. For the simple harmonic signal as a boundary condition (11) the solution to Eq. (1) was obtained by Cole⁴ and Mendousse.⁵

$$V(\sigma, \theta) = -4\epsilon \frac{\sum_{n=1}^{\infty} \exp(-n^2 \epsilon \sigma) n (-1)^n I_n(1/2 \epsilon) \sin n \theta}{I_0(1/2 \epsilon) + 2 \sum_{n=1}^{\infty} \exp(-n^2 \epsilon \sigma) (-1)^n I_n(1/2 \epsilon) \cos n \theta}. \quad (13)$$

Here, I_n denotes modified Bessel functions.

Of special interest might be to replace the quotient between two Fourier series (13) by an analytic single Fourier series. This has been done by Enflo³ through a recursion algorithm

II. SOLUTION FOR SEVERAL FREQUENCIES

Any solution for a periodic boundary condition may have the form of the ansatz (9):

$$U_{mfq} = \sum_{k=-\infty}^{\infty} c_k e^{-k^2 \epsilon \sigma} e^{ik\theta}, \quad (14)$$

where the subscript *mfq* stands for *multifrequency*. This is the equivalent to the solution

$$V_{mfq}(\sigma, \theta) = i2\epsilon \frac{\sum_{k=-\infty}^{\infty} k c_k e^{-k^2 \epsilon \sigma} e^{ik\theta}}{\sum_{k=-\infty}^{\infty} c_k e^{-k^2 \epsilon \sigma} e^{ik\theta}}. \quad (15)$$

Let (15) be the solution to a boundary condition consisting of a number of L frequencies, the α_l 's are integers:

$$V_{mfq}(\sigma=0, \theta) = \sum_{l=-L}^L b_l e^{i\alpha_l \theta}. \quad (16)$$

Each one of the L frequencies has on their own a known solution V_l in a form similar to (13).

Because this is a nonlinear problem they cannot be superposed to form the solution to the multiple frequency boundary condition—except at the distance $\sigma=0$. At this distance no nonlinear (nor linear) evolution has taken place, and superposition is allowed. So this is where, in the remainder of this section, the coefficients of the multifrequency wave will be extracted from the known coefficients of the single frequencies.

Some care must be taken to make sure the individual V 's are expressed in the same dimensionless variables because, if

the dimensionless variables of each individual frequency are inserted straight into (13), these dimensionless variables will be defined for each particular frequency. They will then not stand for the same real variables. If they are to be combined they must be expressed by some definite variables—the same for all individual frequencies. Let the choices be [compare (2)–(5)]

$$V^* = \frac{V}{V_{00}}, \quad (17)$$

$$\theta^* = \omega_0 \tau, \quad (18)$$

$$\sigma^* = \frac{\beta}{c_0^2} \omega_0 V_{00} x, \quad (19)$$

$$\epsilon^* = \frac{1}{2\beta} \frac{b \omega_0}{c_0 V_{00} \rho_0}. \quad (20)$$

All single-frequency boundary conditions written like

$$V^*(\sigma^*=0, \theta^*) = \alpha \sin(\alpha \theta^* + \gamma) \\ = \frac{a}{2} [\exp(i(\alpha \theta^* + \gamma)) - \exp(-i(\alpha \theta^* + \gamma))], \quad (21)$$

will get the results in the same dimensionless parameters by making the following replacements in (13),

$$V = \frac{V^*}{a}, \quad (22)$$

$$\theta = \alpha \theta^* + \gamma, \quad (23)$$

$$\sigma = a \alpha \sigma^*, \quad (24)$$

$$\epsilon = \frac{\alpha}{a} \epsilon^*. \quad (25)$$

So the solutions to the individual frequencies will be

$$V^* = a V(\sigma^*, \theta^*) = -4\alpha \epsilon^* \frac{\sum_{n=1}^{\infty} \exp(-n^2 \alpha^2 \epsilon^* \sigma^*) n (-1)^n I_n\left(\frac{a}{2\alpha \epsilon^*}\right) \sin n(\alpha \theta^* + \gamma)}{I_0\left(\frac{a}{2\alpha \epsilon^*}\right) + 2 \sum_{n=1}^{\infty} \exp(-n^2 \alpha^2 \epsilon^* \sigma^*) (-1)^n I_n\left(\frac{a}{2\alpha \epsilon^*}\right) \cos n(\alpha \theta^* + \gamma)}. \quad (26)$$

Now we may write the superposition at zero propagated distance (15),

$$V_{mfq}^*(\sigma^*=0, \theta^*) = i2\epsilon^* \frac{\sum_{k=-\infty}^{\infty} k c_k e^{ik\theta^*}}{\sum_{k=-\infty}^{\infty} c_k e^{ik\theta^*}}, \quad (27)$$

$$= \sum_{l=1}^L V_l^*(\sigma^*=0, \theta^*) = \{\text{with (26)}\}, \quad (28)$$

$$= \sum_{l=1}^L \frac{\sum_{n=1}^{\infty} -2\alpha_l \epsilon^* n (-1)^n I_n \left(\frac{a_l}{2\alpha_l \epsilon^*} \right) [\exp(in(\alpha_l \theta^* + \gamma_l)) - \exp(-in(\alpha_l \theta^* + \gamma_l))]}{\sum_{n=1}^{\infty} (-1)^n I_n \left(\frac{a_l}{2\alpha_l \epsilon^*} \right) [\exp(in(\alpha_l \theta^* + \gamma_l)) + \exp(-in(\alpha_l \theta^* + \gamma_l))]}, \quad (29)$$

$$= \sum_{l=1}^L -2\alpha_l \epsilon^* \frac{\sum_{n=-\infty}^{\infty} n (-1)^n I_n \left(\frac{a_l}{2\alpha_l \epsilon^*} \right) \exp(in(\alpha_l \theta^* + \gamma_l))}{\sum_{n=-\infty}^{\infty} (-1)^n I_n \left(\frac{a_l}{2\alpha_l \epsilon^*} \right) \exp(in(\alpha_l \theta^* + \gamma_l))}, \quad (30)$$

$$= \sum_{l=1}^L -2\alpha_l \epsilon^* \frac{\sum_{n=-\infty}^{\infty} n A_n^{(l)} \exp(in\alpha_l \theta^*)}{\sum_{n=-\infty}^{\infty} A_n^{(l)} \exp(in\alpha_l \theta^*)}, \quad (31)$$

with

$$A_n^{(l)} = I_n \left(\frac{a_l}{2\alpha_l \epsilon^*} \right) (-1)^n \exp(in\gamma_l). \quad (32)$$

The relation between A_n and A_{-n} is

$$A_{-n}^{(l)} = A_n^{(l)} \exp(-i2n\gamma_l). \quad (33)$$

The next step is to identify (27) with (31) in each frequency, with $n^{(l)}$ being the n belonging to frequency α_l , and $A_n^{(l)}$,

$$k = n^{(1)}\alpha_1 + n^{(2)}\alpha_2 + \dots + n^{(L)}\alpha_L = \sum_{l=1}^L n^{(l)}\alpha_l, \quad (34)$$

As the numerator is simply the derivative of the denominator, the identification may be done in either the numerator or denominator, yielding

$$\begin{aligned} \sum_{k=-\infty}^{\infty} c_k e^{ik\theta^*} &= \sum_{n^{(1)}=-\infty}^{\infty} A_n^{(1)} \exp(in^{(1)}\alpha_1 \theta^*) \\ &\times \sum_{n^{(2)}=-\infty}^{\infty} A_n^{(2)} \exp(in^{(2)}\alpha_2 \theta^*) \dots \\ &\sum_{n^{(L)}=-\infty}^{\infty} A_n^{(L)} \exp(in^{(L)}\alpha_L \theta^*). \end{aligned} \quad (35)$$

This equation is solved for each k . The coefficients become

$$c_k = \sum_{k=\sum n^{(l)}\alpha_l} A_n^{(1)} A_n^{(2)} \dots A_n^{(L)}. \quad (36)$$

Equation (36), together with (15), is the solution to Burgers' equation (1), valid for any number of frequencies with arbitrary amplitudes and phases.

This new solution has many advantages. It is the first exact multifrequency solution in explicit form, which includes dissipation. It is suitable as a replacement for numerical methods. It works fast on a computer, from a couple of seconds for two initial frequencies up to a few minutes for 200. Once the solution is obtained, it is valid for all distances (which, of course, is difficult to match by pure numerics).

Approximate analytical expressions describing the behavior in limiting cases might, of course, be derived from

this exact solution. But as the parameters in a multifrequency condition contains many parameters like relative amplitudes, relative frequencies, and relative phases, one cannot expect to arrive at general asymptotic expressions. Actually, the exact individual frequencies in form of a straightforward Fourier series may be obtained analytically in the same way as for a single-frequency.³ But the final analytical results would be intricate and is therefore not included in this article.

The most analytical way of calculating the coefficients $A_n^{(l)}$ is through the modified Bessel functions in the explicit solution (13). A faster way on a computer is to use a recursion method, formulated by Gallia, which is described in Appendix A.

When the dissipation ϵ^* becomes small, a larger number of terms will have to be included in the series. Due to the slow convergence of these terms and the difference in size of the coefficients, there is a lower limit in ϵ below which the calculations lead to numbers exceeding the numerical representation in computers. This is the only limitation in the presented method, the same as for the Cole–Mendousse solution (13).

III. NUMERICAL RESULTS

In this section we present some results showing the evolution of multifrequency signals. Comparisons with numerical and analytical methods are made. The numerical method's accuracy is validated in Appendix B, where it is compared to well-known single-frequency solutions.

The first example is $V_0 = \sin 7\theta + \sin 9\theta$ for the dissipation over nonlinearity ratio $\epsilon^* = 0.05$, where the difference frequency, $9 - 7 = 2$, is not the same as the lowest frequency, which is equal to 1.

In Fig. 1 it is seen how combination frequencies are created. Note how the lowest frequency is not noticeable, as it cannot be created directly from 7 and 9. It is created from, e.g., $(4 \times 7 = 28) - (3 \times 9 = 27) = 1$, giving a very small contribution. After some distance, $\sigma = 5$, the difference fre-

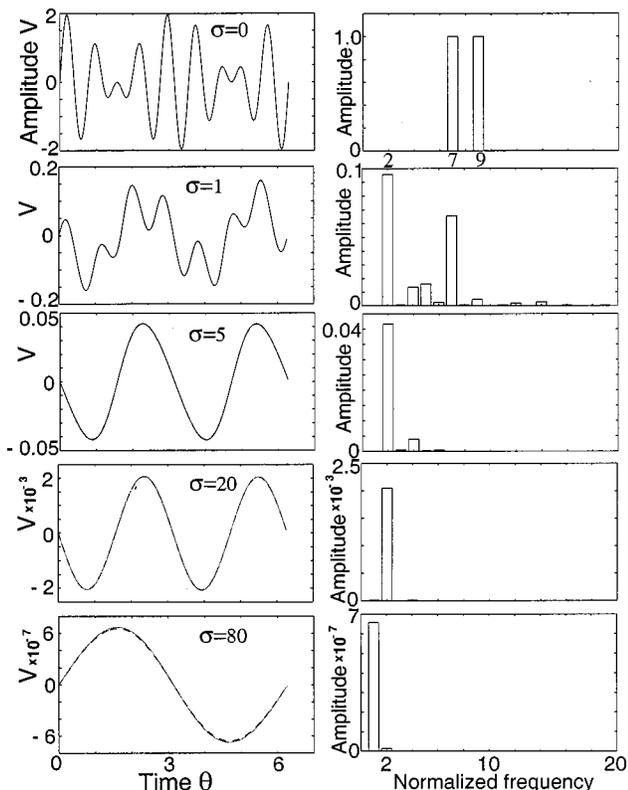


FIG. 1. The difference frequency $9 - 7 = 2$ is not the same as the lowest frequency equal to 1. Here $V_0 = \sin 7\theta + \sin 9\theta$, $\epsilon^* = 0.05$. The numerical solution is a dashed line and the multi frequency solution is solid.

quency has become dominant and many of the higher harmonics have been damped out. The wave will stay in this approximate shape for a long distance; see, for example, $\sigma = 20$. One might believe this to be the final stage. But later on, at $\sigma = 80$, the higher damping of frequency 2 compared to that of frequency 1, has made the lowest frequency the only one remaining. The comparison with a numerical algorithm indicates that both solutions seem to be correct.

To make a comparison with the Fenlon solution, it is necessary to stay in the preshock region, which in this case means up to the distance $\sigma_s \approx 1/(7 + 9) = 0.0625$. In Fig. 2 is the multifrequency solution and the numerical solution for a dissipation of $\epsilon = 0.05$ and the Fenlon solution, which is inherently for zero dissipation, is shown. The expected difference is present.

Next, the input is an *N*-wave that is represented in the initial wave with 210 frequencies; see Fig. 3. The calculation is also performed with 210 frequencies and is compared with

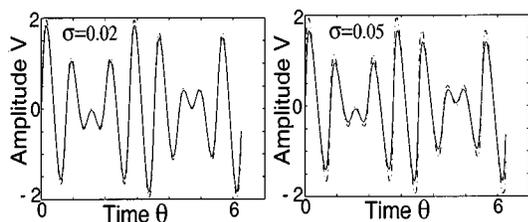


FIG. 2. The present solution (solid), the numerical solution (dashed), and the Fenlon solution (dash-dotted) in its region of validity before shock formation, $V_0 = \sin 7\theta + \sin 9\theta$, $\epsilon^* = 0.05$.

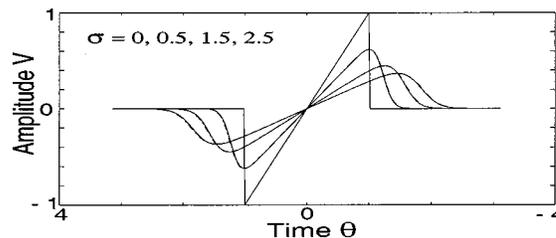


FIG. 3. The initial *N* wave and its evolution according to the multifrequency solution (solid), compared with a numerical algorithm (dash-dotted), $\epsilon^* = 0.04$.

a numerical solution. This shows not only that a large number of harmonics may be used in the input but also that the propagation of pulses are possible to calculate.

IV. CONCLUSIONS

A solution for multifrequency plane waves propagating through a dissipative and nonlinear medium has been presented. Any frequencies, number of frequencies, amplitudes, and phases can be chosen as input. It originates from a solution for a pure sinusoidal wave, known from Cole and Mendousse, expressed as a ratio between Bessel function series. The solution for multiple frequencies is calculated from the knowledge that the solution has to have a specific form and that the unknown coefficients in this form can be extracted at zero distance from the sum of the single-frequency solutions. As the single frequency solutions are exact, so is the solution for the multiple-frequency waves. For the same reason, the limitation in that the series converge slowly when the dissipation over nonlinearity ratio is very large is inherited from the single-frequency solution. The solution depends on the separation of variables, which means that once the coefficients are calculated, one automatically has a solution valid for any distance.

The validity of the method has been shown by comparison to a numerical algorithm. The method is considerably faster than regular numerical methods for large distances, in particular, when the number of input frequencies are not great.

The solution lends itself to the extraction of analytical limiting solutions for any given set of parameters.

ACKNOWLEDGMENTS

I was inspired to do this work when looking at the recursion formula work by Ken Gallia, University of Texas at Austin. The idea came to me after my postdoc years 1995–1997 at the University of Texas supported by TFR (the Swedish Research Council for Engineering Sciences), but the work there prepared me for this article. I would like to thank the people at the Acoustics group at the University of Texas and especially my mentor Bengt Enflo, KTH, for his continual advice.

APPENDIX A: THE BESSEL RECURSION ALGORITHM

This appendix describes the recursion formula developed by Gallia at the University of Texas.

Put Eq. (21) equal to a solution of the form (15) at $\sigma = 0$:

$$V(\sigma=0, \theta) = \sin \theta = \frac{1}{i2} e^{i\theta} - \frac{1}{i2} e^{-i\theta} \\ = i2\epsilon \frac{\sum_{k=-\infty}^{\infty} k A_k e^{ik\theta}}{\sum_{k=-\infty}^{\infty} A_k e^{ik\theta}}. \quad (\text{A1})$$

Multiply by the denominator

$$i2\epsilon \sum_{k=-\infty}^{\infty} k A_k e^{ik\theta} = \sum_{m=-\infty}^{\infty} A_m e^{im\theta} \left[\frac{1}{i2} e^{i\theta} - \frac{1}{i2} e^{-i\theta} \right]. \quad (\text{A2})$$

Here the coefficients A_k are regarded as unknowns. One of the possible methods of obtaining these coefficients is to solve this as a system of equations.

But the fastest way to get the solution is through the identification of the exponents giving us the recursion formula

$$-4\epsilon k A_k = A_{k-1} - A_{k+1}, \quad (\text{A3})$$

or, after rearrangement and insertion of the chosen normalized variable ϵ^* from (20),

$$A_{k-1} = A_{k+1} + 4 \frac{\alpha \epsilon^*}{a} k A_k. \quad (\text{A4})$$

Starting with almost any values of $A_{K+K_{\text{plus}}}$ and $A_{K+K_{\text{plus}}-1}$, where K is the number of terms in the series and K_{plus} is approximately 10, the coefficients are calculated backward. The recursion ends with

$$A_0 = A_2 + 4 \frac{\alpha \epsilon^*}{a} A_1. \quad (\text{A5})$$

These coefficients have been multiplied with a factor that is dependent on the starting values. But as the final solution is a ratio this factor will disappear.

APPENDIX B: EVALUATION OF NUMERICAL METHOD

The numerical program used to validate the multifrequency solution is a split step superposition algorithm in which the nonlinear part is performed in time domain and the dissipative in frequency domain. The nonlinear part is an exact solution to Eq. (1) with the third term absent and the dissipative part is the solution to (1) when the second term is absent. The numerical code will, in this appendix itself, be validated by comparisons with known solutions existing for a single frequency. Three solutions are represented in Fig. B1. At all distances is present the Cole–Mendousse solution and the numerical solution. At distance $\sigma=0.5$ the third solution is the Fubini solution. The Fubini solution does not include dissipation, which makes it differ from the other two

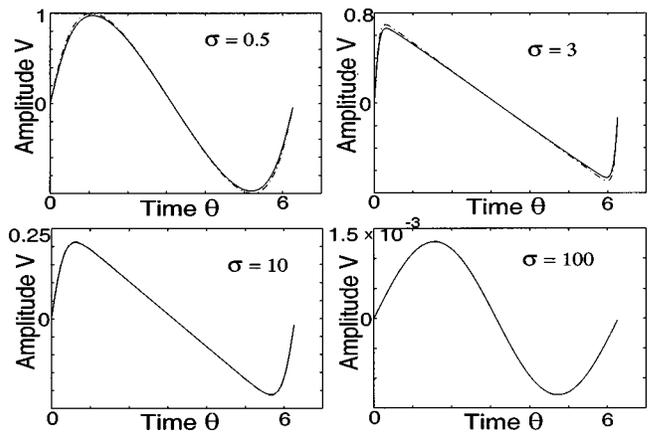


FIG. B1. Here $V_0 = \sin \theta$, $\epsilon^* = 0.05$. The Cole–Mendousse solution (solid line), the numerical solution (dashed), and, at $\sigma=0.5$, the Fubini solution (dash-dotted); at $\sigma=3, 10$, and 100 the Fay solution (dash-dotted). The solutions coincide well at the distances where they are valid.

primarily by its larger amplitude. At distances $\sigma=3, 10$, and 100 the third solution is the Fay solution. For the distance $\sigma=3$ the Fay solution is still not very close to being exact, as its lower limit of validity is approximately $\sigma=4$, which means that at $\sigma=10$ all three solutions coincide. The Fay and the Cole–Mendousse solutions are valid for indefinitely large values. At $\sigma=100$, the numerical solution is coinciding fairly well, which proves its stability and suitability as a comparing tool for the multifrequency solution.

- ¹G. E. Fubini, “Anomalia nella propagazione di onde acustiche di grande ampiezza,” *Alta Freq.* **4**, 530 (1935).
- ²R. D. Fay, “Plane sound waves of finite-amplitude,” *J. Acoust. Soc. Am.* **3**, 222 (1931).
- ³B. O. Enflo, “Some analytical results on nonlinear acoustic wave propagation in diffusive media,” *Radiofizika* **36**, 665–686 (1993).
- ⁴J. D. Cole, “On a quasi-linear parabolic equation occurring in aerodynamics,” *Q. Appl. Math.* **9**, 225–236 (1951).
- ⁵J. S. Mendousse, “Nonlinear dissipative distortion of progressive sound waves at moderate amplitudes,” *J. Acoust. Soc. Am.* **25**, 51–54 (1953).
- ⁶D. T. Blackstock, “Thermoviscous attenuation of plane, periodic, finite-amplitude sound waves,” *J. Acoust. Soc. Am.* **36**, 534–542 (1964).
- ⁷J. M. Burgers, *The Nonlinear Diffusion Equation* (Reidel, Dordrecht, 1977).
- ⁸F. H. Fenlon, “An extension of the Fubini series for a multiple-frequency CW acoustic source of finite amplitude,” *J. Acoust. Soc. Am.* **51**, 284–289 (1972).
- ⁹E. Hopf, “The partial differential equation $u_t + uu_x = \mu$,” *Commun. Pure Appl. Math.* **3**, 201–230 (1950).
- ¹⁰O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Plenum, New York, 1977).
- ¹¹C. M. Hedberg, “One-saddle point solution to Burgers’ equation for harmonic signals,” submitted to *J. Acoust. Soc. Am.*
- ¹²S. I. Soluyan and R. V. Khokhlov, “Propagation of acoustic waves of finite amplitude in a dissipative medium, *Vestn. Mosk. Univ. (Series III)*,” *Fiz. Astron.* **3**, 52–61 (1961); or in Beyer, *Nonlinear Acoustics in Fluids* (Van Nostrand Reinhold, New York, 1984), pp. 193–206.
- ¹³C. M. Hedberg, “Nonlinear propagation through a fluid of waves originating from a biharmonic sound source,” *J. Acoust. Soc. Am.* **96**, 1821–1828 (1994).

Nonlinear oscillation of a spherical gas bubble in acoustic fields

Masaharu Kameda

Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan

Yoichiro Matsumoto

Department of Mechanical Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan

(Received 28 August 1998; revised 15 September 1999; accepted 16 September 1999)

Radial motion of a spherical air bubble in acoustic fields is observed experimentally. The radius-time curves and frequency responses are obtained from the experiment for comparison with a numerical calculation. The calculation is based on a mathematical model in which the thermo-fluid mechanics of the gas in the bubble is precisely described. An oscillatory pressure field is generated in a cylindrical cell, which consists of two piezoceramic transducers and a glass cylinder. A new bubble generator is developed. It is able to generate a bubble filled with an arbitrary kind of gas. The bubble motion is observed by high-speed photography. The time history of the bubble radius is measured from the pictures. The pressure field has a frequency of 19.2 kHz and its amplitude is up to 40 kPa. The bubble has an initial radius within the range from 0.1 mm to 0.25 mm. A highly viscous silicone oil, whose kinematic viscosity is 100 mm²/s, is used for the liquid to keep the spherical shape of the bubble. A quantitatively good agreement between the experimental and numerical results is obtained. The difference between experiment and theory based on the polytropic approximation for the gas is briefly discussed. © 1999 Acoustical Society of America. [S0001-4966(99)07812-1]

PACS numbers: 43.25.Yw, 43.30.Lz [MAB]

INTRODUCTION

The motion of bubbles due to pressure changes is a fundamental subject to understand the dynamics of cavitating flow. It is also important for understanding the acoustic properties of a liquid containing many bubbles.

The radial motion of a single gas bubble plays the most dominant role among the several different motions of a bubble liquid in an acoustic field. There are a large number of investigations on the radial motion of the bubble, and we can find their results in several reviews (e.g., Refs. 1, 2, and 3).

The most common approach to understand the radial motion of a bubble may be to solve the well-known Rayleigh–Plesset equation,

$$R \frac{d^2 R}{dt^2} + \frac{3}{2} \left(\frac{dR}{dt} \right)^2 = \frac{1}{\rho_l} (p_{lw} - p_{l\infty}), \quad (1)$$

where t , R , ρ_l , p_{lw} , $p_{l\infty}$ denote the time, bubble radius, liquid density, liquid pressure at the surface of the bubble, and liquid pressure infinitely far from the bubble, respectively. The pressure p_{lw} is written as taking the surface tension and the liquid viscosity into account:

$$p_{lw} = p_{gw} - \frac{2\sigma}{R} - \frac{4\mu_l}{R} \frac{dR}{dt}, \quad (2)$$

where p_{gw} , σ , and μ_l are the gas pressure at the bubble wall, the surface tension coefficient, and the liquid viscosity, respectively.

The Rayleigh–Plesset equation is the equation for the radial motion of a single bubble in an incompressible viscous

liquid. Also we have the equation for the radial motion of a bubble in a compressible liquid,^{4–6} and for the motion of a vapor–gas bubble.⁷

On the other hand, we have to determine the pressure at the bubble wall p_{lw} to solve the Rayleigh–Plesset equation (1). For determining the pressure p_{lw} , we have to know the internal phenomena of the bubble interior. The bubble usually consists of noncondensable gas and vapor.

The simplest model for the gas in the bubble is the polytropic approximation, in which the internal gas pressure p_g is calculated from

$$p_g R^{3\kappa} = \text{const} \quad (\kappa = \text{const}), \quad (3)$$

where κ denotes a polytropic index.

There are many investigations based on this polytropic approximation. For example, Lauterborn and co-workers^{8,9} solved the Rayleigh–Plesset equation (1) with Eq. (3) numerically, and showed a lot of interesting results on the nonlinear nature of this system. Francescutto and Nabergoj¹⁰ showed a complete analysis of the main resonance as well as subharmonic and superharmonic resonances of the Rayleigh–Plesset equation with Eq. (3).

However, the polytropic approximation has a serious limitation for many problems, because it cannot correctly describe the thermal behavior of the bubble interior. Indeed, we have a considerable spatial nonuniformity of the temperature in the bubble, which needs to be taken into account. Unfortunately, no analytical solutions for that thermal behavior exist except for the small-amplitude case.¹¹ The behavior is studied numerically by many researchers.^{12–15} Kamath and Prosperetti¹⁶ show that the numerical results of the

Rayleigh–Plesset equation with Eq. (3) are quite different from those including more sophisticated prediction on the thermal behavior of the bubble interior.

Although we have a lot of fruitful results on the dynamics of the radial motion of a bubble, they are almost all analytical or numerical ones. We have only a few useful experimental results. The work done by Holt and Crum¹⁷ is one of them.

Holt and Crum performed a direct measurement of the radial motion of a levitated bubble by means of a light scattering method, and compared their experimental results with the numerical predictions based on the model derived by Prosperetti.¹³ They found a substantial discrepancy between the experimental and numerical results in the second harmonic response. However, we may not conclude that the model is insufficient to predict the radial motion of a bubble, because we have to consider the onset of shape oscillation. They did not observe the shape of the bubble in their experiment, so that they were not able to know whether the shape oscillation occurred.

This paper describes an experimental and numerical investigation on the radial motion of a spherical gas bubble in acoustic fields. We observe the radial motion by means of high-speed photography in order to prove that the bubble retains its spherical shape. We use a resonance cell, originally proposed by Crum,¹⁸ for generating the acoustic field. Since we would like to determine initial conditions of the bubble precisely, we do not use an acoustically generating gas bubble in the resonant cell. We have developed a new bubble generator which can produce a single small bubble. We compare the experimentally obtained radial motion with the numerical prediction in which we calculate the thermal behavior of the bubble interior precisely. To obtain the precise prediction of the thermal behavior, we solve the conservation equations for the gas in the bubble by means of the method originally developed by Takemura.^{15,19} At the same time, we perform the numerical calculation based on a refined polytropic approximation to discuss its limitations.

I. EXPERIMENTS

A. Experimental apparatus

Figure 1 shows the schematic diagram of the experimental apparatus. The apparatus is divided into three parts: a resonance cell, a bubble generator, and an optical system.

The resonance cell is a cylindrical container consisting of two cylindrical piezoceramic transducers and a glass cylinder. The cell has an inner diameter of 40 mm and a height of 100 mm. We surround the cell with a rectangular box with optical windows. The space between the cell and the box is filled with liquid to reduce optical distortions due to refraction at the surface of the glass cylinder. The piezoceramic transducers, poled to be driven in the thickness mode, have an inner diameter of 40 mm, an outer diameter of 60 mm, and a height of 25 mm. A function generator and a high-frequency power amplifier are used as the driving source.

The bubble generator¹⁹ is installed in the bottom of the cell. Figure 2 shows its cross section. It consists of a solenoid, a iron bar, a spring, and a casing. A pinhole is attached

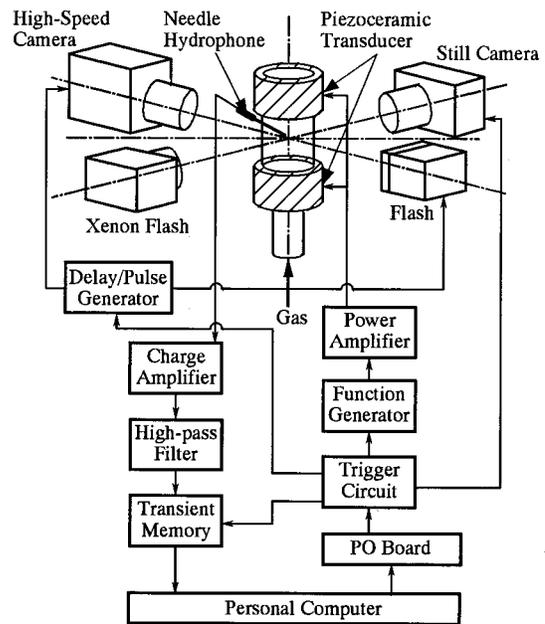


FIG. 1. Experimental setup.

on the top of the casing. A semispherical rubber is attached on the top of the iron bar. An arbitrary kind of gas is introduced in the casing. The gas is pressurized up to slightly higher than the liquid pressure at the bottom of the resonant cell. When we switch the solenoid on, the iron bar moves down. Then when we switch it off, the bar springs up and hits the pinhole. Consequently, we can make a small bubble by this hitting.

The size of this bubble depends on the diameter of pinhole and the pressure difference between the gas in the casing and the liquid at the bottom of the cell. We use a pinhole

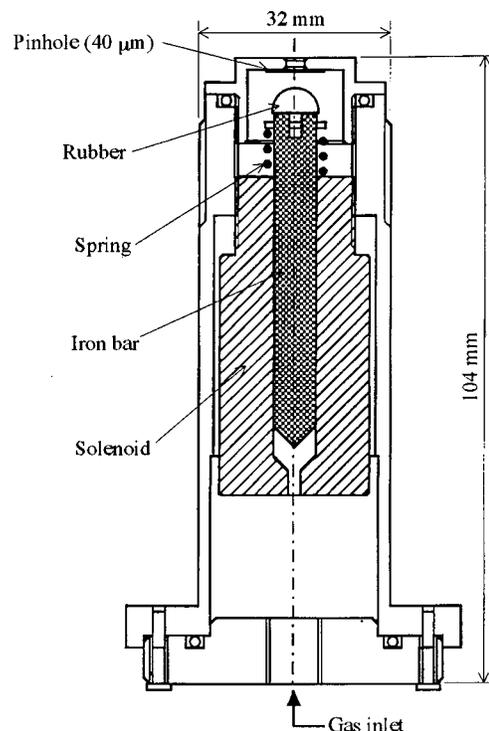


FIG. 2. Bubble generator.

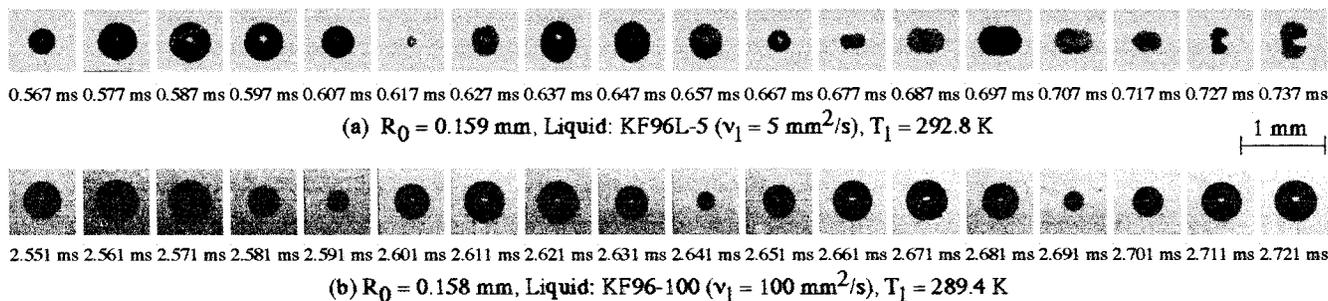


FIG. 3. Transient shape change of an air bubble in silicone oil. (a) The kinematic viscosity $\nu_1 = 5$ mm²/s. (b) $\nu_1 = 100$ mm²/s.

having a diameter of 40 μm in the present experiment. The pressure difference is maintained 2 kPa by using a precision pressure controller (Druck DPI500). Then, we can obtain a gas bubble having an initial radius of 0.2 mm.

A PVDF needle hydrophone (Imotec Type 80-0.5-4.0) is used to measure the pressure change far from the bubble. It has a diameter of 1.2 mm and a resonance frequency of 20 MHz. The output signals from the hydrophone is stored in a digital transient memory (Autonics APC-5104) through a charge amplifier (Kistler 5011A) and a high-pass filter with a cutoff frequency of 1 kHz. The transient memory has a sampling frequency of 2 MHz with 10-bit resolution.

High-speed photography is performed to observe the shape change of the bubble. Shadowgraph images of the bubble are taken by an image converter camera (Nac ULTRANAC). The framing rate is fixed at the value of 100 000 frames/s. The total number of frames in each shot is 24. A Xenon flash is used as a light source. A still camera is used simultaneously to take a photograph of the bubble just before the oscillations start.

B. Experimental conditions

We test the motion of a single gas bubble having an initial radius within the range from 0.1 mm to 0.25 mm in the present experiment. The gas in the bubble is air.

We generate an acoustic field with the frequency of 19.2 kHz. The pressure amplitude is up to 40 kPa. The initial pressure is atmospheric. The temperature is room temperature.

We use two different silicone oils with kinematic viscosities of 5 mm²/s and 100 mm²/s at 298.15 K. Their density and surface tension are $\rho_l \approx 965$ kg/m³ and $\sigma \approx 20.9 \times 10^{-3}$ N/m at 298.15 K. Their most remarkable feature is that they have an extremely low vapor pressure. For example, the oil with a kinematic viscosity of 100 mm²/s has a vapor pressure of $O(10^{-4})$ mmHg. Therefore, we can neglect the vapor in the bubble.

C. Viscous effects on bubble shape

Figure 3 shows two photograph sequences of the bubble motion. Figure 3(a) shows the bubble in the low-viscosity silicone oil, while Fig. 3(b) is that in the highly viscous one. The initial bubble radius is 0.159 mm for Fig. 3(a) and 0.158 mm for Fig. 3(b). The pressure amplitude is 40 kPa.

Although the bubble in the low-viscosity silicone oil retains its spherical shape until the first ten periods of pressure

oscillation, its shape suddenly becomes nonspherical at the elapsed time of 0.6 ms. On the other hand, the bubble in the highly viscous oil keeps its spherical shape over 50 periods of the pressure oscillation.

The onset of the shape oscillation depends on several parameters such as liquid viscosity, surface tension, amplitude of radial motion, bubble radius, and frequency of applied acoustic field.^{2,20} Generally speaking, the high viscosity of the surrounding liquid suppresses the onset of the shape oscillation. We find that all the bubbles tested in the present experiment lose their spherical symmetry in the low-viscosity silicone oil even if we set the pressure amplitude to be 10 kPa. On the other hand, we can obtain the radial motion of a spherical gas bubble in the highly viscous silicone oil for all cases up to the pressure amplitude of 40 kPa.

Since we focus our attention on the radial motion of a spherical gas bubble, from now on, we discuss the results only with the highly viscous silicone oil. We measure the time history of bubble radii from the high-speed photographs such as Fig. 3(b) to compare it with the numerical predictions described in the following. The initial bubble radius is determined from the picture obtained through the still camera.

II. NUMERICAL CALCULATION

The Keller equation^{4,5} is used for the equation of the radial motion of a bubble. The conservation equations including heat transfer are solved directly to predict the motion of gas in the bubble.

We employ the following assumptions.

- (1) The bubble consists of gas without vapor.
- (2) The bubble retains its spherical shape during its oscillation.
- (3) The gas in the bubble obeys the perfect gas law.
- (4) The viscosity of gas is ignored.
- (5) No mass diffusion occurs.
- (6) The temperature at the bubble wall keeps a constant value.

Assumptions (1) and (2) are consistent with the present experiments.

The normal viscous stress in the radial direction of the gas is small compared to the gas pressure,¹⁹ so that assumption (4) is valid.

Although the total mass of the gas in the bubble changes due to the rectified diffusion during the oscillation, its rate of

change is quite small^{18,21} compared to the phenomena discussed here. We study the bubble oscillation up to a maximum of 60 periods after it starts to move. Thus we can assume condition (5).

Prosperetti *et al.*¹³ discussed the validity of assumption (6). According to their result, the temperature at the bubble wall T_w is estimated as

$$\frac{T_w - T_l}{T_{gc} - T_w} \sim \left(\frac{\lambda_g c_{pg} \rho_g}{\lambda_l c_{pl} \rho_l} \right), \quad (4)$$

where T_l denotes the undisturbed liquid temperature, T_{gc} is the temperature at the center of the bubble, ρ the density, λ the thermal conductivity, and c_p the specific heat at constant pressure. The subscripts l and g represent the liquid and gas phase, respectively. The ratio on the right-hand side in Eq. (4) for the present study has $O(10^{-4})$. Therefore we can assume condition (6).

The Keller equation is

$$\begin{aligned} & \left(1 - \frac{1}{C_l} \frac{dR}{dt} \right) R \frac{d^2 R}{dt^2} + \frac{3}{2} \left(1 - \frac{1}{3C_l} \frac{dR}{dt} \right) \left(\frac{dR}{dt} \right)^2 \\ &= \frac{1}{\rho_l} \left(1 + \frac{1}{C_l} \frac{dR}{dt} + \frac{R}{C_l} \frac{d}{dt} \right) (p_{lw} - p_{ls}). \end{aligned} \quad (5)$$

Here ρ_l and C_l are the density and speed of sound of the undisturbed liquid, while p_{ls} is the surrounding liquid pressure. The liquid pressure at the bubble wall p_{lw} is defined in Eq. (2).

The conservation equations for the gas in the bubble are

$$\frac{\partial \rho_g}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \rho_g u_g) = 0, \quad (6)$$

$$\frac{\partial \rho_g u_g}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \rho_g u_g^2) = - \frac{\partial p_g}{\partial r}, \quad (7)$$

$$\begin{aligned} \frac{\partial \rho_g e_g}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \rho_g e_g u_g) &= - \frac{p_g}{r^2} \frac{\partial}{\partial r} (r^2 u_g) \\ &+ \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \lambda_g \frac{\partial T_g}{\partial r} \right). \end{aligned} \quad (8)$$

Here r and u denote the radial coordinate and the radial velocity component. The point $r=0$ corresponds to the bubble center. The specific internal energy e_g is

$$e_g = c_{vg} T_g, \quad (9)$$

where c_v denotes the specific heat at constant volume.

The equation of state for the perfect gas is

$$p_g = \rho_g \mathfrak{R} T_g, \quad (10)$$

where \mathfrak{R} denotes the gas constant divided by the molecular mass of the gas.

The boundary conditions at the bubble center are

$$u_g = 0, \quad \frac{\partial \rho_g}{\partial r} = 0, \quad \frac{\partial T_g}{\partial r} = 0. \quad (11)$$

Those at the bubble wall ($r=R$) are

$$u_g = \frac{dR}{dt}, \quad T_g = T_l (= \text{const}). \quad (12)$$

The set of equations described above is solved by means of the method originally developed by Takemura.^{15,19} A second-order Runge–Kutta method is applied to solve the Keller equation (5). The CIP (Cubic Interpolated pseudo-Particle) method²² is used to solve the conservation equation for momentum Eq. (7) and that for energy Eq. (8). A finite volume method is used to solve the conservation equation for mass Eq. (6). The outline of the present numerical method is described in Appendix.

III. RESULTS AND DISCUSSION

In this section, we show some experimental results and compare them with the present numerical predictions. Hereafter, we call the governing equations described in Sec. II as the *present model*.

In addition to them, we introduce some numerical results based on a refined polytropic approximation to show its limitation. The words *polytropic model* indicates this polytropic approximation. In this approximation, we use the Keller equation (5) with the following liquid pressure at the bubble wall p_{lw} , which is

$$p_{lw} = p_{g0} \left(\frac{R_0}{R} \right)^{3\kappa_e} - \frac{2\sigma}{R} - \frac{4(\mu_l + \mu_{th})}{R} \frac{dR}{dt}, \quad (13)$$

where p_{g0} denotes the initial gas pressure in the bubble, and $p_{g0} = p_{l0} + 2\sigma/R_0$. κ_e and μ_{th} are the ‘‘effective’’ polytropic index and thermal viscosity. As studied by Prosperetti,¹¹ these quantities are well defined only in the framework of a linear theory. Nevertheless, Eq. (13) seems to be a reasonable extension to the nonlinear case, and should work better than the standard polytropic approximation in which no thermal dissipation is included. Thus following Prosperetti,¹¹ we take

$$\kappa_e = \frac{1}{3} \text{Re } \bar{F}, \quad (14)$$

$$\mu_{th} = \frac{p_{g0}}{4\omega} \text{Im } \bar{F}. \quad (15)$$

The function \bar{F} is

$$\bar{F} = \frac{3\gamma\eta^2}{\eta[\eta + 3(\gamma - 1)A_-] - 3i(\gamma - 1)(\eta A_+ - 2)}, \quad (16)$$

$$\eta = R_0 \sqrt{\frac{2\omega}{\chi}} \left(\chi = \frac{\lambda_{g0}}{c_{vg}\rho_{g0}} \right), \quad A_{\pm} = \frac{\sinh \eta \pm \sin \eta}{\cosh \eta - \cos \eta},$$

where γ is the ratio of specific heats, while ρ_{g0} and λ_{g0} are the initial density and thermal conductivity of the gas phase.

In the present calculation, we consider the temperature dependence of the liquid density ρ_l and the kinematic liquid viscosity ν_l , which are calculated from²³

$$\rho_l = \frac{0.965 \times 10^3}{1 + 9.5 \times 10^{-4}(T_l - 298.15)} [\text{kg/m}^3], \quad (17)$$

$$\nu_l = \exp \left(\frac{1.757 \times 10^3}{T_l} - 1.288 \right) [\text{mm}^2/\text{s}]. \quad (18)$$

The speed of sound in the silicone oil C_l is estimated from the Tait formula, which is

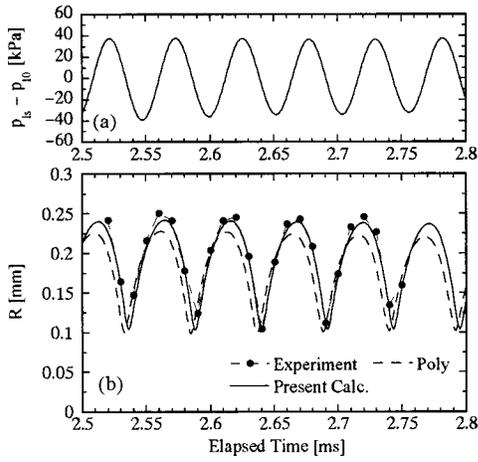


FIG. 4. Radius-time history of an air bubble whose initial radius $R_0 = 0.158$ mm. $f = 19.2$ kHz, $p_{10} = 104$ kPa, and $p_{1a} \approx 40$ kPa. The filled symbols represent the measured bubble radii. Surrounding pressure is shown in (a). The solid and broken lines in (b) are the results predicted by the present and the refined polytropic models.

$$C_l = \sqrt{\frac{n(p_{10} + B)}{\rho_l}}, \quad (19)$$

where $n = 8.669$ and $B = 105.6$ MPa.²³ When we assume $p_{10} = 101.3$ kPa and $\rho_l = 965$ kg/m³ ($T_l = 25^\circ\text{C}$), we have $C_l = 974.5$ m/s.

We take the temperature dependence on the thermal conductivity of the gas into account. It is estimated by the Sutherland law²⁴ as

$$\lambda_g = \lambda_{g,0^\circ\text{C}} \frac{273.15 + C_s}{T_g + C_s} \left(\frac{T_g}{273.15} \right)^{3/2}, \quad (20)$$

where $\lambda_{g,0^\circ\text{C}}$ denotes the thermal conductivity at 0°C , and C_s is the so-called Sutherland constant. Their values for air are $\lambda_{g,0^\circ\text{C}} = 24.07 \times 10^{-3}$ W/(m·K) and $C_s = 118.3$ K, respectively.

The surrounding liquid pressure p_{ls} can be described by the following equation except for the first ten periods:

$$p_{ls} = p_{10} - p_{1a} \cos 2\pi ft, \quad (21)$$

where p_{10} denotes the undisturbed surrounding pressure, p_{1a} the pressure amplitude, and f the driving frequency.

A. Radius-time curves

We will show three typical radius-time curves in this subsection. It is convenient to recall the linear natural frequency of the bubble oscillation:

$$f_0 = \frac{1}{2\pi R_0} \sqrt{\frac{1}{\rho_l} \left(3\kappa_e p_{g0} - \frac{2\sigma}{R_0} \right)}, \quad (22)$$

where κ_e is the effective polytropic index shown in Eq. (14).

Figure 4 shows a comparison of the experimentally measured radius-time history with the numerical predictions. The experimental data are obtained from the photograph shown in Fig. 3(b). The initial bubble radius R_0 is 0.158 mm, the undisturbed liquid temperature T_l is 16.2°C , and the pressure amplitude p_{1a} is about 40 kPa. The surrounding pressure is shown in Fig. 4(a). Symbols in Fig. 4(b) indicate the ex-

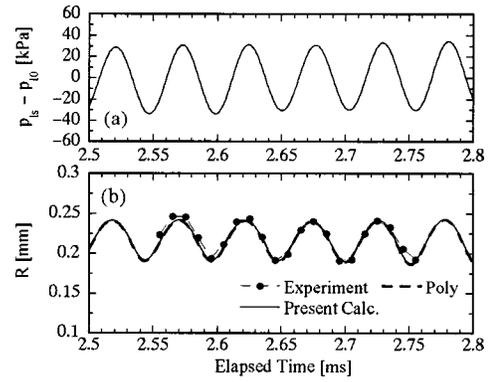


FIG. 5. Radius-time history of an air bubble of $R_0 = 0.214$ mm. $f = 19.2$ kHz, $p_{10} = 104$ kPa, $p_{1a} \approx 30$ kPa, and $T_l = 286.8$ K. The notations are the same as those shown in Fig. 4.

perimental data. For the case shown in Fig. 4, the effective polytropic index and thermal viscosity are $\kappa_e = 1.307$ and $\mu_{th} = 48.6 \times 10^{-3}$ Pa·s. The natural frequency f_0 is 20.6 kHz, so that the ratio of the driving frequency $f = 19.2$ kHz to the natural frequency is 0.933. This means the bubble $R_0 = 0.158$ mm is near the main resonance peak.

The solid and broken lines in Fig. 4(b) are the numerical results based on the present and polytropic models, respectively. In this figure, the elapsed times of the experimental results are shifted in a few microseconds, because we find a certain difference between the experimental and numerical results in terms of their oscillating phase. This difference may be caused by some uncertainty of the experimental setup. Thus we discuss only the amplitude of radial oscillation of bubbles.

The numerical result based on the present model agrees quantitatively well with the experimental one, while the polytropic numerical result differ from them. The maximum and minimum radii in a period are $R_{\max} \approx 0.25$ mm and $R_{\min} \approx 0.1$ mm, so that the amplitude of bubble oscillation $(R_{\max} - R_{\min})/2R_0$ is 0.45. That is, Fig. 4 shows a large-amplitude oscillation. This implies the polytropic approximation is no longer valid for such a case, even though we take the refined polytropic model.

Figures 5 and 6 show two other radius-time curves. The initial bubble radius is $R_0 = 0.214$ mm for Fig. 5 and R_0

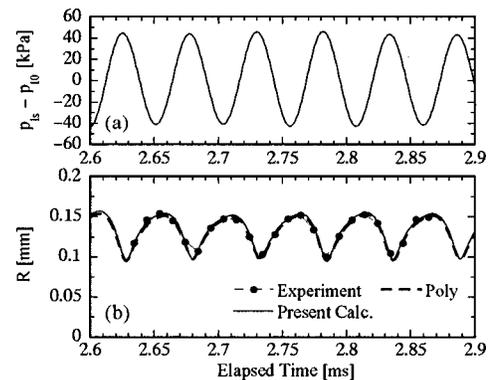


FIG. 6. Radius-time history of an air bubble of $R_0 = 0.126$ mm. $f = 19.2$ kHz, $p_{10} = 104$ kPa, $p_{1a} \approx 40$ kPa, and $T_l = 282.5$ K. The notations are the same as those shown in Fig. 4.

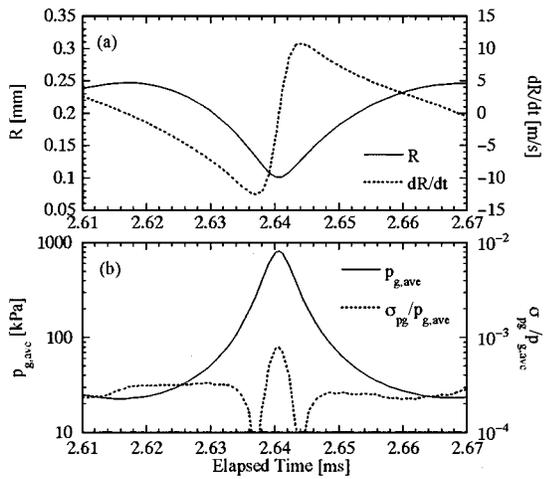


FIG. 7. The numerically predicted time histories of radius, velocity at the bubble wall, averaged gas pressure, standard deviation for pressure distribution for the steady oscillation of the bubble shown in Fig. 4.

$=0.126$ mm for Fig. 6. The amplitude of bubble oscillation $(R_{\max} - R_{\min})/2R_0$ is 0.12 for Fig. 5 and 0.22 for Fig. 6. The ratio f/f_0 is 1.26 (Fig. 5) and 0.741 (Fig. 6). These two curves are far from the main resonance peak.

For these cases, there is no substantial difference between the numerical results based on the present and refined polytropic models. They agree well with the experimental data. These agreements suggest that the refined polytropic model works well for the not-so-large-amplitude oscillation.

B. Behavior of gas in the bubble

We now turn to a detailed observation of the bubble oscillation based on the numerical results with the present model.

First, we examine the spatial uniformity of the pressure field in the bubble. Figure 7 shows time history of the average gas pressure in the bubble $p_{g,ave}$ and the standard deviation of $p_{g,ave}$. The calculation conditions are equal to those for Fig. 4. The standard deviation σ_{pg} is defined as

$$\sigma_{pg} = \left[\sum_{i=1}^N (p_{g,i} - p_{g,ave})^2 / N \right]^{1/2}, \quad p_{g,ave} = \left(\sum_{i=1}^N p_{g,i} \right) / N,$$

where N is the number of grids for discretizing in the bubble and $p_{g,i}$ denotes the pressure at the i -th grids. In addition, the time histories of bubble radius and velocity of the bubble wall are also shown in Fig. 7(a).

The normalized standard deviation of the pressure field is less than 10^{-3} in all the time. This result means that we can assume that the gas pressure is spatially uniform in the bubble. The maximum velocity of the bubble wall is 12 m/s, so that the maximum Mach of the bubble wall of 3.5×10^{-2} is quite small. Prosperetti *et al.*¹³ derived a partial differential equation for the energy of the gas in the bubble with spatially uniform pressure. Their approach is valid for the case discussed here.

Figure 8 shows some temperature profiles at different times in the bubble whose radius-time curve is shown in Fig. 4. The instants of time attached to the figure correspond to

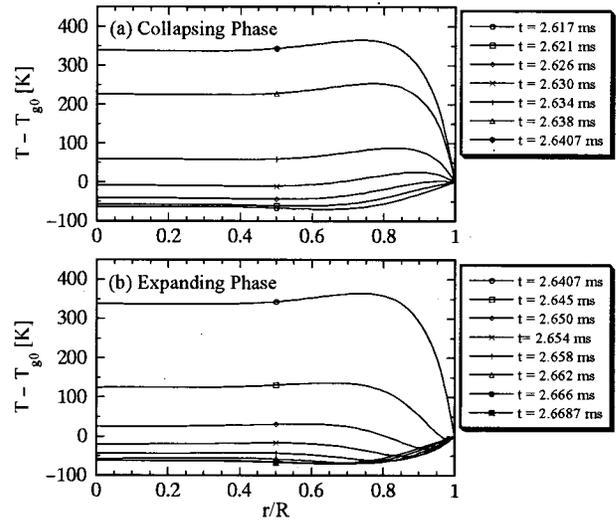


FIG. 8. The numerically predicted spatial distribution of temperature in the bubble shown in Fig. 4.

the elapsed time shown in Fig. 4. Figure 8(a) and (b) shows the change of profiles when the bubble is in its collapsing and expanding phases, respectively.

Although the profile is spatially uniform near the center of the bubble, we can find a considerable thermal boundary layer near the bubble wall. We can also see that the profiles are not monotonic due to the heat exchange between the liquid and the gas near the bubble wall. Nigmatulin *et al.*¹² and Prosperetti *et al.*¹³ have already shown nonuniform temperature distribution similar to those shown in Fig. 8.

C. Frequency response curves

In this subsection, we will describe the relation of the bubble between its initial radius and its oscillating amplitude.

For the convenience of the discussion in this subsection, we define the linear resonance radius R_{res} as the equilibrium radius of a bubble whose linear resonance frequency f_0 defined in Eq. (22) equals the driving frequency f . When we use the effective polytropic index κ_e to calculate the resonance frequency f_0 , we obtain the resonance radius $R_{\text{res}} = 0.17$ mm in the present experiment.

Figure 9 shows the maximum and minimum radii of bubbles in their steady oscillation. The pressure amplitude of the applied acoustic field is 40 kPa.

The filled symbols are the experimental data obtained from the radius-time curves such as Fig. 4. The maximum bubble radius R_{\max} increases suddenly near the initial bubble radius R_0 of 0.15 mm ($R_0/R_{\text{res}} = 0.88$). Then it keeps a constant value, 0.25 mm, in the range of $0.15 \text{ mm} \leq R_0 \leq 0.225 \text{ mm}$ ($0.88 \leq R_0/R_{\text{res}} \leq 1.3$).

The solid and broken lines are the numerical predictions based on the present model described in Sec. II. The numerical predictions agree quantitatively well with the experimental data.

Figures 10 and 11 show resonance curves for two different pressure amplitude, $p_{la} = 20$ and 40 kPa. The vertical

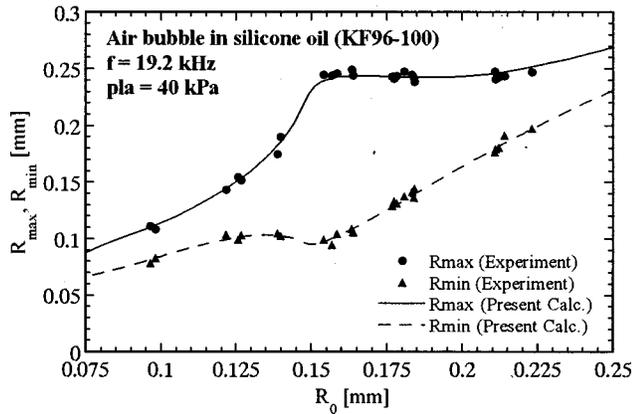


FIG. 9. Minimum and maximum bubble radii versus initial bubble radius. Symbols represent the experimental results. The lines are the results predicted by the present model.

axis is the oscillating amplitude $(R_{\max} - R_{\min})/2R_0$. The filled symbols are the experimental data.

When the pressure amplitude p_{la} is 20 kPa, the position of the peak of the main resonance is close to the linear resonance radius R_{res} of 0.17 mm. On the other hand, the peak shifts toward lower values of R_0 , when the amplitude $p_{la} = 40$ kPa. As already discussed in Ref. 8, this is a typical feature of a “softening nonlinearity.”

The experimental data for the small initial bubble radius slightly scatter because of the error coming not only from the resolution of the pictures but also from the hazy bubble surface in the pictures.

We used a scanner for photo-films (Nikon Coolscan II)

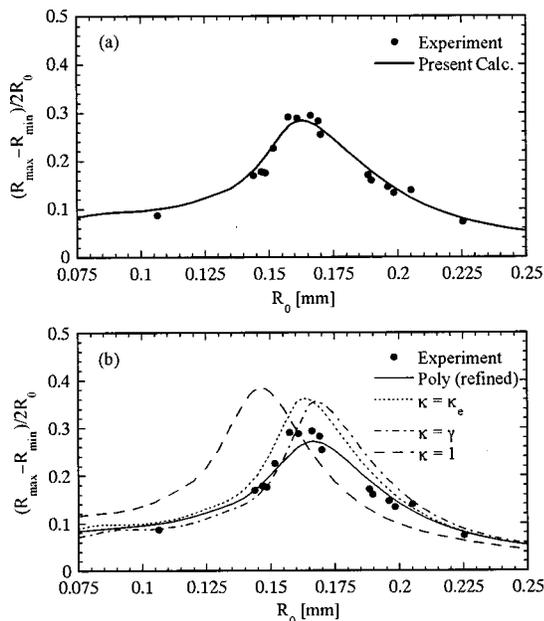


FIG. 10. Normalized oscillation amplitude $[(R_{\max} - R_{\min})/2R_0]$ versus initial bubble radius. $p_{la} = 20$ kPa. Symbols represent the experimental results. The line in (a) shows the numerical result predicted by the present models. The lines in (b) are the predicted results based on the polytropic approximation. The solid line represents the refined polytropic model in which the effective polytropic index and thermal viscosity are taken into account. The dotted, broken, and dash-and-dotted lines show the standard polytropic approximations in which the effective thermal viscosity are neglected.

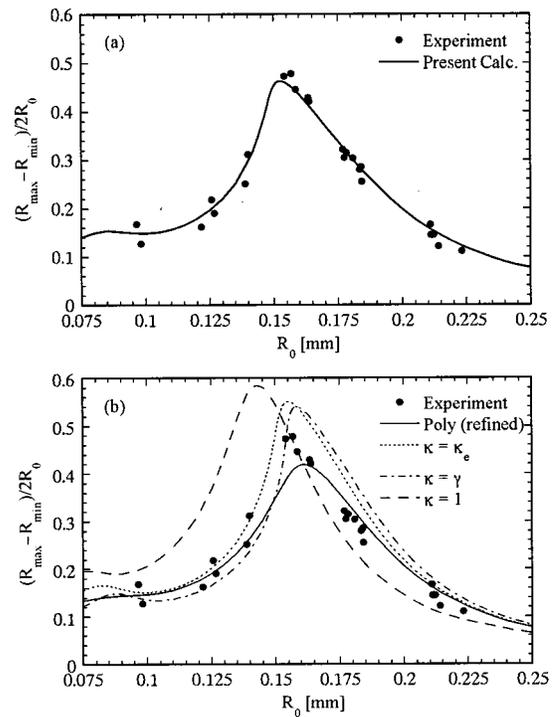


FIG. 11. Normalized oscillation amplitude $[(R_{\max} - R_{\min})/2R_0]$ versus initial bubble radius. $p_{la} = 40$ kPa. The notations are the same as those shown in Fig. 10.

to digitize the pictures. Then we used a photo-retouching software (Adobe PhotoShop) for finding the edge of bubbles to measure their radius. A pixel in the present digitized images equals $4 \mu\text{m}$.

On the other hand, the hazy zone may be caused not only by diffraction but also by “out-of-focus.” The hazy zone around the bubble surface spreads 6 or 7 pixels (about $25 \mu\text{m}$) in a digitized image. Thus we determine the bubble surface, at which the intensity is approximately one-half that which would be expected based on a geometrical optics. This determination may cause the error typically $5 \mu\text{m}$ in the bubble radius.

For the larger bubbles ($R_0 \geq 0.15$ mm), those uncertainties is sufficient to discuss the difference between the experimental result and the two different predictions, although it may be not enough for the smaller bubbles such as $R_0 \leq 0.1$ mm.

The lines shown in Figs. 10(a) and 11(a) show the numerical predictions based on the present model. We can find a good agreement between the experimental data and the prediction based on the present model. This agreement proves the validity of the present model.

The lines shown in Figs. 10(b) and 11(b) show the predictions based on the polytropic approximation. We test not only the refined polytropic model [Eqs. (13)–(16)] but also the “standard” polytropic approximations in which we neglect the effective thermal viscosity μ_{th} in Eq. (13). We take three values for the polytropic index for the standard polytropic approximations: $\kappa = 1$, κ_e , and γ .

Although the refined polytropic model works well within the range of the amplitude $(R_{\max} - R_{\min})/2R_0 < 0.2$, the position and height of the peak of the main resonance are

different from the experimental data. The predicted height of the peak is lower than the experiment. The predicted initial radius at which the oscillating amplitude has a maximum value is larger than the experiment. On the other hand, the predictions based on the standard polytropic approximations are obviously different from the experimental data. Especially, the predicted maximum oscillating amplitudes are larger than the experiment. These results clearly show that the polytropic approximation has a serious limitation even though we take the refined polytropic model.

IV. CONCLUSIONS

This paper describes an experimental and numerical study of the nonlinear oscillation of a single spherical gas bubble in applied acoustic fields. The primary aim of this study is to serve a good experimental data for quantitative comparison with the theory of bubble dynamics. Second, we try to clarify the limitation of the polytropic approximation for the gas in the bubble. The initial bubble radius we deal with here is in the range from 0.1 mm to 0.25 mm. The driving frequency of the acoustic field is 19.2 kHz. The pressure amplitude is up to 40 kPa.

We have developed a bubble generator by which we can obtain a bubble filled with an arbitrary kind of gas. We observe the oscillating air bubble by high-speed photography. We use a silicone oil whose kinematic viscosity is 100 mm²/s to make the bubble retain its spherical shape during its oscillation. The radius-time curves are obtained for the steady oscillation.

To compare with the experimental data, we numerically solve a set of conservation equations for the gas in which the thermal diffusion is taken into account. The equations are common for compressible inviscid fluids.

There is a quantitatively good agreement between the experimental data and the numerical predictions not only for the radius-time curves but also for the frequency responses.

For the small amplitude motion $(R_{\max} - R_{\min})/2R_0 < 0.2$, the results based on the polytropic approximation, in which the thermal dissipation is taken into account, also agrees well with the experimental data. However, for the larger amplitude, the prediction based on the polytropic approximation differs from the experimental data.

ACKNOWLEDGMENTS

We are deeply grateful to Dr. Fumio Takemura who developed not only the numerical method but also the bubble generator early on this study. We also wish to express our gratitude to Professor Andrea Prosperetti who gave us several helpful comments on this manuscript. We appreciate the help of our students M. Yamada, Y. Shimizu, T. Yano, and T. Chujo. This study was supported in part by a fellowship of Sasakawa Foundation.

APPENDIX: THE NUMERICAL METHOD

1. Bubble radius

We assume to know all the quantities at $t=t^n$. The bubble radius R and the wall velocity dR/dt are evolved by a second-order explicit Runge–Kutta method:

$$R^{n+1/2} = R^n + \frac{1}{2} \Delta t \frac{dR^n}{dt}, \quad (\text{A1a})$$

$$\frac{dR^{n+1/2}}{dt} = \frac{dR^n}{dt} + \frac{1}{2} \Delta t \frac{d^2R^n}{dt^2}, \quad (\text{A1b})$$

$$R^{n+1} = R^n + \Delta t \frac{dR^{n+1/2}}{dt}, \quad (\text{A1c})$$

$$\frac{dR^{n+1}}{dt} = \frac{dR^n}{dt} + \Delta t \frac{d^2R^{n+1/2}}{dt^2}, \quad (\text{A1d})$$

where $\Delta t = t^{n+1} - t^n$. The second derivative of R is obtained from the Keller equation (5) as

$$\begin{aligned} \frac{d^2R}{dt^2} = & \left\{ \frac{1}{\rho_l} \left(1 + \frac{1}{C_l} \frac{dR}{dt} \right) \left(p_{gw} - \frac{2\sigma}{R} - \frac{4\mu_l}{R} \frac{dR}{dt} - p_{ls} \right) \right. \\ & + \frac{R}{\rho_l C_l} \left[\frac{dp_{gw}}{dt} + \frac{2\sigma}{R^2} \frac{dR}{dt} + \frac{4\mu_l}{R^2} \left(\frac{dR}{dt} \right)^2 - \frac{dp_{ls}}{dt} \right] \\ & \left. - \frac{3}{2} \left(1 - \frac{1}{3C_l} \frac{dR}{dt} \right) \left(\frac{dR}{dt} \right)^2 \right\} / \left[\left(1 - \frac{1}{C_l} \frac{dR}{dt} \right) R \right. \\ & \left. + \frac{4\mu_l}{\rho_l C_l} \right]. \quad (\text{A2}) \end{aligned}$$

To obtain $(d^2R/dt^2)^n$, we set all the quantities of the right-hand side of Eq. (A2) to be those at the n -th time level. Similar way is used to obtain $(d^2R/dt^2)^{n+1/2}$. dp_{gw}/dt in Eq. (A2) are calculated as

$$\frac{dp_{gw}^n}{dt} = \frac{2(p_{gw}^n - p_{gw}^{n-1/2})}{\Delta t^n}, \quad (\text{A3a})$$

$$\frac{dp_{gw}^{n+1/2}}{dt} = \frac{2(p_{gw}^{n+1/2} - p_{gw}^n)}{\Delta t}, \quad (\text{A3b})$$

where $\Delta t^n = t^n - t^{n-1}$.

To solve the equations mentioned above, we have to determine the gas pressure at the bubble wall p_{gw} by solving the conservation equations for the gas in the bubble.

2. Gas phase

a. Boundary fitted coordinates

The conservation equations for the gas in the bubble are transformed with an boundary fitted coordinate:

$$y = r/R(t), \quad (\text{A4})$$

where $y=0$ and 1 represent the center and wall of the bubble, respectively.

Using Eq. (A4), Eqs. (6)–(8) are readily transformed into the following equations:

$$\frac{\partial}{\partial t}(r^2 r_y \rho_g) + \frac{\partial}{\partial y}[r^2 \rho_g (u_g - r_t)] = 0, \quad (\text{A5})$$

$$\frac{\partial u_g}{\partial t} + (u_g - r_t) \frac{1}{r_y} \frac{\partial u_g}{\partial y} = - \frac{1}{\rho_g r_y} \frac{\partial p_g}{\partial y}, \quad (\text{A6})$$

$$\begin{aligned} \frac{\partial T_g}{\partial t} + (u_g - r_t) \frac{1}{r_y} \frac{\partial T_g}{\partial y} = & -(\gamma - 1) \frac{T_g}{r^2 r_y} \frac{\partial}{\partial y} (r^2 u_g) \\ & + \frac{1}{\rho_g c_{vg}} \frac{1}{r^2 r_y} \frac{\partial}{\partial y} \left(r^2 \frac{\lambda_g}{r_y} \frac{\partial T_g}{\partial y} \right). \end{aligned} \quad (\text{A7})$$

Here r_t and r_y represent $\partial r / \partial t$ and $\partial r / \partial y$, and are written as

$$r_t = y \frac{dR}{dt}, \quad (\text{A8a})$$

$$r_y = R. \quad (\text{A8b})$$

To transform Eq. (8) into Eq. (A7), we use the equation of state [Eq. (10)] and the relation $c_{vg} = \mathfrak{R} / (\gamma - 1)$. Equation (A5) has a conservative form because we will apply a finite-volume method to solve it. On the other hand, Eqs. (A6) and (A7) have nonconservative forms because we will use the CIP (Cubic Interpolated pseudo-Particle) method²² for solving them.

b. The CIP method

Now we are going to introduce the CIP method in brief. Let us consider a one-dimensional hyperbolic equation:

$$\frac{\partial \phi}{\partial t} + C \frac{\partial \phi}{\partial x} = 0. \quad (\text{A9})$$

Even if the propagation velocity C is a function of t and x , we can locally use a solution:

$$\phi(x_i, t + \Delta t) \approx \phi(x_i - C\Delta t, t), \quad (\text{A10})$$

where x_i is the grid point.

Since the value of ϕ is given only at grid points x_i , $\phi(x_i - c\Delta t, t)$ must be approximated by using these values. In the CIP method, $\phi(x, t)$ between x_{i-1} and x_i are interpolated with a cubic polynomial:

$$F_i(x) = a_i X^3 + b_i X^2 + \phi'_{i-1} X + \phi_{i-1}, \quad (\text{A11})$$

where ϕ_{i-1} and ϕ'_{i-1} are the values and the first spatial derivatives of ϕ at x_{i-1} , and $X = x - x_{i-1}$. If we know ϕ_i and ϕ'_i for all the grid points, we may determine a and b in terms of ϕ and ϕ' as follows:

$$a_i = -2(\phi_i - \phi_{i-1}) / \Delta x^3 + (\phi'_i + \phi'_{i-1}) / \Delta x^2, \quad (\text{A12a})$$

$$b_i = 3(\phi_i - \phi_{i-1}) / \Delta x^2 - (\phi'_i + 2\phi'_{i-1}) / \Delta x, \quad (\text{A12b})$$

where $\Delta x = x_i - x_{i-1}$.

The solution of Eq. (A9) during a short period Δt can be approximated by Eq. (A10), then the values of ϕ_i and ϕ'_i at $t + \Delta t$ are predicted as follows:

$$\begin{aligned} \phi_i^{n+1} &= F_{i+1}(x_i - C\Delta t) \\ &= a_{i+1}(-C\Delta t)^3 + b_{i+1}(-C\Delta t)^2 + \phi'_i(-C\Delta t) + \phi_i, \end{aligned} \quad (\text{A13a})$$

$$\begin{aligned} \phi_i'^{n+1} &= dF_{i+1}(x_i - C\Delta t) / dx \\ &= 3a_{i+1}(-C\Delta t)^2 + 2b_{i+1}(-C\Delta t) + \phi'_i, \end{aligned} \quad (\text{A13b})$$

where the superscript $n+1$ means the time at $t + \Delta t$. This expression is derived for $C < 0$. For $C \geq 0$, we can obtain similar expressions for ϕ_i^{n+1} and $\phi_i'^{n+1}$ as

$$\begin{aligned} \phi_i^{n+1} &= F_i(x_i - C\Delta t) \\ &= a_i(\Delta x - C\Delta t)^3 + b_i(\Delta x - C\Delta t)^2 \\ &\quad + \phi'_{i-1}(\Delta x - C\Delta t) + \phi_{i-1}, \end{aligned} \quad (\text{A14a})$$

$$\begin{aligned} \phi_i'^{n+1} &= dF_i(x_i - C\Delta t) / dx \\ &= 3a_i(\Delta x - C\Delta t)^2 \\ &\quad + 2b_i(\Delta x - C\Delta t) + \phi'_{i-1}. \end{aligned} \quad (\text{A14b})$$

To apply this method for the nonlinear equations such as Eq. (A6), we will use the following method. Let us consider a hyperbolic equation:

$$\frac{\partial \phi}{\partial t} + C \frac{\partial \phi}{\partial x} = \psi, \quad (\text{A15})$$

where the source term ψ is a function of t and x . We split this equation into two phases; the nonadvection phase $\partial \phi / \partial t = \psi$ and the advection phase $\partial \phi / \partial t + C \partial \phi / \partial x = 0$.

The equation for the nonadvection phase is simply solved with a finite difference:

$$\phi_i^* = \phi_i + \Delta t \psi_i, \quad (\text{A16})$$

where $*$ on ϕ means the time after one time step in the nonadvection phase. If ψ includes some diffusion terms, we can apply an implicit method for solving the equation for the nonadvection phase. The spatial derivative ϕ'^* can be estimated as²²

$$\phi_i'^* = \phi'_i + \frac{\phi_{i+1}^* - \phi_{i-1}^* - \phi_{i+1} + \phi_{i-1}}{2\Delta x}. \quad (\text{A17})$$

After ϕ and ϕ' are advanced in the nonadvection phase, the CIP solver is applied to the advection phase. The interpolated profile is determined from Eqs. (A11) to (A14b) using ϕ^* and ϕ'^* .

c. Discretization

A staggered mesh is used to discretize Eqs. (A5)–(A7) and the equation of state (10). With Δy as the mesh size, we define

$$y_i = i \times \Delta y, \quad (i = 0, 1, 2, \dots, N), \quad (\text{A18a})$$

$$y_{i+1/2} = (i + 1/2) \times \Delta y, \quad (i = -1, 0, 1, \dots, N), \quad (\text{A18b})$$

where N satisfies $N\Delta y = 1$. The density ρ_g , pressure p_g and temperature T_g are defined at the center of the computational cell $y_{i+1/2}$, while the velocity u_g is defined at the cell-surface y_i .

The conservation equation for the mass Eq. (A5) is discretized by a finite-volume method:

$$\begin{aligned} & \frac{1}{3}[(r_{i+1}^{n+1/2})^3 - (r_i^{n+1/2})^3] \rho_{g,i+1/2}^{n+1/2} \\ &= \frac{1}{3}[(r_{i+1}^n)^3 - (r_i^n)^3] \rho_{g,i+1/2}^n - \frac{1}{2} \Delta t \{ [(r_{i+1}^{n+1/2})^2 \rho_{g,i+1}^n \\ & \quad \times (u_g - r_t)_{i+1}^{n+1/2}] - [(r_i^{n+1/2})^2 \rho_{g,i}^n (u_g - r_t)_i^{n+1/2}] \}, \end{aligned} \quad (\text{A19a})$$

$$\begin{aligned} & \frac{1}{3}[(r_{i+1}^{n+1})^3 - (r_i^{n+1})^3] \rho_{g,i+1/2}^{n+1} \\ &= \frac{1}{3}[(r_{i+1}^n)^3 - (r_i^n)^3] \rho_{g,i+1/2}^n \\ & \quad - \Delta t \{ [(r_{i+1}^{n+1})^2 \rho_{g,i+1}^{n+1/2} (u_g - r_t)_{i+1}^{n+1}] \\ & \quad - [(r_i^{n+1})^2 \rho_{g,i}^{n+1/2} (u_g - r_t)_i^{n+1}] \}, \end{aligned} \quad (\text{A19b})$$

where $\rho_{g,i}$ is determined by $(\rho_{g,i+1/2} + \rho_{g,i-1/2})/2$.

The equation for the velocity u_g Eq. (A6) is discretized by following the CIP method. The nonadvection phase in Eq. (A6) is discretized to obtain $u_{g,i}^{n+1/2}$:

$$u_{g,i}^* = u_{g,i}^n - \frac{\Delta t/2}{\rho_{g,i}^n r_{y,i}^n} \frac{p_{g,i+1/2}^n - p_{g,i-1/2}^n}{\Delta y}. \quad (\text{A20})$$

Then, the CIP method is applied to solve the advection phase in Eq. (A6),

$$\frac{\partial u_g}{\partial t} + \frac{(u_g^n - r_t^n)}{r_y^n} \frac{\partial u_g}{\partial y} = 0. \quad (\text{A21})$$

We apply a similar procedure to obtain $u_{g,i}^{n+1}$ by using the quantities at the $(n+1/2)$ -th time level.

The equation for the temperature T_g Eq. (A7) is also discretized by following the CIP method. Since Eq. (A7) has a diffusion term, we discretize the nonadvection phase in Eq. (A7) in an implicit form:

$$\begin{aligned} T_{g,i+1/2}^* &= T_{g,i+1/2}^n - \frac{\Delta t}{2} \frac{(\gamma-1) T_{g,i+1/2}^*}{(r_{i+1/2}^{n+1/2})^2 r_{y,i+1/2}^{n+1/2}} \\ & \quad \times \frac{(r_{i+1}^{n+1/2})^2 u_{g,i+1}^{n+1/2} - (r_i^{n+1/2})^2 u_{g,i}^{n+1/2}}{\Delta y} \\ & \quad + \frac{\Delta t/2}{\rho_{g,i+1/2}^{n+1/2} c_{vg}} \frac{1}{(r_{i+1/2}^{n+1/2})^2 r_{y,i+1/2}^{n+1/2} \Delta y} \\ & \quad \times \left[(r_{i+1}^{n+1/2})^2 \frac{\lambda_{g,i+1}^n}{r_{y,i+1}^{n+1/2}} \frac{T_{g,i+3/2}^* - T_{g,i+1/2}^*}{\Delta y} \right. \\ & \quad \left. - (r_i^{n+1/2})^2 \frac{\lambda_{g,i}^n}{r_{y,i}^{n+1/2}} \frac{T_{g,i+1/2}^* - T_{g,i-1/2}^*}{\Delta y} \right]. \end{aligned} \quad (\text{A22})$$

We apply the Thomas algorithm for solving a tri-diagonal system of the discretized equation (A22). Then, we obtain $T_{g,i}^{n+1/2}$ by solving the following equation by the CIP method:

$$\frac{\partial T_g}{\partial t} + \frac{(u_g^{n+1/2} - r_t^{n+1/2})}{r_y^{n+1/2}} \frac{\partial T_g}{\partial y} = 0. \quad (\text{A23})$$

The pressure p_g is determined by using the equation of state (10) as

$$p_{g,i+1/2}^{n+1/2} = \rho_{g,i+1/2}^{n+1/2} \mathfrak{R} T_{g,i+1/2}^{n+1/2}. \quad (\text{A24})$$

3. Numerical procedure

The following procedures are carried out step by step.

- (1) The bubble radius $R^{n+1/2}$ and the velocity at the bubble wall $dR/dt^{n+1/2}$ are obtained from Eqs. (A1a) and (A1b). The time step Δt is determined to satisfy a condition that the Courant number $[(|u_g| + C_g) \Delta t / (r_y \Delta y)]$ of all the grid point is less than unity. Here C_g denotes the speed of sound in the gas.
- (2) The velocity $u_g^{n+1/2}$ is obtained by solving Eqs. (A20) and (A21). The velocity at the center of the bubble $u_{g,0}^{n+1/2}$ is set to be zero, while the velocity at the bubble wall $u_{g,N}^{n+1/2}$ is set to be $dR/dt^{n+1/2}$.
- (3) The density $\rho_g^{n+1/2}$ is obtained from Eq. (A19a).
- (4) The temperature $T_g^{n+1/2}$ is obtained by solving Eqs. (A22) and (A23). To satisfy the boundary condition [Eqs. (11) and (12)], $T_{g,-1/2}^{n+1/2}$ and $T_{g,N+1/2}^{n+1/2}$ are determined as

$$T_{g,-1/2}^{n+1/2} = T_{g,1/2}^{n+1/2}, \quad T_{g,N+1/2}^{n+1/2} = 2T_l - T_{g,N-1/2}^{n+1/2}.$$
- (1) The pressure $p_g^{n+1/2}$ is obtained from Eq. (A24). The pressure at the bubble wall $p_{gw}^{n+1/2} (= p_{g,N})$ is determined as it satisfies Eq. (A2) with the quantities at $(n+1/2)$ th time level and the equation for the gas at the bubble wall:

$$\frac{d^2 R^{n+1/2}}{dt^2} = - \frac{1}{\rho_{g,N}^{n+1/2}} \frac{p_{gw}^{n+1/2} - p_{g,N-1/2}^{n+1/2}}{r_N^{n+1/2} - r_{N-1/2}^{n+1/2}}, \quad (\text{A25})$$
 where the density $\rho_{g,N} = p_{gw} / (\mathfrak{R} T_l)$.
- (2) The bubble radius R^{n+1} and the wall velocity dR/dt^{n+1} are obtained from Eqs. (A1c) and (A1d).
- (3) The velocity u_g^{n+1} , density ρ_g^{n+1} , temperature T_g^{n+1} , and pressure p_g^{n+1} are obtained by carrying out the similar procedures mentioned above.

We have chosen values of N between 100 and 200 to check the accuracy of calculation. As a result, we confirm that differences of R_{\min} are less than $0.05 \mu\text{m}$ for $R_0 = 0.15 \text{ mm}$ with $p_{la} = 40 \text{ kPa}$. Thus we have set the number of grids to be 100 for all the calculations appearing in this manuscript. In addition, we find that the total mass of the gas contained in the bubble remains a constant value. The error of the total mass during the calculation is less than 10^{-13} to its initial value.

¹M. S. Plesset and A. Prosperetti, "Bubble dynamics and cavitation," *Annu. Rev. Fluid Mech.* **9**, 145–185 (1977).

²A. Prosperetti, "Bubble phenomena in sound fields," *Ultrasonics* **22**, 69–77, 115–124 (1984).

³Z. C. Feng and L. G. Leal, "Nonlinear bubble dynamics," *Annu. Rev. Fluid Mech.* **29**, 201–243 (1997).

⁴J. B. Keller and I. I. Kolodner, "Damping of underwater explosion bubble oscillation," *J. Appl. Phys.* **27**, 1152–1161 (1956).

⁵J. B. Keller and M. Miksis, "Bubble oscillation in large amplitude," *J. Acoust. Soc. Am.* **68**, 628–633 (1980).

⁶A. Prosperetti and A. Lezzi, "Bubble motion in a compressible liquid. Part I. First-order theory," *J. Fluid Mech.* **168**, 457–478 (1986).

⁷S. Fujikawa and T. Akamatsu, "Effects of the nonequilibrium condensation of vapour on the pressure wave produced by the collapse of a bubble in a liquid," *J. Fluid Mech.* **97**, 481–512 (1980).

⁸W. Lauterborn, "Numerical investigation of nonlinear oscillations of gas bubbles in liquids," *J. Acoust. Soc. Am.* **59**, 283–293 (1976).

⁹U. Parlitz, V. Englisch, C. Scheffczyk, and W. Lauterborn, "Bifurcation

- structure of bubble oscillators," J. Acoust. Soc. Am. **88**, 1061–1077 (1990).
- ¹⁰A. 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).
- ¹¹A. Prosperetti, "The thermal behaviour of oscillating gas bubbles," J. Fluid Mech. **222**, 587–616 (1991).
- ¹²R. I. Nigmatulin, N. S. Khabeev, and F. B. Nagiev, "Dynamics, heat and mass transfer of vapour-gas bubbles in a liquid," Int. J. Heat Mass Transf. **24**, 1033–1044 (1981).
- ¹³A. Prosperetti, L. A. Crum, and K. W. Commander, "Nonlinear bubble dynamics," J. Acoust. Soc. Am. **83**, 502–514 (1988).
- ¹⁴Y. Matsumoto and F. Takemura, "Influence of internal phenomena on gas bubble motion (Effects of thermal diffusion, phase change on the gas-liquid interface and mass diffusion between vapor and noncondensable gas in the collapsing phase)," JSME International Journal Series B **37**, 288–296 (1994).
- ¹⁵F. Takemura and Y. Matsumoto, "Influence of internal phenomena on gas bubble motion (Effects of transport phenomena and mist formation inside bubble in the expanding phase)," JSME International Journal Series B **37**, 736–743 (1994).
- ¹⁶V. Kamath and A. Prosperetti, "Numerical integration method for gas bubble motion," J. Acoust. Soc. Am. **85**, 1538–1548 (1989).
- ¹⁷R. G. Holt and L. A. Crum, "Acoustically forced oscillations of air bubbles in water: Experimental results," J. Acoust. Soc. Am. **91**, 1924–1932 (1992).
- ¹⁸L. A. Crum, "Measurement of the growth of air bubbles by rectified diffusion," J. Acoust. Soc. Am. **68**, 203–211 (1980).
- ¹⁹F. Takemura, "Internal phenomena of gas bubble motion," Ph.D. dissertation, The University of Tokyo, Tokyo, Japan (1993) (in Japanese).
- ²⁰A. Prosperetti, "Viscous effects on perturbed spherical flows," Q. Appl. Math. **34**, 339–352 (1977).
- ²¹V. Kamath and A. Prosperetti, "Mass transfer during bubble oscillation," in *Frontiers of Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Elsevier, London, 1990), pp. 503–508.
- ²²T. Yabe and E. Takei, "A new higher Godunov method for general hyperbolic equations," J. Phys. Soc. Jpn. **57**, 2598–2601 (1988).
- ²³Shin-etsu Chemical Co. (Eds.), *A technical note—Performance test results of KF96* (Shin-etsu Chemical, Tokyo, 1991) (in Japanese).
- ²⁴R. C. Reid, J. M. Prausnitz, and T. K. Sherwood, *The Properties of Gases and Liquids*, 3rd ed. (McGraw-Hill, New York, 1977).

Two-dimensional effects in the edge sound of vortices and dipoles

G. Schouten^{a)}

Department of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands

(Received July 7 1998; revised July 7 1999; accepted 29 July 1999)

Edge sound consists of the pressure waves generated by the fast modification of the local velocity field when a flow inhomogeneity passes near an edge. The modifications take place in the near field; they locally allow a description in terms of incompressible flow. The incompressible disturbances are referred to as pseudosound. The disturbances progress into the far field as diffraction waves that, in 3D space, dominate the pseudosound. In 3D space the diffraction waves have a 2D character in their time history; they behave as the halfth time derivative of the pseudosound. This latter effect is not manifest in the analogous, purely 2D diffraction problem where the diffracted wave in the far field has the form of retarded pseudosound. In this paper, an answer is presented to the intriguing question: what is the rationale behind this difference in behavior? The analysis is based on time integral representations having as a kernel the closed form of Green's functions for half-plane problems in 2D and in 3D space. The Green's functions are presented in this paper. The integral representations inherently represent the relations between wave-sound and incompressible pseudosound. After performing an integration by parts, appropriate approximations reveal the incompressible near-field behavior or the far-field wave behavior. Examples are given of the edge sound of moving vortices and dipoles in 2D and in 3D space. In free 2D space the potential field of a vortex is proportional to the logarithm of the distance; the field of a dipole is proportional to the inverse of the distance. However, the edge sound pressure of vortices and of dipoles does not show a difference in far field behavior. It falls off with the square root of distance in 2D space and linear with distance in 3D space. © 1999 Acoustical Society of America. [S0001-4966(99)06311-0]

PACS numbers: 43.28.Ra, 43.50.Nm [LCS]

INTRODUCTION

Edge sound consists, in an elementary form, of the pressure disturbances that are generated by the relatively fast modification of the velocity field of vortices or dipoles when they move in the vicinity of a sharp edge in a medium that is at rest at infinity. One may think of turbulence as to consist of a time dependent distribution of moving vortices or dipoles. Whereas the general theory of sound generated aerodynamically has been founded by Lighthill (1952), (1954), and (1982), the importance of edge sound has been brought forward by Curle (1955), Powell (1964) and many other authors. Ffowcs Williams and Hall (1970) consider the production of dipole edge sound by turbulence. For the description of the resulting sound field they use the far field approximation of the integral representation of Green's function for the half-plane found by Macdonald (1915). Following Howe (1975), Kambe *et al.* (1985) use in their analysis a low frequency approximation to the Fourier transform of the integral representation of Macdonald (1915). They find a deformation of the edge sound signal by the halfth time derivative (see Lighthill, 1978) that is characteristic for wave propagation in 2D space. Crighton (1972), in his analysis of the 2D edge sound of a moving vortex, using an outer wave solution that is matched to the incompressible inner solution, does not use any Green's function representation. The wave he finds in his Eq. (3.6) is just the retarded incompressible pseudo-

sound [the term 'pseudosound' is defined in Ribner's (1962) report]. Intuitively one would have expected a deformation of the signal caused by the typically 2D halfth time derivative, a deformation that is manifest in the diffraction on an edge in 3D space.

For an estimation of the magnitude of the sound field, the above treatments will do. The intriguing question that remained unanswered is why, in three-dimensional space, the diffracted field is the halfth time derivative of the retarded incompressible diffracted field and in two-dimensional space it is just the retarded incompressible diffracted field. Especially in 2D space one would expect the wave solution to be more complicated than just a retarded incompressible field.

From the fluid dynamic point of view, the mechanism generating the sound field is to be understood in terms of disturbances in the incompressible flow field. When a singularity (i.e., a vortex or a dipole) moves relatively slowly in the vicinity of an edge, the flow field needs modification to accommodate the edge. One might think that the modification is more severe in the field of the dipole than in the field of the vortex as the dipole has a more singular near field. In free 2D space the potential field of a vortex is independent of the distance; the field of a dipole is proportional to the inverse of the distance. However, the edge sound of vortices and of dipoles does not show a difference in far field behavior. It falls off with the square root of the distance.

Considering dipole sound, generated by a periodic dipole in the origin of 3D free space, the pressure in the far

^{a)}Electronic mail: G.Schouten@LR.TUDELFT.NL

field of the wave is a factor R/c stronger than the retarded incompressible pressure, where R is the radial distance in 3D space and c is the speed of sound. In 2D free space the pressure in the wave field is proportional to the retarded halfth time derivative (Lighthill, 1978; see the Appendix) of the incompressible pressure multiplied by the factor $(r/c)^{1/2}$, r being the radial distance in 2D space.

The far field of the pressure wave (the edge sound) that is generated when a dipole of constant strength crosses the edge of a half-plane in 3D space is proportional to the halfth time derivative of the retarded incompressible diffracted pressure field amplified by the factor $(r/c)^{1/2}$. This behavior is similar to a far field wave response in unbounded 2D space. The diffracted wave field in 2D space, however, does not show this typically two-dimensional behavior. There the dominating part has the form of the retarded incompressible diffracted field.

It is our aim to discuss these differences in the far field behavior of edge sound in 3D and in 2D space and to relate them to the incompressible fields. The discussion is based on integral representations having Green's functions as a kernel. In the limit $c = \infty$ they must yield the incompressible field, i.e. the near field in a fluid with finite c . A problem in establishing the relation is that in the limit of infinitely large velocity of sound the Green's function for the 2D wave equation does not simply degenerate with $c \rightarrow \infty$ into Green's function for the incompressible 2D Laplace equation. In the analysis presented in this paper it is shown that, after performing an integration by parts, the resulting general expressions are equally amenable to approximation of the incompressible near field behavior as they are to approximation of the far field wave behavior.

I. GENERAL FORMULATIONS

A. Model and outline

The sound field to be considered consists of the pressure wave that is generated by the modification of the velocity field of vortices or dipoles when they move in the vicinity of the edge of a half-plane in a medium that is at rest at infinity. The pressure disturbance p is approximated by the local time derivative of the flow potential Φ , $p = -\rho_0(\partial\Phi/\partial t)$, where ρ_0 is the density of the fluid at rest. The potential Φ of the inviscid flow field is governed by the wave equation (1),

$$\left\{ \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right\} \Phi = q(x, y, z, t), \quad (1)$$

with the right-hand side representing an appropriate singularity distribution $q(x, y, z, t)$. The solution to this linear inhomogeneous equation, satisfying sound-hard boundary conditions on a half-plane, is written as a distribution integral of the appropriate Green's functions,

$$\Phi = \int q(\xi, \eta, \zeta, \tau) G(x, y, z, t; \xi, \eta, \zeta, \tau) d\xi d\eta d\zeta d\tau. \quad (2)$$

An appropriate Green's function, $G(x, y, z, t; x_0, y_0, z_0, t_0)$, satisfies the inhomogeneous equation that has a pulse-function, i.e., $\delta(x-x_0)\delta(y-y_0)\delta(z-z_0)\delta(t-t_0)$, on the right-hand side; it also satisfies the sound-hard boundary

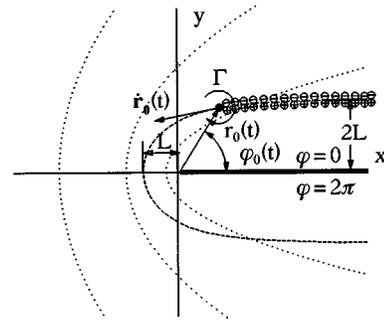


FIG. 1. Trajectory of a free vortex Γ moving around a half-plane. The nearest distance is L , far from the edge the vortex moves parallel to the half-plane at a distance $2L$. The potential of the vortex is represented by the potential of a dipole distribution of intensity Γ/m^2 on the trajectory.

conditions. The physical meaning of Green's function is that it is a pulse response; it represents the potential Φ , in the observation point (x, y, z) at the time t , of the flow field that emerges when a unit of volume is delivered instantaneously at the instant $t = t_0$ in a source point (x_0, y_0, z_0) .

The potential of a vortex of strength Γ is described in 3D space by a dipole distribution of density Γ per unit of area, normal to the surface of an arbitrary area bounded by the vortex. For a single line vortex moving in 2D space we situate the growing dipole distribution on the trajectory of the vortex (see Fig. 1). A sound-hard diffracting half-plane extends from $x=0$ to $x=\infty$. The two sides are identified in circular cylinder- or in spherical coordinates by $\varphi=0$ and $\varphi=2\pi$.

The edge-sound fields generated by the moving dipoles or vortices are expressed in terms of integrals using the known closed forms of the Green's functions for half-plane problems in inviscid, incompressible or compressible fluids in 2D or 3D spaces. The exact integrals are, after performing an integration by parts, readily amenable to approximation of the incompressible near field behavior as well as to approximation of the far field wave behavior.

In Sec. IB the pressure integrals are formulated and in Sec. IC the expressions for the Green's functions are called to mind. Integration by parts of the formal integrals, performed and discussed in Sec. ID, then reveals the near field and the far field behavior of the edge sound. In Secs. II and III, examples of edge sound pressure in some elementary cases in 2D and 3D space are presented and discussed. In Sec. IV a general discussion follows with conclusions.

B. Edge sound integrals

1. A vortex moving in 2D space

The pressure p in the field of the moving vortex is described as the local time derivative of the unsteady potential field $\Phi(x, y, t)$. The potential field of the moving vortex is described as the potential of a distribution of dipoles of intensity Γ normal to the trajectory $(r_0(t), \varphi_0(t))$ of the vortex (see Fig. 1). The distribution represents a jump of strength Γ in the potential field. The dipoles come to life locally on the trajectory as a step function in time from the instant that the vortex has passed a point $\mathbf{r}_0(\tau)$ of its trajectory:

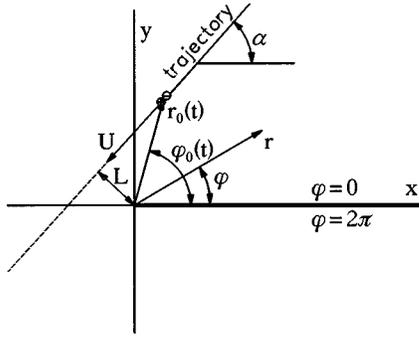


FIG. 2. Straight line trajectory crossing the sharp edge of a half-plane at a distance L .

$$\begin{aligned}
 p &= -\varrho_0 \frac{\partial \Phi}{\partial t} \\
 &= -\varrho_0 \frac{\partial}{\partial t} \int_{-\infty}^t d\tau \int \int d\varrho' \dot{\mathbf{r}}_0(\tau) \\
 &\quad \times \mathbf{\Gamma} \cdot \nabla_0 \delta(\varrho' - \mathbf{r}_0(\tau)) \left\{ \int_{\tau}^t G(\mathbf{r}, t; \varrho', \tau') d\tau' \right\} \\
 &= -\varrho_0 \int_{-\infty}^{t^*} d\tau \dot{\mathbf{r}}_0(\tau) \times \mathbf{\Gamma} \cdot \nabla_0 G(r, \varphi, t; r_0(\tau), \varphi_0(\tau), \tau). \quad (3)
 \end{aligned}$$

As the Green's function consists of different terms the upper boundary must be adapted in each term. The upper boundary for the primary wave refers to the primary front with the retarded time $t^* = t - \tilde{r}(t^*)/c$, where $\tilde{r}(t^*) = (r^2 + r_0^2(t^*) - 2rr_0(t^*)\cos(\varphi - \varphi_0(t^*)))^{1/2}$. For the diffracted wave the upper boundary is the retarded time $t^0 = t - (r + r_0(t^0))/c$, referring to the diffraction front. In the third line of (3), the Green's function is present in a step function integral beginning at τ , the instant that the vortex has passed the point $\varrho' = r_0(\tau)$ on its trajectory and a dipole with the moment $\dot{\mathbf{r}}_0(\tau) \times \mathbf{\Gamma} d\tau$ comes to life.

2. A dipole moving in a 2D or 3D space

The considered dipoles are moving with a constant velocity U on a straight line trajectory in the x, y plane perpendicular to the edge that is directed along the z axis (see Fig. 2). The potential field of a moving dipole of constant intensity \mathbf{m} is simpler than that of the moving vortex in that the dipole does not leave a trail in the form of a discontinuity in the potential field in space. In a compressible fluid there is a transient "wake" in the field behind the fronts that are described by a square root singularity (as will become apparent in the expressions for the Green's functions further on). The dipoles that momentarily come to life on the trajectory leave these "wakes" in the field. Another difference with the field of a moving vortex is that a moving dipole has its axis along the trajectory whereas the moving vortex is modeled with dipoles perpendicular to the trajectory. The time derivative of the integral turns the moving dipole into a longitudinal quadrupole for the pressure field function:

$$\begin{aligned}
 p &= -\varrho_0 \frac{\partial \Phi}{\partial t} \\
 &= -\varrho_0 \frac{\partial}{\partial t} \int_{-\infty}^t d\tau \int \int d\varrho' \mathbf{m} \cdot \nabla_0 \delta(\varrho' - \mathbf{r}_0(\tau)) \\
 &\quad \times G(\mathbf{r}, t; \varrho', \tau) \\
 &= -\varrho_0 \frac{\partial}{\partial t} \int_{-\infty}^{t^*} d\tau \mathbf{m} \cdot \nabla_0 G(r, \varphi, t; r_0(\tau), \varphi_0(\tau), \tau). \quad (4)
 \end{aligned}$$

C. Green's functions for the half-plane

1. Green's function in 3D space

The Green's function for the half-plane in incompressible flow in three-dimensional space was constructed in closed form by Sommerfeld in 1897 [the correction to this paper, Carslaw (1898–1899), only refers to the Green's function for the strip]:

$$\begin{aligned}
 G_{3D \text{ Laplace}} &= \delta(t - t_0) \\
 &\times \left\{ \frac{1}{4\pi R} \left(-1 + \frac{1}{\pi} \tan^{-1} \frac{R}{2\sqrt{rr_0} \cos((\varphi - \varphi_0)/2)} \right) \right. \\
 &\quad \left. + \frac{1}{4\pi R'} \left(-1 + \frac{1}{\pi} \tan^{-1} \frac{R'}{2\sqrt{rr_0} \cos((\varphi + \varphi_0)/2)} \right) \right\}, \quad (5)
 \end{aligned}$$

where (r, φ) are circular cylinder coordinates in the $x-y$ plane, and (r_0, φ_0) represents the source point; $R = (r^2 + r_0^2 - 2rr_0 \cos(\varphi - \varphi_0) + (z - z_0)^2)^{1/2}$; and $R' = (r^2 + r_0^2 - 2rr_0 \cos(\varphi + \varphi_0) + (z - z_0)^2)^{1/2}$. The z coordinate is directed along the edge. The distance from observer to the image source point (at $\varphi = 4\pi - \varphi_0$) is referred to with a prime as R' .

The Green's function for a half-plane in the field of the wave equation was first formulated in terms of eigenfunctions by Macdonald (1915). It was Cagniard (1935) who found the closed form solution to the diffraction of a step function wave on a half-plane by the method of Laplace transforms. The time dependent Green's function (6) is simply found by differentiation of Cagniard's solution:

$$\begin{aligned}
 G_{3D \text{ wave}} &= -\frac{\delta(t - t_0 - R/c)}{4\pi R} \\
 &- \frac{1}{4\pi^2} \frac{2\sqrt{rr_0} \cos((\varphi - \varphi_0)/2)}{\sqrt{(t - t_0)^2 - (D/c)^2} (c^2(t - t_0)^2 - R^2)} \\
 &- \frac{1}{4\pi^2} \frac{2\sqrt{rr_0} \cos((\varphi + \varphi_0)/2)}{\sqrt{(t - t_0)^2 - (D/c)^2} (c^2(t - t_0)^2 - R'^2)}, \quad (6)
 \end{aligned}$$

where the distance $D = [(r + r_0)^2 + (z - z_0)^2]^{1/2} = [R^2 + 4rr_0 \cos^2((\varphi - \varphi_0)/2)]^{1/2}$ is the distance from source point to observation point via the edge. The zero of the square root in the denominator of (6) describes a spindle shaped diffraction front, centered around the edge and touching the spherical primary front, marked by the delta function. The velocity field inside the diffraction front, in the plane $z = z_0$, is shown

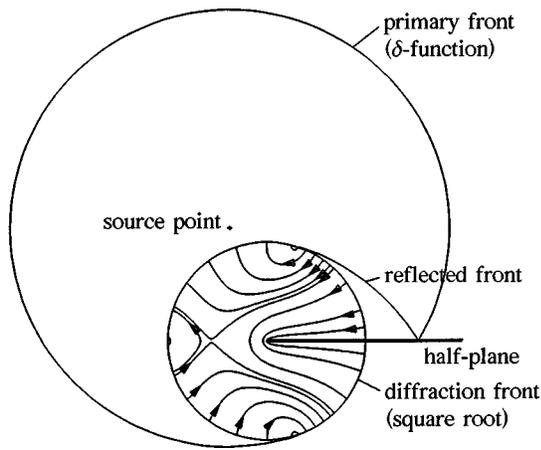


FIG. 3. Fronts and streamlines in the plane of symmetry of Green's function for the 3D wave equation.

in Fig. 3. The expression (6) must be read with the condition that on the "illuminated" side of the half-plane, a reflected wave must be added in the region $0 < \varphi < \pi - \varphi_0$ and that on the "shadow" side the primary wave must be omitted in the region $\varphi_0 + \pi < \varphi < 2\pi$.

More general expressions for Green's functions for wedge problems, i.e., where the half-"plane" has two sides at a finite wedge angle, are to be found in the books of van de Pol and Bremmer (1964) and of Felsen and Marcuvitz (1973).

2. Green's function in 2D space

The Green's function for the two-dimensional Laplace equation, satisfying sound-hard boundary conditions on the half-plane $0 < x$ (or $\varphi=0$ and $\varphi=2\pi$), is simply found from potential theory by conformal mapping (see Milne-Thomson, 1968):

$$G_{2D \text{ Laplace}} = \delta(t-t_0) \left\{ \frac{1}{4\pi} \log(r^2 + r_0^2 - 2rr_0 \cos(\varphi - \varphi_0))^{1/2} - \frac{1}{4\pi} \log \left(r + r_0 + 2\sqrt{rr_0} \cos\left(\frac{\varphi - \varphi_0}{2}\right) \right) + \frac{1}{4\pi} \log \left(r + r_0 - 2\sqrt{rr_0} \cos\left(\frac{\varphi + \varphi_0}{2}\right) \right) \right\}. \quad (7)$$

The formulation of the Green's function for the two-dimensional wave equation has taken some time, seen in the historical perspective of the three-dimensional developments. Turner (1956) made a first step. In later publications, Friedlander (1958), Lauwerier (1962), and de Jager (1964) to mention some, the Green's function evolved into the erroneous form

$$G_{2D \text{ wave}} = -\frac{1}{2\pi} \frac{1}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}} + \frac{1}{4\pi} \frac{H(t-t_0 - (r+r_0)/c)}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}} - \frac{1}{4\pi} \frac{H(t-t_0 - (r+r_0)/2)}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}}. \quad (8)$$

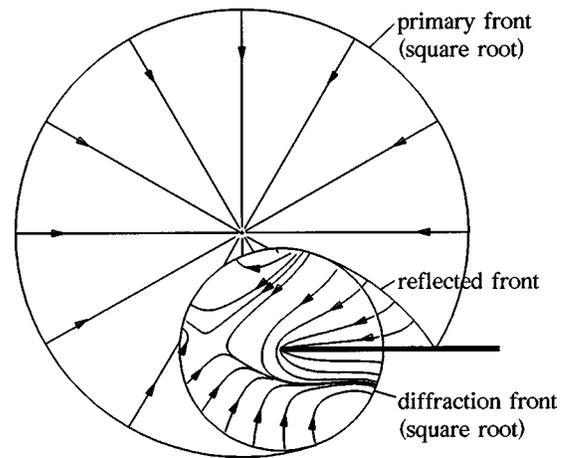


FIG. 4. Fronts and streamlines of Green's function for the 2D wave equation.

In this form $H(t-t_0 - (r+r_0)/c)$ is Heaviside's unit step function describing the boundary of the diffraction circle. The form (8) has always been suspect from the fluid dynamic point of view as it does not show the expected diffraction effect at the edge of the half-plane, i.e., the singular flow around the edge is not taken care of by (8).

The correct expression for this Green's function is obtained in a direct way by integration over z_0 of the 3D function (6). The integration boundaries must be carefully selected in order that the observation point lies inside or on the spindle-shaped diffraction front of the contributing pulses, i.e., $(z-z_0)^2 < c^2(t-t_0)^2 - (r+r_0)^2$. The correct expression for the Green's function is

$$G_{2D \text{ wave}} = -\frac{1}{2\pi} \frac{1}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}} + \frac{1}{4\pi^2} \frac{1}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}} \times \sin^{-1} \frac{\sqrt{(t-t_0)^2 - (r+r_0)^2/c^2}}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}} - \frac{1}{4\pi^2} \frac{1}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}} \times \sin^{-1} \frac{\sqrt{(t-t_0)^2 - (r+r_0)^2/c^2}}{\sqrt{(t-t_0)^2 - (\tilde{r}/c)^2}}. \quad (9)$$

A peculiarity of the Green's function for the wave equation in unbounded 2D space is that it does not have a regular transition to that of Laplace's equation governing incompressible flow ($c=\infty$). This is due to the explicit combination of t and \tilde{r}/c under the square root sign. The undisturbed field is essentially a far field in the whole of 2D space. The incompressible logarithmic near field behavior around the source point comes forward in a limit procedure after integration of the pulse response to form a transient time response. Expansion of the diffracted field terms in (9) around the edge shows the correct square root singularity in the ap-

pearance of a factor $r^{1/2}$. So in the velocity field inside the diffraction circle the singular flow around the sharp edge is manifest. The streamline pattern inside the diffraction circle is shown in Fig. 4.

D. Integration by parts yields the 2D effects in the far field

The formal pressure integrals (3) and (4) can be made to show the near field and the far field behavior when they are integrated by parts. Introducing the appropriate Green's function into (3) and (4), both integrals, omitting the factors that are not essential for the analysis, reduce to the simplified form

$$I = \int_{-\infty}^{t^*} \mathbf{m}(\tau) \cdot \nabla_0 g(r, \varphi, t; r_0(\tau), \varphi_0(\tau), \tau) d\tau. \quad (10)$$

In 2D space g has the form $1/\sqrt{(t-\tau)^2 - (\tilde{r}(\tau)/c)^2}$, the square root singularity at the zero of the root marks the primary wave front. The front of the diffracted wave is marked by the zero of $\sqrt{(t-\tau)^2 - (r+r_0)^2/c^2}$. In 3D space, g has the analogous form $1/\sqrt{(t-\tau)^2 - (D/c)^2}$, and the singularity is on the spindle-shaped diffraction front where $D=c(t-\tau)$. The square root forms are representative of the 2D aspects in the Green's functions (6) and (9), as explained in more detail in Lighthill (1978) and in the Appendix. The dipole intensity $\mathbf{m}(\tau)$ in (10) represents either the time dependent factor $\dot{\mathbf{r}}_0(\tau) \times \mathbf{\Gamma}$ in Eq. (3) or the constant \mathbf{m} in Eq. (4).

We first consider the 2D case. Having in mind the Green's function (9) the integration boundary in (10) may refer to the primary front as well as to the diffraction front. Performing the $\{\mathbf{m} \cdot \nabla_0\}$ operation before the integration by parts in the 2D case then yields the terms

$$I = I_1 + I_2 = \frac{\mathbf{m} \cdot \nabla_0 \tilde{r}(\tau)}{\tilde{r}(\tau)} \frac{t-\tau}{\sqrt{(t-\tau)^2 - (\tilde{r}(\tau)/c)^2}} \Big|_{-\infty}^{t^*} - \int_{-\infty}^{t^*} \frac{\partial}{\partial \tau_i} \left(\frac{\mathbf{m} \cdot \nabla_0 \tilde{r}(\tau_i)}{\tilde{r}(\tau_i)} \frac{t-\tau}{\sqrt{(t-\tau)^2 - (\tilde{r}(\tau_i)/c)^2}} \right) d\tau. \quad (11)$$

The first term comes from the integration of the explicit function of τ under the square root. The contribution at the lower boundary vanishes as at $\tau = -\infty$, the distance $\tilde{r}(\tau)$ is supposed to be infinitely large. The contribution of the first term depends on whether the upper boundary of the integration in Eq. (10) marks the primary front or the diffraction front. When it refers to the primary front $t^* = (t - \tilde{r}(t^*)/c)$ the contribution is singular and must be discarded. (see Hadamard, 1952). When the upper boundary refers to the diffracted field and the retarded time is not t^* but $t^0 = t - (r + r_0(t))/c$ the contribution is finite. In combination with the other factors and terms of Green's function it will yield the retarded incompressible diffracted field. The finite contribution I_1 , referring to t^0 at the diffraction front, has the form

$$I_1 \approx \frac{\mathbf{m} \cdot \nabla_0 \tilde{r}^0}{2\pi \tilde{r}^0} \frac{r+r_0^0}{2\sqrt{rr_0^0} \cos((\varphi - \varphi_0^0)/2)}. \quad (12)$$

In the near field, the retardation is negligible and the form represents the time dependent incompressible diffracted field. As in the following we will consider the response in the far field, the difference between the retarded times t^* and t^0 is no longer important and we will refer the approximate form of I_1 to the retarded time $t^* \approx t - r/c$.

The second term of (11) is the integral of the time derivative operating on $\tilde{r}(\tau)$ indicated by $\tilde{r}(\tau_i)$. [This time derivative is the nonlocal derivative ($\dot{\mathbf{r}}_0 \cdot \nabla_0$).] Expanding the integrand in terms of $(r_0/r)^{1/2}$ and retaining the dominant term to prepare for a far field approximation of the integral then gives, for the second term I_2 in Eq. (11),

$$I_2 \approx \frac{1}{2\pi} \int_{-\infty}^{t-r/c} \dot{\mathbf{r}}_0 \cdot \nabla_0 \left\{ \frac{\mathbf{m} \cdot \nabla_0 \tilde{r}}{r} \right\} \frac{t-\tau}{\sqrt{(t-\tau)^2 - (r/c)^2}} d\tau. \quad (13)$$

By the expansion the difference between t^* and t^0 has disappeared. Then putting $(t-\tau) \approx r/c$ in the far field the integral gets a form suitable for interpretation as the halfth time derivative of the retarded free-space incompressible field $\mathbf{m} \cdot \nabla_0 \tilde{r}/4\pi\tilde{r}$ and a scaling factor $(\pi r/2c)^{1/2}$ (see the Appendix). The halfth time derivative has the dimension $t^{-1/2}$ and its order of magnitude is estimated to be $O[(|\dot{\mathbf{r}}_0|/r_0)^{1/2}] \approx O[(U/L)^{1/2}]$, where U is the reference velocity of the vortex or the dipole and L is the minimal distance at which it crosses the edge. The resulting field I_2 would dominate the free-space incompressible field by an order $[(Ur/cL)^{1/2}]$. The expression (10) reduces, using (12) and (13), in the far field, to the form

$$I \approx \left[\frac{\mathbf{m} \cdot \nabla_0 (\tilde{r} - r_0)}{8\pi\sqrt{rr_0} \cos((\varphi - \varphi_0)/2)} \right]^* - \left(\frac{\pi r}{2c} \right)^{1/2} \int_{-\infty}^{t^*} \frac{\partial}{\partial \tau} \left\{ \frac{\mathbf{m} \cdot \nabla_0 \tilde{r}}{4\pi r} \right\} \frac{1}{\sqrt{\pi(t^* - \tau)}} d\tau. \quad (14)$$

Estimating the order of magnitude of the terms we find that the integral term I_2 , representing the halfth time derivative, is smaller than I_1 by a factor of the order $O[(\pi/2)U/c]^{1/2}$. (This factor is not necessarily very small. When $U/c = 0.1$ it already attains the value 0.4.) Nevertheless, the integration by parts performed on Eq. (10) shows that the resulting term having the form of the retarded incompressible diffracted field, for small values of U/c , dominates the edge sound field of a vortex or a dipole. It is larger by a factor $(U/c)^{-1/2}$ than the term that gets the shape of the halfth time derivative.

Now consider the 3D case. The square root function g here only occurs in the diffraction terms. The integration by parts covers the explicit presence of τ in the integrand, so it includes the factor $1/(c^2(t-\tau)^2 - R^2)$. The term I_1 resulting from the integration by parts has the form

$$I_1 = \frac{1}{4\pi^2} \times \mathbf{m} \cdot \nabla_0 \left\{ -\frac{1}{R} \tan^{-1} \frac{R\sqrt{1-D^2/(c^2(t-\tau)^2)}}{2\sqrt{rr_0^*} \cos((\varphi-\varphi_0^*)/2)} \right\} \Big|_{-\infty}^{t-D/c} \quad (15)$$

It vanishes in the upper boundary; the contribution from the lower boundary would yield the incompressible diffracted field of a stationary dipole. For a dipole coming from infinity the contribution vanishes. The diffracted field then is represented solely by the remaining integral I_2 . This integral is interpreted in the far field as the halfth time derivative of the incompressible diffracted field amplified by the factor $(R/c)^{1/2}$. In the near field the integral just represents the incompressible diffracted field.

The noteworthy phenomenon is that the 3D diffracted field shows the typical 2D effect of the halfth time derivative whereas the dominating term of the 2D diffracted field does not.

II. EDGE SOUND DUE TO SINGULARITIES MOVING NEAR A SHARP EDGE IN 2D SPACE

A. Examples

1. Vortex moving on a straight line near an edge

Pressure signals in an incompressible field are readily obtained from potential theory (see Acheson, 1992). The associated responses in a compressible fluid then follow simply by application of the above result, i.e., applying a retardation r/c to the diffracted field.

Consider a vortex Γ moving with a constant velocity U along a straight line trajectory (see Fig. 2) at an angle α with the positive x axis passing the edge at a minimum distance L . (Actually the vortex must be forced to follow such a straight line trajectory.) The trajectory is described by

$$\begin{aligned} \varphi_0(t) &= \alpha + \tan^{-1}(L/Ut); \\ r_0(t) &= (L^2 + (Ut)^2)^{1/2} = L/\sin(\varphi_0(t) - \alpha); \\ \dot{\mathbf{r}}_0 &= -U(\cos \alpha, \sin \alpha). \end{aligned} \quad (16)$$

The edge sound pressure signal of a vortex moving on a straight line trajectory (16) near a half-plane is then formally found from Eq. (3) using the compressible Green's function (9). As explained in Sec. ID, in 2D space a good approximation to the compressible edge sound pressure is to be found in the retarded incompressible pressure. The incompressible pressure signal is readily obtained from the pressure integral (3) using the trajectory (16) and the incompressible dipole potential derived from the Green's function (7). Details on incompressible flows are to be found in textbooks on fluid dynamics (e.g., Acheson, 1992; Milne-Thomson, 1968). The pressure, $p(t)$, is proportional to the potential of a dipole at (r_0, φ_0) with moment $U\Gamma$ perpendicular to the trajectory. Taking only the dominant term and applying the required retardation then yields the pressure signal (i.e., the edge sound) in a medium with finite sound velocity c ,

$$\begin{aligned} p &\approx -\varrho_0 \frac{U\Gamma \cos(\varphi/2) \sin((\alpha - \varphi_0^*)/2)}{2\pi\sqrt{rr_0^*}} \\ &= -\varrho_0 \frac{U\Gamma \cos(\varphi/2)}{2\pi r^{1/2} L^{1/2}} \sqrt{\sin(\varphi_0^* - \alpha) \sin((\alpha - \varphi_0^*)/2)}, \end{aligned} \quad (17)$$

where $\varphi_0(t^*)$ is approximated by $\varphi_0^* \approx \varphi_0(t - r/c)$.

2. Vortex moving around a half-plane

A vortex moving force-free around the edge of a half-plane follows a trajectory that is described by

$$\varphi_0(t) = 2 \tan^{-1}(2L/(-Ut)); \quad r_0(t) = L/\sin(\varphi_0(t)/2) \quad (18)$$

(Howe, 1975; Schouten, 1998). The trajectory is plotted in Fig. 1. The velocity $\dot{\mathbf{r}}_0$ of the free vortex is not constant, it has the time dependent components

$$\dot{r}_0 \approx -\frac{U}{2} \cos\left(\frac{\varphi_0(t)}{2}\right); \quad (19)$$

$$r_0 \dot{\varphi}_0 = U \sin\left(\frac{\varphi_0(t)}{2}\right); \quad \text{where } U = \frac{\Gamma}{4\pi L}.$$

The reference velocity U here is the velocity of the vortex when it moves parallel to the half-plane at a great distance from the edge. As explained in Sec. ID, the edge-sound pressure signal of the free vortex is found as the retarded incompressible pressure signal. The incompressible pressure signal is readily obtained from the pressure integral (3) using the trajectory (19) and the incompressible dipole potential derived from the Green's function (7). Retaining only the dominant term and applying the required retardation then yields an expression for the pressure as

$$\begin{aligned} p &\approx -\varrho_0 \frac{\Gamma^2 \cos(\varphi/2)}{32\pi^2 r^{1/2} L^{3/2}} \\ &\times \sin(\varphi_0^*) \sqrt{\sin(\varphi_0^*/2)} \{1 + O[\sqrt{(\pi/2)U/c}]\}. \end{aligned} \quad (20)$$

Introducing the reference velocity U would substitute the factor $\Gamma^2 L^{-3/2}$ by $4\pi U\Gamma L^{-1/2}$.

3. Dipole moving along a straight line near an edge in 2D space

The edge sound pressure signal, in 2D space, of a dipole moving on a straight line is found in a manner analogous to the procedure followed to obtain the edge sound of the vortex (17), but now using the pressure integral (4) instead of (3). The dominant term in the pressure signal gets the form

$$\begin{aligned}
p &\approx -\varrho_0 \frac{\partial}{\partial t} \frac{m \cos(\varphi/2)}{2\pi\sqrt{rr_0^*}} \cos(\alpha - \varphi_0^*/2) \\
&= \varrho_0 \frac{mU \cos(\varphi/2)}{4\pi\sqrt{rr_0^*}^3} \cos(2\alpha - 3\varphi_0^*/2) \\
&= \varrho_0 \frac{mU \cos(\varphi/2)}{4\pi r^{1/2} L^{3/2}} (\sin(\varphi_0^* - \alpha))^{3/2} \cos(2\alpha - 3\varphi_0^*/2).
\end{aligned} \tag{21}$$

The dipole has a more singular local velocity field than the vortex; its near field velocity behaves as $1/\tilde{r}^2$ whereas the vortex has just $1/\tilde{r}$. This might lead to the conjecture that when an edge appears in the field of a moving dipole, the effects would be more severe than when the same happens to the field of a moving vortex. The results, however, do show that when either of the two crosses the edge at a minimum distance L with a velocity U , the effects are comparable when the dipole strength m is $O[\Gamma L]$. The quantity ΓL is half the equivalent dipole moment of the free vortex in the fluid when it moves parallel to the half-plane at a great distance from the edge.

When the dipole moves on a straight line with a constant velocity U , its edge pressure field contains the factor $Umr^{-1/2}L^{-3/2}$ whereas the vortex had the factor $U\Gamma r^{-1/2}L^{-1/2}$. The factors are identical when we substitute the value \mathbf{m} of the dipole by the value ΓL of the vortex.

B. Discussion

The problem of a vortex generating edge sound when it rounds the edge of a half-plane has been treated by Crighton (1972) and Howe (1975). Using different methods, both find the retarded incompressible pressure signal as the wave solution. The procedure of integration by parts, as presented in Sec. 1D, results in the appearance, next to the retarded pseudosound, of the halfth time derivative term working on the primary field. It reveals why Crighton's pressure signal, his Eq. (3.6), which he obtained by matching a wave solution to the near-field Laplace-solution, does not show the halfth time derivative aspect that one would expect to be manifest in a two-dimensional field. The edge sound is dominated by the retarded incompressible diffracted field. But what is far more important is that the two aspects are entirely different. The farfield halfth time derivative part in Eq. (14) cannot be matched to the diffracted near field. It is not clear how the halfth time derivative aspect can come forward anyhow in a matched asymptotic expansion procedure in two-dimensional or in three-dimensional space. In two-dimensional space, it happens to be of slightly minor importance. In three-dimensional space it is the dominant term in the diffracted field, as comes forward in Eq. (24).

Howe, whose results are identical to those of Crighton, makes use of a low-frequency approximation to the Green's function in 2D space. This approximation is nothing but a retarded incompressible function. Making use of that function excludes at forehand the typical two-dimensional behavior of the halfth time derivative. In the derivation above in Eq. (14) it has become clear that the retarded incompressible

part indeed is the representative part of the diffracted field in purely two-dimensional problems. In three-dimensional space it is not.

III. EDGE SOUND GENERATED BY A DIPOLE CROSSING THE EDGE OF A HALF-PLANE IN 3D SPACE

A dipole \mathbf{m} moving with velocity \mathbf{U} , where \mathbf{U} has the direction of \mathbf{m} , represents the motion of a spherical volume $V = m/(3U)$ (see Milne-Thomson, 1968). Edge sound is generated when the volume passes near an edge. Consider a dipole with moment \mathbf{m} , moving with a constant velocity U (in the direction of \mathbf{m}), along a straight line trajectory (16) perpendicular to the edge. After inserting the Green's function, Eq. (6), in the expression (4), consider that part of the potential Φ that contains the diffraction term of the primary field of the Green's function (the image field is treated in an analogous way). It has the form

$$\begin{aligned}
I &= \frac{1}{4\pi^2} \int_{-\infty}^{t^0} (\mathbf{m} \cdot \nabla_0) \\
&\times \frac{2\sqrt{rr_0(\tau)} \cos((\varphi - \varphi_0(\tau))/2)}{\sqrt{(t-\tau)^2 - (D(\tau)/c)^2} (c^2(t-\tau)^2 - R^2(\tau))} d\tau.
\end{aligned} \tag{22}$$

Integrating by parts, before performing the $(\mathbf{m} \cdot \nabla_0)$ operation, yields two terms. One part, coming from the integration of the explicit presence of τ , yields the two-valued \tan^{-1} function that is representative of the incompressible field; its contributions from the boundary at the diffraction front, $\tau = t^0 = t - D^0/c$ and the boundary $\tau = -\infty$, vanish. The other part consists of the integral of the product of this \tan^{-1} function and the nonlocal time derivative, $(\dot{\mathbf{r}} \cdot \nabla_0)$, of the integrand. In the limit $c = \infty$ it yields the incompressible field, i.e., the near field in a fluid with finite c . Expanding the \tan^{-1} in the integral in terms of $(r_0/r)^{1/2}$ to obtain a far field approximation then yields, for the part of the potential under consideration, the expression

$$\begin{aligned}
\Phi &= -\frac{1}{4\pi^2} \int_{-\infty}^{t^0} \{\dot{\mathbf{r}}_0 \cdot \nabla_0\} \\
&\times \left\{ \mathbf{m} \cdot \nabla_0 \frac{(t-\tau)2\sqrt{rr_0} \cos((\varphi - \varphi_0)/2)}{R^2\sqrt{(t-\tau)^2 - R^2/c^2}} \right\} d\tau.
\end{aligned} \tag{23}$$

Applying the far field approximation $(t-\tau) \approx R/c$, the integral gets a form suitable for interpretation as the halfth time derivative of the factor

$$m \cdot \nabla_0 \left\{ \frac{2\sqrt{rr_0(\tau)} \cos((\varphi - \varphi_0(\tau))/2)}{\tau^2 R^2} \sqrt{\frac{\pi R}{2c}} \right\}.$$

This factor is the product of $\sqrt{(\pi/2)R/c}$ and the dominating term in the far field of the corresponding part of the incompressible diffracted field. This part of the potential (23) in the far field thus gets the approximate form

TABLE I. Far-field behavior of the potential Φ and the pressure signal p of a dipole moving on a straight line trajectory.

	Laplace (incompressible field)		Wave (compressible retarded field)	
	Φ	p	Φ	p
Far field 3D				
Free space	r^{-2}	$r^{-2}\left(\frac{U}{r}\right)$	r^{-2}	$r^{-2}\left(\frac{U}{r}\right)$
Diffracted	$r^{-3/2}L^{-1/2}$	$r^{-3/2}L^{-1/2}\left(\frac{U}{L}\right)$	$r^{-3/2}L^{-1/2}\left(\frac{U}{c}\frac{r}{L}\right)^{1/2}$	$r^{-3/2}L^{-1/2}\left(\frac{U}{L}\right)\left(\frac{U}{c}\frac{r}{L}\right)^{1/2}$
Far field 2D				
Free space	r^{-1}	$r^{-1}\left(\frac{U}{r}\right)$	r^{-1}	$r^{-1}\left(\frac{U}{r}\right)$
Diffracted	$r^{-1/2}L^{-1/2}$	$r^{-1/2}L^{-1/2}\left(\frac{U}{L}\right)$	$r^{-1/2}L^{-1/2}$	$r^{-1/2}L^{-1/2}\left(\frac{U}{L}\right)$

$$\begin{aligned} \Phi &\approx \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \frac{\partial}{\partial \tau} \left\{ \frac{\mathbf{m} \cdot \nabla_0 \{2\sqrt{rr_0} \cos((\varphi - \varphi_0)/2)\}}{R^2} \right. \\ &\quad \times \left. \sqrt{\frac{\pi R}{2c}} \frac{d\tau}{\sqrt{\pi(t^* - \tau)}} \right\} \\ &\approx \sqrt{\frac{\pi R}{2c}} \left(\frac{\partial}{\partial t^*} \right)^{1/2} \frac{\mathbf{m} \cdot \nabla_0 \{2\sqrt{rr_0^*} \cos((\varphi - \varphi_0^*)/2)\}}{4\pi R^2}. \end{aligned} \quad (24)$$

The halfth time derivative introduces a scale factor $O[(|\dot{\mathbf{r}}_0|/L)^{1/2}]$. The term thus is larger than the incompressible diffracted field term by a factor $O[(\pi/2)(R/L) \times (U/c)^{1/2}]$. Introduction of the trajectory (16) and the complete Green's function (6) into (4) and applying the procedure indicated at the end of Sec. ID and sketched above then yields as the dominant term in the far field pressure signal of the dipole

$$\begin{aligned} P_{\text{wave}} &\approx \sqrt{\frac{\pi R}{2c}} \left(\frac{\partial}{\partial t} \right)^{1/2} \\ &\quad \times \left\{ -\varrho_0 \frac{mU \cos(\varphi/2)(\sin \theta)^{1/2}}{4\pi^2 R^{3/2} L^{3/2}} (\sin(\varphi_0^* - \alpha))^{3/2} \right. \\ &\quad \times \left. \cos(2\alpha - 3\varphi_0^*/2) \right\}. \end{aligned} \quad (25)$$

The relation between the radial distances r and R in the far field is $r \approx R \sin \theta$.

Comparing the above pressure signal (25) with the expression (3.12) of Kambe *et al.* (1985), the results are in perfect agreement when the potential of Kambe's ring vortex is represented by that of a dipole with moment $m = \Gamma a^2$, where a^2 is the area of his ring vortex.

IV. DISCUSSION AND CONCLUSIONS

A. The edge sound of moving dipoles in 2D and in 3D spaces

As a reference in the discussion we have in mind the differences in the undisturbed fields of time dependent dipoles in 2D and in 3D space. In 3D free space the far field sound waves behave as the first time derivative of the incompressible near field variations, they dominate the latter by a factor of the order $(R/c)(\partial/\partial t)$. In 2D space the sound waves behave as the halfth time derivative of the incompressible near field variations, they dominate the latter by a factor of the order $(r/c)^{1/2}(\partial/\partial t)^{1/2}$.

The far field pressure signal p , due to the steady motion of a dipole with constant strength on a straight line trajectory, is not a periodic signal. The signal is varying so slowly ($\partial/\partial t \approx U/r$) that in a compressible fluid only minor modifications do occur relative to an incompressible fluid. The potential decays with distance as $1/r$ in 2D space and as $1/r^2$ in 3D space. The pressure signal, being proportional to the local time derivative of the potential, gets an extra factor with dimension t^{-1} and order of magnitude $[U/L]$.

When the dipole moves at a distance r_0 of the order $r_0 \approx L \ll r$ in the vicinity of a sharp edge in 2D or 3D space, a much stronger signal is generated. In an incompressible field the dominating term in the far field of the potential then comes from the diffracted field having an extra factor $(r/L)^{1/2}$. In a compressible medium the time dependent diffracted potential in the far field gets terms involving the halfth time derivative.

In 2D space these terms are not the dominating ones in the diffracted field, they have the order of magnitude $(U/(cL))^{1/2}$ and they are smaller by a factor $(U/c)^{1/2}$ than the retarded incompressible diffracted field terms that have the order of magnitude $[1/(rL)^{1/2}]$.

In 3D space the diffraction terms involving the halfth time derivative are the dominating ones; they have the order of magnitude $[r^{-1}L^{-1}(U/c)^{1/2}]$. The pressure signal, being proportional to the local time derivative of the potential, gets

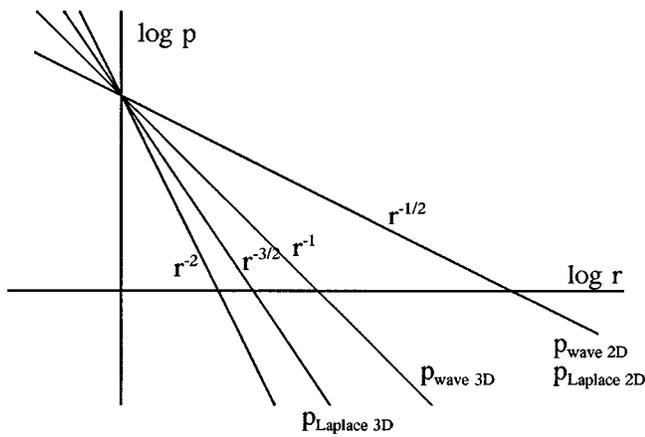


FIG. 5. The far-field pressure decay of the diffracted field in 2D and 3D space.

an extra factor with dimension t^{-1} and order of magnitude $[U/L]$.

The results discussed above are put together in Table I. In Fig. 5 the edge sound pressure decay with distance in 2D and 3D space is presented in a logarithmic plot.

Diffraction on an edge introduces a two-dimensional effect in 3D space in the form of a halfth time derivative of the incompressible diffracted field and an amplification factor $(r/c)^{1/2}$. In 2D space this two-dimensional effect is not markedly present in the diffracted field. It is dominated by a term having the form of the retarded incompressible diffracted field.

It is worthwhile to compare the magnitude of the edge sound with the pressure variations that are observed when a dipole moves in free space. Consider the 2D and 3D far fields for the case of the constant strength dipole moving on a straight line in free space; the far field then is a very slowly varying field and is of minor importance. The fast adaptation of the flow field when the dipole closely passes a sharp edge generates the quickly varying diffracted field, the fast pressure variations of which are the edge sound. The diffracted field in 2D space is different from that in 3D space. In 2D space the diffracted field behaves as the retarded incompressible diffracted field; it behaves as retarded pseudosound. The edge sound field in 2D space would dominate the pressure disturbance due to a freely moving dipole by a factor $(r/L)^{3/2}$.

In 3D space the diffracted wave in the far field is characterized as the halfth time derivative of the retarded incompressible diffracted field. It would dominate the latter by a factor $(R/c)^{1/2}(\partial/\partial t)^{1/2} \approx (R/L)^{1/2}(U/c)^{1/2}$. The edge sound would dominate the pressure disturbance due to a dipole moving in free space by a factor $[(r/L)^2(U/c)^{1/2}]$.

In the description of the edge sound due to the free motion of a vortex Γ rounding the edge of a half-plane, the edge effect and the free space effects both are present. The motion of the vortex on a curved line continuously generates a strong dipole signal in the potential field. The dipoles are continuously being formed perpendicular to the trajectory. This trajectory would be a forced trajectory in free 2D space. The generation of the dipoles would generate a sound wave behaving as the halfth time derivative of the incompressible

field variations; the wave dominates the latter by a factor of the order $(U/L)^{1/2}(r/c)^{1/2}$. These waves differ by only a factor $(U/c)^{1/2}$ in magnitude with the diffraction waves that behave as retarded pseudosound and that dominate the total sound field.

B. Transition from near field to far field of the edge sound

The procedure of integration by parts, applied to the edge sound integrals (3) and (4) with the exact expressions (6) and (9) for the Green's functions as a kernel, leads to a result that consists of two terms as shown in Eq. (14). In 2D space the first of these two terms yields the dominating term. It has the form of retarded incompressible diffracted field; it does not show a typical transition from near field to far field behavior. The second term shows a transition from an incompressible near field to a typical two-dimensional wave effect in the far field where it can be interpreted as a halfth time derivative.

In 3D space the second term, the integral in Eq. (14), yields the diffracted field. This integral shows the transition from an incompressible near field to a far field that is typical for wave behavior in free 2D space.

Kambe *et al.* (1985) considered the edge sound of a ring vortex passing near the edge of a half-plane. They used, in their analysis, a low-frequency approximation to the Fourier transform of the integral representation of Macdonald (1915). The approximation of Kambe, which dominates in the far field, vanishes when the compressibility of the medium vanishes, i.e., when $c \rightarrow \infty$. The closed form integral (23) still contains both the near field and the far field aspects of the edge sound field.

C. Amplitude of the edge sound of vortices and dipoles

Although a dipole has a more singular local velocity field than a vortex (its 2D velocity field behaves as $1/r^2$ whereas the vortex has just $1/r$), the edge sound in the 2D far field of either of them decays as $r^{-1/2}$. In two-dimensional space the edge sound of a vortex moving on a straight line contains the amplitude factor $U\Gamma/(r^{1/2}L^{1/2})$. A dipole moving on a straight path has the factor $Um/(r^{1/2}L^{3/2})$, so a vortex of strength Γ is comparable to a dipole of intensity $m = \Gamma L$. A vortex that moves on a free trajectory does so with its highest velocity near the edge; its amplitude factor is $\Gamma^2/(r^{1/2}L^{3/2})$. Introducing the reference velocity $U = \Gamma/4\pi L$, the amplitude gets the form $U\Gamma/(r^{1/2}L^{1/2})$.

In three-dimensional space it does not make sense to compare edge sound of a line vortex to that of a dipole as only ring vortices can exist. A ring vortex Γ with a ring area a^2 is equivalent to a dipole with intensity $m = \Gamma a^2$. The edge sound pressure signal (25) of a dipole m crossing the edge of a half-plane at a distance L with velocity U has an amplitude factor $\varrho m U (U/c)^{1/2} / (RL^2)$.

D. Green's functions for the half-plane in 3D and in 2D space

The 3D Green's function for the half-plane, obtained by differentiating the unit step response (Cagniard, 1935), is presented in Eq. (6). The 2D Green's function for the half-plane is obtained by the method of descent, i.e., integration with respect to z_0 , of the 3D Green's function. The result, as presented in Eq. (9), must be considered an improvement to the (erroneous) result (8), as known in the literature (Turner, 1956; Lauwerier, 1962; van der Pol and Bremmer, 1964). The closed expressions (6) and (9) for the Green's functions form an efficient basis for the analysis of nonperiodic diffraction signals in the time domain.

APPENDIX A

1. Halfth derivative

The halfth derivative of a function $f(t)$ is defined by Lighthill (1978) as

$$\left(\frac{d}{dt}\right)^{1/2} q(t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^t \frac{\dot{q}(\tau)}{\sqrt{t-\tau}} d\tau. \quad (\text{A1})$$

The potential field of a time dependent dipole, described on the right-hand side of the 2D wave equation as $\mathbf{m} \cdot \nabla_0 \delta(\mathbf{r} - \mathbf{r}_0) f(t)$, is written in terms of the free space Green's function,

$$\begin{aligned} \Phi &= \frac{-1}{2\pi} \int_{-\infty}^{t-\tilde{r}/c} \mathbf{m} \cdot \nabla_0 \frac{f(\tau)}{\sqrt{(t-\tau)^2 - (\tilde{r}^2/c^2)}} d\tau \\ &= \frac{-m\tilde{r} \cos(\beta)}{2\pi c^2} \int_{-\infty}^{t-\tilde{r}/c} \frac{f(\tau) d\tau}{((t-\tau)^2 - (\tilde{r}/c)^2)^{3/2}}, \end{aligned} \quad (\text{A2})$$

where β is the angle between the vector from the dipole to the observation point and the dipole vector \mathbf{m} . Integration by parts then yields, provided that $f(t)$ vanishes at $t = -\infty$,

$$\Phi = \frac{m \cos(\beta)}{2\pi\tilde{r}} \int_{-\infty}^{t-\tilde{r}/c} \frac{\dot{f}(\tau)(t-\tau) d\tau}{\sqrt{(t-\tau)^2 - (\tilde{r}/c)^2}}. \quad (\text{A3})$$

As in the far field the most important contribution will come from the neighborhood of the zero of the square root in the denominator, we rewrite the integral (A3) in terms of $\tau' = \tau + \tilde{r}/c$. For small values of $t - \tau'$ the dipole potential may be written in the approximate form

$$\begin{aligned} \Phi &\approx \frac{m \cos(\beta)}{2\pi\tilde{r}} \int_{-\infty}^{t-\tilde{r}/c} \sqrt{\frac{\tilde{r}}{2c}} \frac{\dot{f}(\tau) d\tau}{\sqrt{t-\tau-\tilde{r}/c}} \\ &= \frac{m \cos(\beta)}{2\pi\tilde{r}} \sqrt{\frac{\pi\tilde{r}}{2c}} \frac{1}{\sqrt{\pi}} \int_{-\infty}^t \frac{\dot{f}(\tau' - \tilde{r}/c) d\tau'}{\sqrt{t-\tau'}}. \end{aligned} \quad (\text{A4})$$

The potential of a dipole in an incompressible medium is obtained from the compressible form (A3) by putting $c = \infty$,

$$\Phi = \frac{m \cos(\beta)}{2\pi\tilde{r}} f(t). \quad (\text{A5})$$

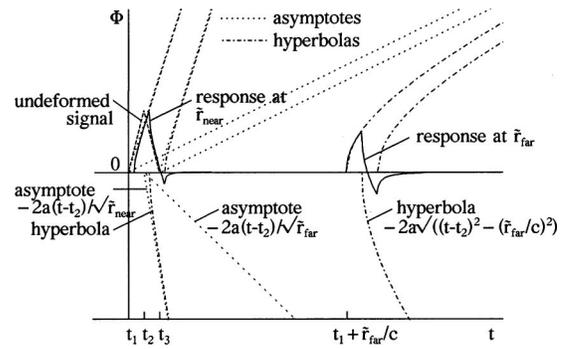


FIG. A1. Transition from undeformed response at $\tilde{r}_{\text{near}}=0.01$ m to halfth derivative response at $\tilde{r}_{\text{far}}=0.4$ m; velocity of sound $c=330$ m/s; $t_1=0.0$, $t_2=0.0001$, $t_3=0.0002$.

In a compressible medium in 2D space the time dependent dipole potential behaves as the half time derivative of the incompressible potential, multiplied by a scale factor $\sqrt{\pi\tilde{r}/(2c)}$.

2. Example, a triangular dipole pulse

The potential of a dipole, growing linearly in time from the instant t_1 , is governed by the inhomogeneous wave equation

$$\nabla^2 \Phi - \frac{\partial^2 \Phi}{\partial t^2} = a(t-t_1)H(t-t_1)(\mathbf{m} \cdot \nabla_0) \delta(r-r_0), \quad (\text{A6})$$

where $H(t-t_1)$ is Heaviside's unit step function. Introduction of the weight function on the right-hand side above into the integral (A3) and performing the integration yields the expression for the potential Φ :

$$\begin{aligned} \Phi &= \frac{a\mathbf{m} \cdot \nabla_0 \tilde{r}}{2\pi\tilde{r}} \sqrt{(t-t_1)^2 - (\tilde{r}/c)^2} \\ &= \frac{am \cos \beta}{2\pi\tilde{r}} \sqrt{(t-t_1)^2 - (\tilde{r}/c)^2}. \end{aligned} \quad (\text{A7})$$

The hyperbola asymptotically approaches the unretarded dipole source function $a(t-t_1)$. Making the approximation in the far field $(t-t_1) + (\tilde{r}/c) \approx (2\tilde{r}/c)$, the expression (A7) shows that the halfth time derivative of $a(t-t_1)H(t-t_1)$ is proportional to $(t-t_1)^{1/2}$. The transition of a signal shape into a far field response as its halfth time derivative is demonstrated in the response to a dipole source-term with a triangular time function, (A8), constructed as a sequence of three constant-slope signals [as in (A6)],

$$\begin{aligned} f(t) &= a(t-t_1)H(t-t_1) - 2a(t-t_2)H(t-t_2) \\ &\quad + a(t-t_3)H(t-t_3). \end{aligned} \quad (\text{A8})$$

The response is plotted in Fig. A1. The triangular input function is transformed into a hyperbolic response.

At some distance \tilde{r}_{far} from the dipole-source point the response starts growing at the retarded time $t_1 + \tilde{r}_{\text{far}}/c$. The shape of the response, the halfth derivative of the input function, is some intermediate form between the shape of the input signal and its first derivative.

- Acheson, D. J. (1992). *Elementary Fluid Dynamics* (Oxford U.P., New York, 1992).
- Cagniard, L. (1935). "Diffraction d'une onde progressive par un écran en forme demi-plan," J. Phys. Radium **VI**(VII), 310–318.
- Cannell, P., and Ffowcs Williams, J. E. (1973). "Radiation from line vortex filaments exhausting from a two-dimensional semi-infinite duct," J. Fluid Mech. **58**(1), 65–80.
- Carslaw, H. S. (1898–1899). "Some multiform solutions of the partial differential equations of Physical mathematics and their applications," Proc. London Math. Soc. **28**, 395–429.
- Crighton, D. G. (1972). "Radiation from vortex filament motion near a half-plane," J. Fluid Mech. **51**(2), 357–362.
- Curle, N. (1955). "The influence of boundaries upon aerodynamic sound," Proc. R. Soc. London, Ser. A **231**, 505–514.
- de Jager, E. M. (1964). "Applications of distributions in mathematical physics," Ph.D. thesis, Amsterdam.
- Felsen, L. B., and Marcuvitz, N. (1973). *Radiation and Scattering of Waves* (Prentice-Hall, Englewood Cliffs, NJ).
- Ffowcs Williams, J. E., and Hall, L. H. (1970). "Aerodynamic sound generation by turbulent flow in the vicinity of a scattering half-plane," J. Fluid Mech. **40**(4), 657–670.
- Friedlander, F. G. (1958). *Sound Pulses* (Cambridge U.P., Cambridge).
- Hadamard, J. (1952). *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Dover, New York, 1952).
- Howe, M. S. (1975). "Contribution to the theory of aerodynamic sound with application to excess jet noise and the theory of the flute," J. Fluid Mech. **71**(4), 625–673.
- Kambe, T., Minota, T., and Ikushima, Y. (1985). "Acoustic wave emitted by a vortex ring passing near the edge of a half-plane," J. Fluid Mech. **155**, 77–103.
- Lauwerier, H. A. (1962). *The Diffraction of a Cylindrical Pulse by a Half-plane* (Mathematisch centrum, Amsterdam), Rep. TW 86.
- Lighthill, M. J. (1952). "On sound generated aerodynamically. Part I General theory," Proc. R. Soc. London, Ser. A **211**, 564–587.
- Lighthill, M. J. (1954). "On sound generated aerodynamically. Part II Turbulence as a source of sound," Proc. R. Soc. London, Ser. A **222**, 1–32.
- Lighthill, M. J. (1978). *Waves in Fluids* (Cambridge U.P., Cambridge).
- Lighthill, M. J. (1982). "Early Development of an 'Acoustic Analogy' Approach to Aeroacoustic Theory," AIAA J. **20**, 449–450.
- Macdonald, H. M. (1915). "A class of diffraction problems," Proc. London Math. Soc. **14**, 410–427.
- Milne-Thomson, L. M. (1968). *Theoretical Hydrodynamics* (Macmillan, New York).
- Powell, A. (1964). "Theory of vortex sound," J. Acoust. Soc. Am. **36**, 177–195.
- Ribner, H. S. (1962). *Aerodynamic Sound from Fluid Dilatations*, UTIA Report No. 86.
- Schouten, G. (1998). *The Two-Dimensional Sound Field of a Vortex Moving Around the Sharp Edge of a Half-Plane* (Delft U.P., Delft).
- Sommerfeld, A. (1897). "Über verzweigte Potentiale im Raum," Proc. London Math. Soc. **30**, 121–163 [corrections to this paper in Carslaw (1898)].
- Turner, R. D. (1956). "The diffraction of a cylindrical pulse by a half-plane," Q. Appl. Math. **16**, 63–73.
- van der Pol, B., and Bremmer, H. (1964). *Operational Calculus* (Cambridge U.P., Cambridge).

Development of a velocity gradient underwater acoustic intensity sensor

Kevin J. Bastyr^{a)} and Gerald C. Lauchle

The Pennsylvania State University, Graduate Program in Acoustics, State College, Pennsylvania 16802

James A. McConnell

The Pennsylvania State University, Graduate Program in Acoustics, State College, Pennsylvania 16802
and Acoustech Corporation, P.O. Box 139, State College, Pennsylvania 16804

(Received 29 March 1999; revised 29 July 1999; accepted 3 September 1999)

A neutrally buoyant, underwater acoustic intensity probe is constructed and tested. This sensor measures the acoustic particle velocity at two closely spaced locations, hence it is denoted a “ $\mathbf{u}\text{-}\mathbf{u}$ ” intensity probe. A new theoretical derivation infers the acoustic pressure from this one-dimensional velocity gradient, permitting the computation of one component of acoustic intensity. A calibration device, which produces a planar standing-wave field, is constructed and tested. In this calibrator, the performance of the $\mathbf{u}\text{-}\mathbf{u}$ intensity probe compares favorably to that of an acoustic intensity probe which measures both pressure and velocity directly. © 1999 Acoustical Society of America. [S0001-4966(99)05312-6]

PACS numbers: 43.30.Xm, 43.30.Yj, 43.58.Fm [DLB]

INTRODUCTION

Acoustic intensity is a measure of the magnitude and direction of acoustic energy transport. Quantitatively, intensity is the period averaged product of acoustic pressure and acoustic particle velocity at a single point.¹ Fahy² shows that intensity is the product of the complex acoustic pressure, p , and the complex conjugate of the acoustic particle velocity, \mathbf{u}^* :

$$\mathbf{I} = \frac{1}{T} \int_0^T p(t) \mathbf{u}(t) dt = \frac{1}{2} p \mathbf{u}^*. \quad (1)$$

Intensity is a vector quantity: there is a direction associated with it. As a vector quantity, it provides physical information that scalar acoustic pressure simply cannot. In a complex sound field, i.e., one having multiple sound sources, an acoustician measuring intensity will be able to quantify the sound power radiated from each source. The vector nature of intensity also enables the location of individual acoustic sources to be determined by simple triangulation.

In the late 1970's Chung³ developed the cross-spectral formulation of the two microphone, or $p\text{-}p$, technique. Basing his formulation on the linearized Euler's equation, Chung showed that intensity is proportional to the cross spectrum between two pressure microphones. Chung's technique, along with FFT analyzers and microphones with approximately matched phases, marked the beginning of reliable intensity measurements. In 1991, Ng⁴ extended the $p\text{-}p$ intensity measurement technique to underwater environments. In 1995, McConnell *et al.*⁵ invented an underwater acoustic intensity sensor that uses a pair of velocity sensors. This $\mathbf{u}\text{-}\mathbf{u}$ intensity probe is quite similar in principle to the $p\text{-}p$ probe.

The first section of this paper deals with the issues fundamental to the operation of the $\mathbf{u}\text{-}\mathbf{u}$ intensity probe. In the second, the theory and procedures used to calibrate the $\mathbf{u}\text{-}\mathbf{u}$

intensity probe are presented. Verification of the calibration's accuracy is then accomplished by comparing intensity measurements made by the $\mathbf{u}\text{-}\mathbf{u}$ probe to those made by a $p\text{-}\mathbf{u}$ probe,^{6,7} a device that directly measures both particle velocity and pressure.

I. VELOCITY GRADIENT INTENSITY

Measurements of intensity in a single dimension are commonly obtained from a pressure gradient. This technique uses the linearized Euler's equation and a finite difference approximation to infer velocity from the pressure gradient measured by two pressure sensors.² In this paper, a new technique invented by McConnell *et al.*⁵ will be used to infer pressure from the velocity gradient measured by two acoustic velocity sensors.

The linearized equation of continuity,¹ a statement of conservation of mass in a fluid, indicates

$$\frac{\partial \rho}{\partial t} + \rho_0 (\nabla \cdot \mathbf{u}) = 0, \quad (2)$$

where \mathbf{u} is the particle velocity, and ρ_0 is the ambient fluid density. The instantaneous density, ρ , is related to the acoustic pressure, p , through the equation of state. Hence, Eq. (2) takes the form

$$\frac{\partial p}{\partial t} + c^2 \rho_0 (\nabla \cdot \mathbf{u}) = 0. \quad (3)$$

Assuming time harmonic quantities and solving for acoustic pressure yields

$$p = \frac{j \rho_0 c^2}{\omega} (\nabla \cdot \mathbf{u}), \quad (4)$$

where ω is the radian frequency of the acoustic wave and c is the speed of sound in the fluid. A finite difference approximation to the velocity gradient yields

^{a)}Electronic mail: zabour@sabine.acs.psu.edu

$$p \approx \frac{j\rho_0 c^2}{\omega} \left[\frac{\Delta u_x}{\Delta x} + \frac{\Delta u_y}{\Delta y} + \frac{\Delta u_z}{\Delta z} \right], \quad (5)$$

where Δu_x is the difference of the x -component of acoustic particle velocity between two locations separated by a distance Δx (similar definitions apply in the y and z directions). It is apparent that pressure depends upon the components of velocity in all three directions. Thus in a three-dimensional acoustic field, all three components of velocity at each of two locations are required to measure the acoustic pressure. However, if the acoustic field is one-dimensional, as in this experimental investigation, a measurement of pressure can be obtained from two single-dimensional velocity sensors. The velocity at the mid-point between these sensors is approximated by the average of the two known velocities. Therefore, the z -component of Eq. (1) takes the form

$$I_z = \frac{j\rho_0 c^2}{2\omega\Delta z} [u_2 - u_1][u_2^* + u_1^*]. \quad (6)$$

Multiplying out the bracketed terms, and noting that $u_1 u_2^* = G_{12}$, the cross-spectrum of u_1 and u_2 , Eq. (6) becomes

$$I_z = \frac{j\rho_0 c^2}{2\omega\Delta z} [G_{22} + G_{21} - G_{12} - G_{11}]. \quad (7)$$

Next, the total intensity is separated into its two constituent parts—active intensity and reactive intensity. The active and reactive components of intensity are the real and imaginary parts of Eq. (7), respectively. Pressure and velocity are entirely in-phase for active intensity, implying both a *net* transport of acoustic energy, and a nonzero time average. The presence of a local active intensity does not infer that energy is being transported throughout an extended region of the field. For reactive intensity, the pressure and velocity are out-of-phase. Therefore, reactive intensity has a zero time average and corresponds to local *oscillatory* transport of energy.² Note that in practice, an acoustic field that is either purely active, or purely reactive cannot exist.

The derivation of an expression for active intensity begins with the real part of Eq. (7). Since auto-spectra are purely real quantities, that expression becomes

$$I_z^a = \frac{\rho_0 c^2}{2\omega\Delta z} \text{Re}[jG_{21} - jG_{12}]. \quad (8)$$

Substituting G_{12}^* for G_{21} and separating the cross-spectra into their real and imaginary components produces the final form:

$$I_z^a = -\frac{\rho_0 c^2}{\omega\Delta z} \text{Im}[G_{12}]. \quad (9)$$

In a similar manner, a simplified form for the reactive component of intensity can be derived; starting with the imaginary part of Eq. (7),

$$I_z^r = \text{Im} \left[\frac{j\rho_0 c^2}{2\omega\Delta z} [G_{22} + G_{21} - G_{12} - G_{11}] \right]. \quad (10)$$

Separating the cross-spectra into their real and imaginary parts and canceling like terms produces the final form:

$$I_z^r = \frac{\rho_0 c^2}{2\omega\Delta z} [G_{22} - G_{11}]. \quad (11)$$

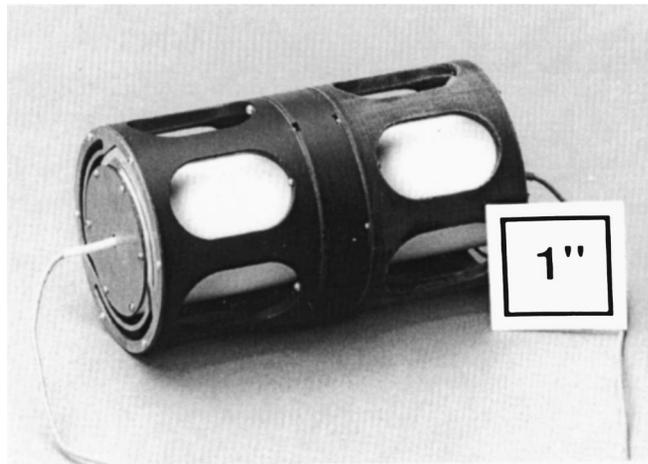


FIG. 1. The velocity gradient, or $u-u$, intensity probe.

As a final note, this development is based on the linearized equation of continuity; thus, the validity of the equations presented is limited to the regime of linear acoustics. In addition, Eqs. (6)–(11) are valid only in one-dimensional acoustic fields.

II. INTENSITY PROBE DESIGN

The $u-u$ probe, as described by McConnell *et al.*,⁵ is a single axis intensity sensor consisting of two moving coil velocity sensors in two separate, coaxially oriented cylindrical bodies. The velocity sensors in this design, GeoSpace GS-14-L3 geophones,⁸ are located in the geometric center of each cylinder. The velocity sensor bodies are neutrally buoyant, meaning that their density is substantially equal to the ambient density of the surrounding fluid medium. The coaxial orientation of the sensor bodies is maintained by suspension springs which connect the ends of each body to a free flooding, acoustically transparent support structure, as illustrated in Figs. 1 and 2. The suspension springs have a high radial stiffness and an extremely low axial stiffness to facilitate the positioning of the sensor without adversely influencing its dynamics.

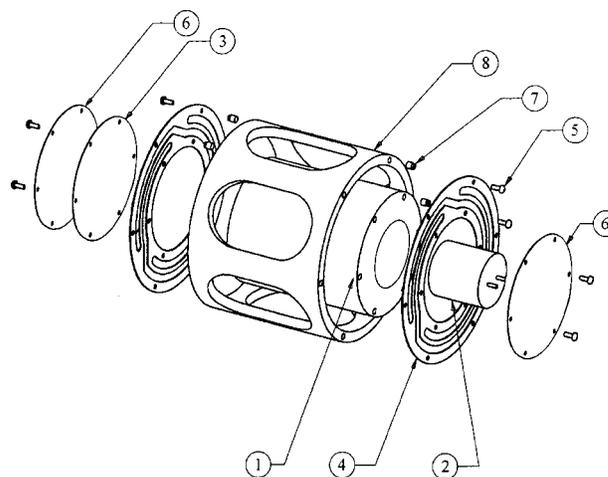


FIG. 2. An exploded view of a single velocity sensor body and its suspension system. 1 is the syntactic foam body, 2 is the geophone, 3 is μ -metal which minimizes magnetic cross-talk, 4 is a suspension spring, 5 are screws, 6 are ballast disks, 7 are inserts, and 8 is the support tube.

A. Fundamentals

Moving coil sensors such as the geophone are primarily used by geophysicists to measure the low frequency movement of the Earth's surface. However, there are distinct advantages to employing geophones as underwater sensors. Geophones have one of the highest signal-to-noise ratios of currently existing transducers, and in comparison to piezoelectric devices, they have a very high receiving sensitivity. In addition, due to their low output impedances, geophones can drive impressive lengths of signal cable without preamplification.

The transducing elements in a geophone consist of a permanent magnet running through the center of the geophone case, and a coil of wire around a spindle. The spindle is constrained by springs to move along the length of the bar magnet. The physics governing the transduction of a geophone is that $e = Blv$. The voltage, e , induced across the ends of the geophone's moving coil equals the product of the transduction coefficient, Bl , and the relative velocity, v , between the coil and the bar magnet running through it.⁹

The mechanical elements of a geophone impose limits on its direct measurement of velocity. The resonance frequency of the geophone is set by its mass-spring system, consisting of the moving coil-spindle and the springs that position it along the bar magnet. For mechanical excitation at frequencies above this resonance frequency, the coil remains virtually motionless while the geophone oscillates around it.⁷ Under these conditions, the output voltage of the geophone is proportional to the case velocity. Below the resonance frequency, the output of the geophone is proportional to jerk, acceleration's time rate of change.

For the geophone's output voltage to be proportional to the acoustic particle velocity, the geophone's velocity must equal the acoustic particle velocity. This can only be achieved if the density of the sensor body matches that of the ambient fluid medium.^{10,11} This condition, known as neutral buoyancy, exists when an object submerged in a fluid neither floats nor sinks. For moving coil sensors, the desired condition is dynamic neutral buoyancy (DNB); i.e., when the dynamic (moving) density of a body matches that of the fluid medium. As previously discussed, for excitation at frequencies above the geophone's resonance frequency, the proof mass remains essentially motionless; therefore, a geophone's dynamic density is less than its static density.

For the GS-14-L3 geophone to become neutrally buoyant, it must be imbedded in a substantially larger volume of positively buoyant material. Syntactic foam, a composite of epoxy binder¹² and tiny hollow glass spheres called microballoons,¹³ is such a material. To obtain foam of very low density, a mixture of 4 parts microballoons to 1 part epoxy binder is used. When the foam has cured, it is machined to allow insertion of a geophone into the geometric center of the body. Due to the foam's low density of ≈ 0.5 g/cc, the length of a 3.80 cm diameter neutrally buoyant right circular cylinder is less than 5 cm. For the calculation of the cylinder's exact length, the masses of all the sensor's components must be known. Table I outlines the mass and volume of each component in a single velocity sensor body, which is illustrated in Fig. 2.

TABLE I. The masses and volumes of the $u-u$ probe's constituent parts. Note that SNB denotes static neutral buoyancy, and DNB denotes dynamic neutral buoyancy.

Component	Mass (g)	Volume (cc)
Foam cylinder	22.7	45.5
Geophone	18.3	3.82
μ -metal	1.43	0.232
Springs (2)	1.82	0.177
Fasteners	0.053	0.006
Ballast for SNB	6.10	0.695
Subtotal	50.4	50.4
Ballast for DNB	2.43	0.276
Total	52.9	50.7

To achieve neutral buoyancy, even the effective mass of the suspension springs must be accounted for. To first order, the moving mass of a cantilever-type spring in a simple harmonic oscillator is equivalent to one quarter the moving mass of the spring.¹⁴

To test for neutral buoyancy, the sensor body is submerged in water. If the sensor neither sinks nor floats, it is statically neutrally buoyant. Testing indicated that a 1% deviation from neutral buoyancy was observable. That is, if the total sensor mass was 1% lower (or higher) than that of a neutrally buoyant body, the sensor would float (or sink). To achieve dynamic neutral buoyancy, the sensor body is ballasted with a mass equal to the geophone's proof mass, without displacing any additional volume. In practice it is easier to add mass in the form of ballast disks, as in Fig. 2. As these disks displace volume, a mass equal to the proof mass plus the mass of water displaced by the ballast disk is added, as listed in Table I.

B. Suspension system

The suspension system consists of two elements: a support tube and two beryllium copper springs. The support tube was machined from nylon-6, a material that exhibits an acoustic impedance substantially equal to that of bulk water. The suspension springs, which are illustrated in Figs. 1 and 2, serve two primary functions: maintaining the coaxial orientation of the sensor bodies and isolating the sensors from any vibration in the structure to which the $u-u$ probe is mounted. Due to the springs' high radial stiffness and extremely low axial stiffness, the sensor body is "free" to move in the axial (sensing) direction, but has no excursion in the radial direction.

The suspension springs were photo-etched from 0.20 mm (8 mil) beryllium copper shim stock. These springs are similar in design to those used to support the proof mass in typical geophones. However, they have two subtle differences: the suspension springs are larger and have two bolting rings. The bolting rings allow the springs to be connected to both the support tube and the sensor body. The springs are fastened using 00-90 screws and threaded press-fit inserts as shown in Fig. 2.

The suspension springs' design gives the sensor mass-spring system a sufficiently low resonance frequency that the suspension dynamics do not adversely affect the probe's performance. Modeling the spring as a cantilever beam allows

the theoretical prediction of its resonance frequency. This model rests on several assumptions. First, the mass of the sensor body is approximated as a point mass at the end of the cantilever beam. Further, the slightly curved cantilevers are modeled as straight beams. This approximation is valid because the amplitude of the beams' deflection is small compared to their overall length. Note that each of the six spring members (i.e., three per spring) support one-sixth the sensor's total mass. The theoretical resonance frequency of the probe-suspension system is 6.28 Hz, as calculated using equations from Thomson.¹⁵

The resonance frequency of the probe-suspension system was experimentally measured underwater. An underwater source acoustically excited a single probe body's geophone, while a hydrophone was used as a reference transducer. The probe-suspension system resonance frequency was determined from the frequency response measured between the two transducers to be (6.13 ± 0.05) Hz. The Q of the probe-mass/suspension-spring resonance is approximately 15, and the system's resonance frequency is substantially lower than that of the geophone (28 Hz). Accordingly, the phase and amplitude errors introduced at frequencies above the geophone's resonance are negligible.⁷

C. Performance issues

The distance separating the acoustic centers of the velocity sensors determines the upper frequency of operation. This distance also affects the amount of undesirable interaction between the sensor bodies. This cross-talk between the sensors manifests itself in three forms: structural, magnetic, and hydrodynamic.

The upper frequency limit of the $u-u$ intensity probe is defined as the frequency at which a 10% (1 dB) bias error occurs in the estimation of acoustic intensity. A bias error is the difference between the measurement of a quantity, and the quantity's true value. In this case, the bias error results from using a finite difference approximation of the velocity gradient in Eq. (5). Following Fahy's² bias error derivation for $p-p$ intensity probes exposed to a plane wave field, the normalized bias error in the $u-u$ probe's estimation of pressure and velocity are

$$\epsilon_b[p] = \frac{\sin(k\Delta z)}{k\Delta z} - 1, \text{ and } \epsilon_b[u_z] = \cos(k\Delta z) - 1, \quad (12)$$

where $k = \omega/c$. Forming the normalized bias error of the intensity estimate, and making a Taylor series expansion¹⁶ yields:

$$\epsilon_b[I_z] \approx -\frac{2}{3}(k\Delta z)^2 + \frac{1}{20}(k\Delta z)^4. \quad (13)$$

The lower frequency limit of the probe is effectively set by the phase mismatch between the two sensors. Of course, the majority of the phase mismatch can be removed through proper calibration, as will be shown later.

In addition to the finite difference error, the sensor spacing also affects the amount of cross-talk between the sensors. Magnetic coupling can occur between two geophones in close proximity, as in the $u-u$ probe. Relative motion be-

tween the magnet of one geophone and the coil of the second will induce an undesirable current in that second coil.

Hydrodynamic coupling between the sensor bodies occurs when slight deviations from neutral buoyancy alter the bodies' velocity from that of the acoustic particle velocity. This nonideal movement of one sensor body disturbs the acoustic field sensed by the second body. Due to their design, the suspension springs may also entrain mass, and contribute to hydrodynamic coupling. The suspension springs employed in the $u-u$ intensity probe are attached to both the moving probe bodies, and the stationary structural support tube. The spring members are neither stationary nor moving with the acoustic particle velocity; therefore, there is mass entrained with the springs' motion.

The mechanical connection between the sensor bodies (via the support tube) gives rise to the possibility of undesirable vibrational coupling between them. However, the vibration isolation provided by the suspension springs renders this negligible.¹⁷

To ensure the proper operation of the $u-u$ probe, the contributions of all these coupling mechanisms occurring in typical operating conditions must be measured. Accordingly, a scenario to simulate these conditions was devised. The coil of one geophone is electrically driven to simulate the motion (effective acoustic velocity) that would occur if the $u-u$ probe were subjected to an acoustic plane wave field of amplitude 150 dB *re*: 1 μ Pa. The voltage induced (in the second geophone) by the combined magnetic and hydrodynamic coupling can be adjusted by the geophone's receiving sensitivity, providing an "effective velocity due to coupling." This velocity can then be compared to the effective acoustic velocity in order to quantify the extent of the coupling.

In water, a plane wave with an amplitude of 150 dB *re*: 1 μ Pa corresponds to an effective plane wave particle velocity of 2.1×10^{-3} cm/s. A GS-14-L3 geophone with known sensitivity was driven with a broadband signal equivalent to 150 dB *re*: 1 μ Pa in each frequency analysis bin.

In this measurement, one body was driven electrically at the aforementioned level, and the motion induced in the other body was measured. This test was performed in water so that the combined hydrodynamic and magnetic effects could be compared to in-air test results, which essentially measured only the magnetic effects. Figure 3 illustrates these velocities normalized by the effective acoustic velocity. The velocity induced by these coupling mechanisms is at most 25 dB below the acoustic disturbance of interest (the effective 150 dB *re*: 1 μ Pa plane wave). Therefore, the effects of the coupling mechanisms are considered to be insignificant.

III. CALIBRATION FACILITY

Calibrating an acoustic intensity sensor underwater is a formidable challenge. Standard underwater intensity and velocity transducers do not exist at this time, so a relative calibration is impossible. Simply attaching an accelerometer to each sensor body is initially an attractive option; however, this adds mass to the body, causing deviations from neutral buoyancy. Further, it can provide no indication of whether or not the probe body's velocity matches the acoustic particle velocity. Calibration in air, where velocity sensors are avail-

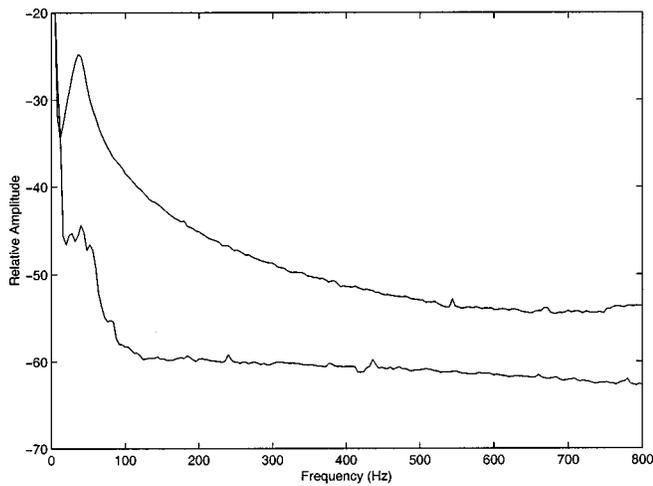


FIG. 3. The ratio of the hydrodynamically and magnetically induced velocity relative to the effective acoustic velocity, the 150 dB *re*: 1 μ Pa plane wave. The upper and lower curves are in water and in air, respectively.

able in the form of hot wire anemometers, is not an option, because the sensor bodies are not neutrally buoyant in air.

A reasonable option for a transducer to employ in an underwater comparison calibration is a hydrophone. The use of a hydrophone in a relative calibration relies on knowledge of the phase relationship between pressure and velocity in the acoustic field. The ideal calibration would take place in a progressive plane wave field, where the phase between pressure and velocity is 0° . Such fields are difficult to achieve, even in a laboratory setting. A standing plane wave field, where the pressure and velocity are 90° out-of-phase, offers an alternative option for calibration. This is perhaps the harshest field in which to perform a calibration, due to the presence of pressure and velocity nodes. However, this type of field can be created in the laboratory, which is a benefit outweighing its numerous detriments.

A. Description of the slow wave calibrator

The slow wave calibrator (SWC) developed in this research is a vertically oriented, water filled, elastic walled, one-dimensional waveguide. The SWC is illustrated in Fig. 4. A piston sound source is located at the lower end, while the opposite end is an air–water interface. The 20.3 cm diameter cylindrical acrylic wall of the SWC is 1.22 m tall, and has a thickness of 0.64 cm. A USRD J9 source,¹⁸ with a piston diameter of 7.5 cm, acts as the sound source for the SWC. The acoustic standing waves in the SWC depend on the acoustic reflection from the water–air interface, which is a pressure-release boundary. Accordingly, the phase between pressure and velocity at any axial location in the SWC is 90° .

One advantage a SWC has over a rigid walled calibrator is that the phase speed of the fluid inside the elastic walled SWC is substantially reduced from that in the bulk medium.^{19,20} This decreased phase speed dramatically shortens the wavelengths, and the accuracy of low frequency measurements made by the *u-u* probe is improved by lessening the influence of transducer phase mismatch. This advantage may be offset by an increase in the finite difference error given by Eq. (13).



FIG. 4. The slow wave calibrator with the *u-u* intensity probe and reference transducers suspended from above. Note that the two cables attached to the USRD J9 source are for the electrical driving signal and the driver's hydrostatic pressure compensation.

B. SWC acoustic performance

The acoustic waveguide nature of the SWC largely contributes to the planarity of the wavefronts. Detailed mapping of the cross section of the SWC with a hydrophone found the wavefronts were to be planar within 1 dB from 28 Hz to 850 Hz in the section used for sensor calibration.

The longitudinal sound speed in the SWC dictates two important parameters of the facility: the modal resonance frequencies, and the plane-wave cutoff frequency. The phase speed of acoustic waves in a fluid medium is proportional to the square root of an elastic modulus. In a large body of fluid, the relevant quantity is the bulk modulus. However, the elastic vessel constraining the fluid in the SWC contributes the majority of the system's compliance. Consequently, the effective sound speed, c_{eff} , in an elastic-walled waveguide²⁰ is given by

$$c_{\text{eff}} = \sqrt{B / \left(\rho_0 \left(1 + \frac{2r\rho_0 c^2}{tE} (1 - \nu^2) \right) \right)}, \quad (14)$$

where B is the bulk modulus of the fluid, ρ_0 is the density of the fluid, r is the inner radius of the duct, c is the bulk phase speed of sound, t is the thickness of the duct wall, E is Young's modulus of the wall material and ν is the Poisson's ratio of the wall material. Using Eq. (14), the theoretical phase speed, c_{eff} , in the SWC is (290 ± 10) m/s. The uncertainty is due to the variance in the acrylic wall's Young's modulus, as specified by the manufacturer.²¹

This theoretically determined sound speed was verified experimentally by two independent measurements. The first was to ensonify the SWC at its resonance frequencies and measure the distance between adjacent nodes. The steep pressure gradient around each node improves the accuracy with which an axially scanned hydrophone can locate each pressure minimum. The distance between two adjacent pressure minima corresponds to a half wavelength of the driving frequency. The nodes of three resonance frequencies of the SWC were measured, and basic acoustics²² (i.e., $c=f\lambda$) indicated that the sound speed was (355 ± 8) m/s, which is less than one-quarter the bulk speed in water. Note that within the uncertainty of the data, the SWC is a nondispersive system.

Recall that the resonance frequencies of a system are the frequencies at which the imaginary part of the total impedance vanishes. The SWC is not a purely acoustic system because the mechanical impedance of the sound source is coupled to the acoustic domain, resulting in a nonnegligible shift of the resonance frequencies. The second method to measure the sound speed effectively uncouples the acoustic and mechanical domains, thus removing the influence of the driver's impedance from measurements of the SWC's resonance frequencies.

The details of this method were presented by McConnell,²³ therefore, only a brief overview will be presented here. The resonances of the acoustic portion of the SWC can be obtained from a measurement of the transfer function between a hydrophone at an arbitrary depth below the surface of the SWC and an accelerometer²⁴ mounted on the piston of the J9. From these frequencies, the speed of sound can easily be determined with the following relationship:

$$c_{\text{eff}} = \frac{2f_n l}{n + \frac{1}{2}}, \quad (15)$$

where $n=0,1,2,\dots$, and l is the length (height) of the SWC. Employing Eq. (15), and averaging together the five lowest measured resonance frequencies, results in a sound speed of (360 ± 3) m/s. All five resonance frequencies fell within 3 Hz of the theoretical values calculated with the assumption of this sound speed.

According to the experimentally determined sound speed, the plane-wave cutoff frequency of the SWC is ≈ 1 kHz. Measurements of the wavefront planarity show abrupt deviations from planar behavior at 900 Hz, supporting this estimate. The wavefronts are planar within 1 dB from 28 Hz to 850 Hz in the section of the SWC used during the sensor's calibration.

The theoretically calculated sound speed is within 20% of the experimentally measured value. The discrepancy is likely due to the uncertainty in many of the published material properties of the acrylic tube. The agreement between the two experimental methods is quite good, however. The sound speed as determined by the transfer function method is (360 ± 3) m/s. This value lies within the experimental uncertainty of the (355 ± 8) m/s value obtained from the pressure null measurements.

C. System transfer functions for calibration

The $u-u$ intensity probe must be calibrated to produce a measurement of intensity. The SWC is used as the calibration facility. Excluding the extreme outer radial area of the SWC, sensors at coincident depths are subjected to the same acoustic field. Because the phase between the pressure and velocity is a known 90° , a calibration relative to pressure hydrophones can be performed. Note that this procedure is a proof of principle; it is neither an absolute calibration, nor the best method of calibration conceivable.

The measurement of active intensity produced by the $u-u$ intensity probe is proportional to the imaginary part of G_{12} , whereas that of reactive intensity is proportional to the difference of G_{22} and G_{11} as illustrated in Eqs. (9) and (11). Spectral quantities such as these are not directly measurable, because they involve particle velocities, whereas the sensors' output is voltage. A transfer function, H , relates an input quantity, x , to a measured output quantity, \hat{x} :

$$x = \frac{\hat{x}}{H}. \quad (16)$$

It follows that

$$G_{11} = \frac{\hat{G}_{11}}{H_{11}}, \quad (17)$$

where \hat{G}_{11} is a voltage spectrum, and H_{11} is the transfer function that relates geophone velocity to output voltage. In order that no assumptions be made about the magnitude and phase of the geophone output voltage, an expression for H_{11} that depends on the reference hydrophone output voltage will be derived. For this reason, the velocity field in the SWC must be expressed in terms of the acoustic pressure. Figures 5 and 6 illustrate the layout of the sensors in the SWC. The expression describing the pressure at a depth z below the free surface of the SWC is:

$$p(z,t) = p_0 \sin(kz) e^{j\omega t}. \quad (18)$$

The particle velocity is:

$$u_z(z,t) = j \frac{p_0}{\rho_0 c} \cos(kz) e^{j\omega t}. \quad (19)$$

Combining these produces:

$$u_{z_1}(d_1) = j \frac{p_1(d_1)}{\rho_0 c \tan(kd_1)} \quad \text{and} \quad u_{z_2}(d_2) = j \frac{p_2(d_2)}{\rho_0 c \tan(kd_2)}. \quad (20)$$

The transfer functions for the pressure and velocity sensors are:

$$u_{z_1} = \frac{\hat{e}_{u_1}}{H_{u_1}}, \quad u_{z_2} = \frac{\hat{e}_{u_2}}{H_{u_2}}, \quad p_1 = \frac{\hat{e}_{p_1}}{H_{p_1}}, \quad p_2 = \frac{\hat{e}_{p_2}}{H_{p_2}}, \quad (21)$$

where subscripts 1 and 2 refer to locations d_1 and d_2 , respectively. Note that the subscript z is suppressed in the voltage and transfer function terms.

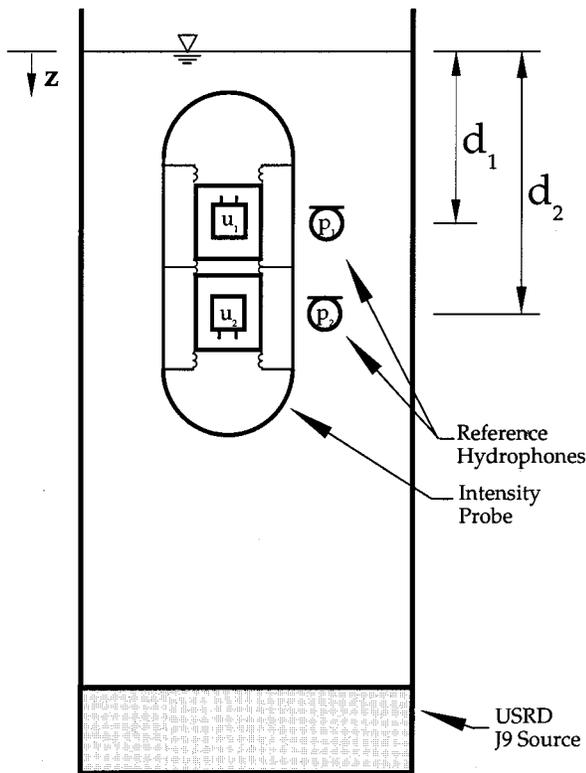


FIG. 5. The orientation of the reference hydrophones and $u-u$ intensity probe in the SWC.

Substituting the acoustic quantities from Eq. (21) into Eq. (20) produces a transfer function for each velocity sensor that involves the voltages measured by the pressure and velocity sensors:

$$H_{u_1} = -j\rho_0 c \tan(kd_1) H_{p_1} \frac{\hat{e}_{u_1}}{\hat{e}_{p_1}} \quad (22)$$

and

$$H_{u_2} = -j\rho_0 c \tan(kd_2) H_{p_2} \frac{\hat{e}_{u_2}}{\hat{e}_{p_2}}.$$

Recall that the quantities H_{11} , H_{12} , and H_{22} , which are used for adjusting the three measured velocity spectra involved in intensity measurements, are defined as:

$$H_{11} = H_{u_1} H_{u_1}^*, \quad H_{12} = H_{u_1} H_{u_2}^*, \quad H_{22} = H_{u_2} H_{u_2}^*. \quad (23)$$

Combining Eqs. (22) and (23) produces:

$$H_{12} = (\rho_0 c)^2 \tan(kd_1) \tan(kd_2) H_{p_1} H_{p_2}^* \frac{\hat{e}_{u_1}}{\hat{e}_{p_1}} \frac{\hat{e}_{u_2}^*}{\hat{e}_{p_2}^*}, \quad (24)$$

$$H_{11} = (\rho_0 c)^2 \tan(kd_1) \tan(kd_1) H_{p_1} H_{p_1}^* \frac{\hat{e}_{u_1}}{\hat{e}_{p_1}} \frac{\hat{e}_{u_1}^*}{\hat{e}_{p_1}^*}, \quad (25)$$

and

$$H_{22} = (\rho_0 c)^2 \tan(kd_2) \tan(kd_2) H_{p_2} H_{p_2}^* \frac{\hat{e}_{u_2}}{\hat{e}_{p_2}} \frac{\hat{e}_{u_2}^*}{\hat{e}_{p_2}^*}. \quad (26)$$

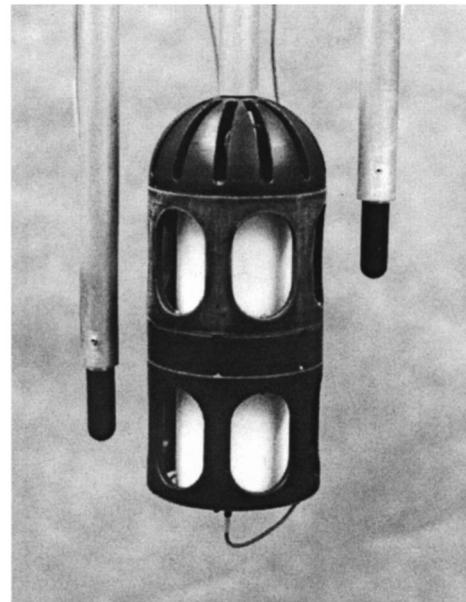


FIG. 6. The arrangement of the sensors in the $u-u$ probe calibration procedure. Note that the aluminum positioning rods are attached to a three axis positioning system from above.

The hydrophone transfer functions can be separated into magnitude and phase components such that,

$$H_{p_{12}} = H_{p_1} H_{p_2}^* = |H_{p_1}| |H_{p_2}| e^{j(\phi_1 - \phi_2)}, \quad (27)$$

where $|H_{p_1}|$ and $|H_{p_2}|$ are the open circuit receiving sensitivities of hydrophones 1 and 2, M_{p_1} and M_{p_2} , respectively. The phase difference between the two hydrophones can be measured using a standard switching technique.^{3,25}

Substituting a simplified version of the transfer function into Eq. (9) produces the final form of the equation for active intensity measurements:

$$I_z^a = - \frac{1}{\omega \rho_0 \Delta z \tan(kd_1) \tan(kd_2) M_{p_1} M_{p_2}} \times \text{Im} \left[\frac{\hat{G}_{12}^c \hat{G}_{p_1 p_2}^c}{\hat{G}_{u_1 u_2}^c} e^{j(\phi_1 - \phi_2)} \right]. \quad (28)$$

Here the transfer function given in Eq. (24) has been simplified by a change in notation: $\hat{G}_{u_1 u_2}^c$ is the cross-spectrum between the geophone output voltages, and $\hat{G}_{p_1 p_2}^c$ is the cross-spectrum between the hydrophone output voltages. The superscript c denotes that both of these quantities are measured in the probe's calibration procedure. The term \hat{G}_{12}^c is the voltage cross-spectrum between the geophone outputs in an intensity measurement situation.

The analogous substitutions into Eq. (11) produce the working expression for reactive intensity:

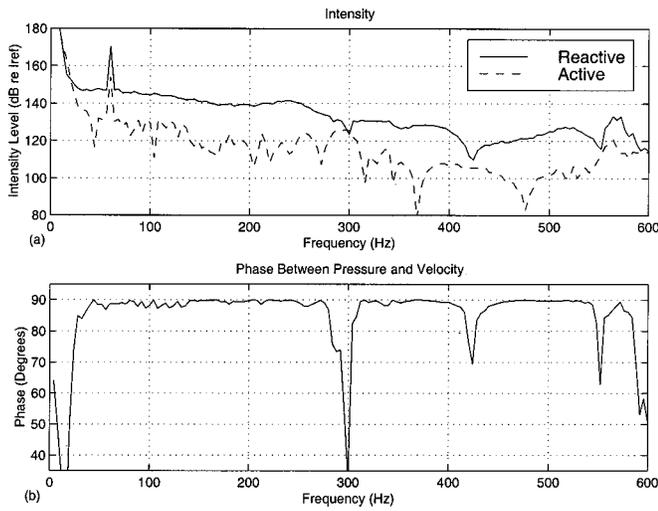


FIG. 7. Intensity and pressure-velocity phase at a depth of 20 cm in the SWC, as measured by the reference hydrophones using the p - p technique. The peak at 60 Hz is due to EMI. Note that I_{ref} is $10^{-12}/\rho c$ W/m².

$$I_z^r = \frac{1}{2\omega\rho_0\Delta z} \times \left[\frac{\hat{G}_{22}\hat{G}_{p_2p_2}^c}{\tan^2(kd_2)M_{p_2}^2\hat{G}_{u_2u_2}^c} - \frac{\hat{G}_{11}\hat{G}_{p_1p_1}^c}{\tan^2(kd_1)M_{p_1}^2\hat{G}_{u_1u_1}^c} \right]. \quad (29)$$

IV. MEASUREMENT OF INTENSITY

The calibration setup depicted schematically in Fig. 5 and photographically in Fig. 6 can also be used to measure the acoustic intensity of the standing-wave field set up within the SWC. The two pressure hydrophones can be used for a p - p intensity measurement that can be compared to the u - u probe's measurement. A second verification of the u - u probe's accuracy is based on a comparison of a u - u intensity measurement with a measurement made by a p - u intensity probe.²⁶

An eight channel Hewlett Packard 3567A dynamic signal analyzer (DSA) was used to process the transducer outputs. Any phase difference between the eight DSA channels will have a significant impact on the measurement accuracy. The greatest amount of amplitude and phase mismatch between any of the DSA channels was measured to be 0.1% and 0.02°, respectively.

The hydrophones²⁷ used for the calibration have a sensitivity of -210 dB *re*: 1 V per μ Pa. These piezoceramic hydrophones have a nominal capacitance of 3.4 nF. A preamp was used with these hydrophones to avoid amplitude and phase errors. The former would be created by the voltage divider consisting of the hydrophone capacitance and the input capacitance of the DSA (100 pF), while latter would be introduced at low frequencies by the low pass filter consisting of the hydrophone's capacitance and the input impedance of the DSA. Ithaco 1201 preamplifiers,²⁸ with a nominal input impedance of 100 M Ω , were used with all the hydrophone measurements presented in this paper.

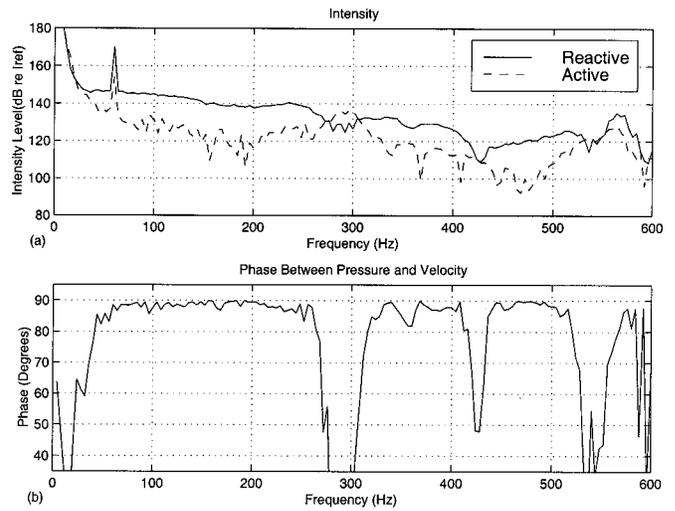


FIG. 8. Intensity and pressure-velocity phase at a depth of 20 cm in the SWC, as measured by the u - u intensity probe. The peak at 60 Hz is due to EMI, and I_{ref} is $10^{-12}/\rho c$ W/m².

The signal driving the J9 source, which ensouffled the SWC, was generated by the DSA. Due to the limitation of the DSA's source voltage and the low sensitivity of the J9, a power amplifier²⁹ was required to boost the driving signal's amplitude. The frequency response of the power amplifier is flat within 0.1 dB over the range of 28 Hz to 2800 Hz.

A. Pressure gradient technique

In effort to remove any phase mismatch between the reference hydrophones in subsequent measurements, their phase difference was measured with a standard switching technique^{3,25} conducted in the SWC with the acoustic centers of the hydrophones located at coincident depths. Recall that the wavefronts in the SWC are planar within 1 dB; therefore, the implementation of the switching technique might seem unnecessary. However, the switching technique lessens the effects of the SWC's standing-wave resonances on the measurement. In this measurement, the coherence function between the hydrophones fell below 0.999 only at 60 Hz, and at frequencies above the plane-wave cutoff frequency (≈ 1 kHz).

A velocity maximum exists for all frequencies at the air-water interface, making it the ideal location for achieving a high signal-to-noise ratio for velocity sensor measurements. However, the pressure at that location is effectively zero for all frequencies, making it the worst possible location for hydrophone measurements. Similarly, for a given frequency, the locations of pressure maxima are velocity minima. Therefore, the location of the sensors in the SWC is essentially an arbitrary choice, because at every location in the SWC except the surface, a velocity minima exists for some frequency. For the following measurements, the u - u intensity probe and reference hydrophones were mounted 20 cm below the free surface of the SWC, which was filled with water until a meniscus formed above the top of the acrylic wall. Completely filling the SWC alleviates any concern of standing-wave formation in the air column above the air-water interface.

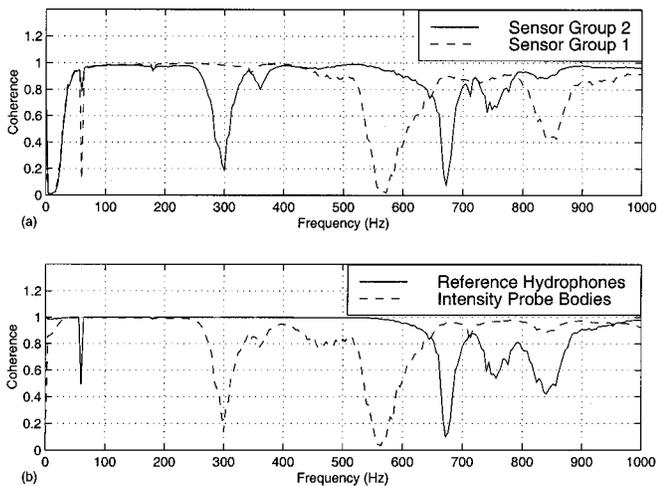


FIG. 9. (a) The coherence functions between each probe body and the corresponding reference hydrophone. (b) The coherence functions between the probe bodies, and between the reference hydrophones.

Intensity is measured with the p - p technique during the calibration of the u - u intensity probe. Equations (28) and (29) predict that if the calibration procedure and intensity measurement are coincident, all the velocity spectra from the u - u probe bodies will cancel, producing a p - p intensity measurement. The phase information determined by the switching technique and the various cross-spectra and physical quantities illustrated in Eqs. (28) and (29) generated the intensity curve in Fig. 7(a). This measurement was repeated without the intensity probe located between the hydrophones, and no significant changes resulted. This indicates that the presence of the u - u probe did not adversely affect the p - p probe intensity measurement.

The reactive component of intensity illustrated in Fig. 7(a) is nominally 20 dB higher than the active component. This is expected, because the energy in the SWC is predominantly reverberant, or reactive. To quantify the extent to which the field is reactive, the phase between pressure and velocity is calculated. An equivalent and simpler method is to calculate the phase between the real and imaginary parts of intensity. This is equivalent to the phase of G_{pu} , which is given in Fig. 7(b). The phase between pressure and velocity should be 90° ; however, the data of Fig. 7(b) indicate that there are deviations from this ideal. Recall that active intensity corresponds to local transport of energy. This occurs to a small extent in the form of sound emitted from the SWC. In addition, viscous and thermal relaxation losses in the acoustic boundary layer, near the tube wall, contribute to the presence of active intensity. Energy transport may also occur through flexure-induced dissipation in the acrylic walls of the SWC. Limitations in the extent to which the phase difference between the hydrophones can be removed from their measurements is another source of error.

Because the random error in a spectrum level is inversely proportional to the square root of the number of spectral averages,³⁰ the random error in the intensity measurements given here (with 256 averages) is ± 0.3 dB.

B. Velocity gradient technique

The procedure for measuring intensity with the u - u probe is analogous to that of the p - p measurement. This measurement employs the same sensors, instrumentation, and SWC. Equations (28) and (29) are again used to compute intensity; however, the calibration procedure must be separate from the measurement of intensity. The results of the u - u intensity measurement are presented in Fig. 8. Note that both the calibration procedure and intensity measurement occurred at the same location in the SWC.

As in the p - p measurements in Fig. 7(b), the phase between pressure and velocity exhibits some deviations from 90° . The extent to which the phase difference between the hydrophones cannot be removed significantly contributes to these deviations. The lack of coherence between the individual velocity sensors and their reference hydrophones at certain frequencies is the most significant source of these phase deviations. This lack of coherence results from presence of standing-waves: at any location in the SWC, there exist frequencies at which either the pressure or velocity is approximately zero. The accuracy of the calibration and subsequent measurements is diminished at these frequencies.

Figure 9 illustrates the various coherence functions between pairs of sensors during the calibration. Cross referencing these plots yields a great deal of information about the standing-wave nature of the SWC. The frequencies of 300 Hz and 560 Hz are antiresonance frequencies of the SWC. These are the frequencies at which the most energy must be input to produce a constant level of acoustic pressure (at a given location) in the system. The coherence functions of several sensor pairs have an extremely low value at these antiresonance frequencies. Further, the frequency range for which the calibration has validity is limited by the overlapping drop-outs in coherence at 560 Hz and 670 Hz.

A comparison of the intensity levels of Figs. 7(a) and 8(a) shows excellent agreement between the p - p and u - u techniques. This agreement is not conclusive proof that the u - u probe functions properly, because the transducers used in the former measurement served as the reference transducers in the latter; however, it certainly is supporting evidence.

C. Pressure-velocity technique

An independent method to verify that the u - u probe properly measures acoustic intensity is now discussed. A measurement was performed using a well calibrated underwater acoustic intensity probe on loan from Acoustech Corporation.²⁶ This sensor is called a p - u probe,^{6,7} because it measures both pressure and velocity directly. Accordingly, the p - u probe does not suffer from the finite difference errors that affect the accuracy of the other intensity transducers used in this investigation.

The signal processing required to measure intensity with the p - u probe is surprisingly simple. By definition, intensity is the cross-spectrum of pressure and velocity. Consequently, the sensor output voltages are adjusted by their measured sensitivities to obtain direct measures of the acoustic quantities. The cross-spectrum of pressure and velocity is computed and then separated into real and imaginary parts. The

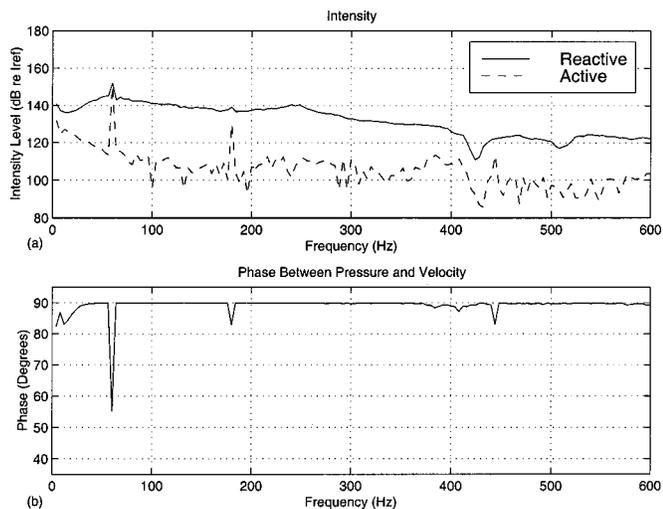


FIG. 10. Intensity and pressure-velocity phase in the SWC, as measured by the p - u intensity probe at a depth of 20 cm. The peak at 60 Hz is due to EML, and I_{ref} is $10^{-12}/\rho c$ W/m².

resulting data appear in Fig. 10. Note that the phase between pressure and velocity, as computed by the p - u probe, is almost a perfect 90° over the entire frequency range of interest. The reactive intensity levels are very similar to those of Figs. 7(a) and 8(a), except at the antiresonance frequencies. The active intensity levels do not compare well for reasons discussed in the next section.

V. SUMMARY AND CONCLUSIONS

A new technique for measuring underwater acoustic intensity has been presented. The theory to infer acoustic pressure from a velocity gradient was derived. It was shown that in a typical acoustic field, six velocity sensors (to measure three velocity gradients) are required for the measurement of the three components of acoustic intensity. In a single dimensional acoustic field, measurement of the acoustic velocity gradient was accomplished using two neutrally buoyant bodies with imbedded velocity sensors (geophones). Active intensity was shown to be proportional to the imaginary part of the cross-spectrum between the velocity sensors, while reactive intensity was proportional to the difference of their auto-spectra.

The suspension system of the neutrally buoyant body maintains their proper coaxial orientation and simplifies the mounting of the probe. This suspension does not adversely affect measurement quality in the frequency range of interest, from 28 Hz to 800 Hz. The lower limit is imposed by the resonance frequency of the geophones.

The geometry of the u - u intensity probe gave rise to potential undesirable interaction between the two sensor bodies. However, when the probe was subjected to an effective 150 dB SPL field, the total coupling between the bodies was shown to be negligible.

A facility to calibrate the u - u intensity probe has been designed and constructed. Termed the slow wave calibrator (SWC), this device is essentially an acoustic standing-wave guide. The elastic walls of the SWC lower the longitudinal phase speed of sound from the bulk speed of 1480 m/s to

(358 ± 5) m/s. The acoustic wavefronts in the SWC have a cross-section that is planar within 1 dB. Thus sensors located at coincident depths are subjected to the same acoustic field.

The u - u probe calibration procedure is based on only one assumption: that the free surface of the calibrator is an ideal pressure-release surface. This assumption implies that the phase between pressure and velocity in the SWC is exactly 90°. A relative calibration of the u - u probe's velocity sensor bodies with pressure hydrophones was therefore possible.

To verify that the two reference hydrophones are phase matched accurately enough to provide a reliable u - u probe calibration, a p - p intensity measurement was made using the hydrophones. The intensity levels measured by the u - u probe differ only slightly from those measured by the p - p probe.

Differences in level are attributed both to the numerous pressure and velocity nulls present in the SWC, and to the antiresonances of the SWC itself. With the exception of these frequencies, the level of the reactive intensity measured by the u - u probe matches that of a p - u probe within 6 dB. This is consistent with the findings of Gabrielson *et al.*⁷ in which a 10 dB difference existed between measurements made by p - u and finite differencing p - p probes in a reactive field.

The u - u probe's acoustic intensity measurement methodology is viable. Measurements of reactive intensity require subtracting two nearly equal signals, which the u - u intensity probe does fairly accurately in the reactive field of the SWC. On the other hand, measurements of active intensity in a highly reactive field using a finite differencing technique, such as a p - p or u - u probe, are difficult to perform accurately. Figures 8(a) and 10(a) show a 20 dB difference in the active intensity levels measured by the u - u and p - u probes. Figure 7(a) indicates that this difference is associated either with the reference transducer phase matching, or the finite difference approximation employed by both the p - p and u - u probes. In future applications, the p - u probe is recommended.

ACKNOWLEDGMENTS

This work has been supported by the Office of Naval Research, Code 321 (Grant No. N00014-96-1-1173) and Code 333 (Grant Nos. N00014-94-1-0708 and N00014-99-1-0259). Particular appreciation is extended to Scott Littlefield and Dr. Kam Ng.

¹A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, Woodbury, NY, 1989).

²F. J. Fahy, *Sound Intensity*, 2nd ed. (E & F Spon, London, 1995).

³J. Y. Chung, "Cross-spectral method of measuring acoustic intensity without error caused by instrument phase mismatch," *J. Acoust. Soc. Am.* **64**, 1613–1616 (1978).

⁴K. W. Ng, "Acoustic intensity probe," U.S. Patent No. 4,982,375 (1 January 1991).

⁵J. A. McConnell, G. C. Lauchle, and T. B. Gabrielson, "Two geophone underwater acoustic intensity probe," U.S. Patent Disclosure (30 August 1995).

⁶T. B. Gabrielson, J. F. McEachern, and G. C. Lauchle, "Underwater acoustic intensity probe," U.S. Patent No. 5,392,258 (21 February 1991).

⁷T. B. Gabrielson, D. L. Gardner, and S. L. Garrett, "A simple neutrally

- buoyant sensor for direct measurement of particle velocity and intensity in water," J. Acoust. Soc. Am. **97**, 2227–2237 (1995).
- ⁸GS-14-L3 Geophone, Geospace Corp., 7334 Gessner, Houston, TX 77040.
- ⁹R. A. Serway, *Physics for Scientists and Engineers*, 3rd ed. updated (Saunders College, Philadelphia, PA, 1990).
- ¹⁰L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed. (Butterworth-Heinemann, New York, 1987).
- ¹¹C. B. Leslie, J. M. Kendall, and J. L. Jones, "Hydrophone for measuring particle velocity," J. Acoust. Soc. Am. **23**, 711–715 (1956).
- ¹²E-Cast F-8 Low Viscosity Epoxy Resin and Hardener 118, United Resin Corp., 4359 Normandy Ct., Royal Oak, MI 48073.
- ¹³K & B Glass Microballoon Filler, K & B Manufacturing, Inc., 2100 College Dr., Lake Havasu, AZ 86403.
- ¹⁴C. M. Harris, *Shock and Vibration Handbook*, 4th ed. (McGraw-Hill, New York, 1996).
- ¹⁵W. T. Thomson, *Theory of Vibration with Applications*, 3rd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1988).
- ¹⁶M. R. Spiegel, Ed., *Mathematical Handbook: Schaum's Outline Series*, 30th ed. (McGraw-Hill, New York, 1992).
- ¹⁷K. J. Bastyr, "A velocity gradient underwater acoustic intensity sensor," MS Thesis, The Pennsylvania State University, December 1998.
- ¹⁸L. E. Ivey, "Underwater electroacoustic transducers," Naval Research Laboratory Specifications Sheet, August 1994. The USRD is now affiliated with the Naval Undersea Warfare Center in Newport, RI.
- ¹⁹M. B. Johnson, "Reciprocity calibration in a compliant cylindrical tube," MS Thesis, Naval Postgraduate School, June 1985. DTIC Report Number ADA 158998.
- ²⁰M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction* (Acoustical Society of America, Woodbury, NY, 1993).
- ²¹Townsend Plastic, "Physical properties of townsend cast acrylic," Manufacturers Specification Sheet, dated 1998.
- ²²E. Skudrzyk, *The Foundations of Acoustics: Basic Mathematics and Basic Acoustics* (Springer-Verlag, New York, 1971).
- ²³J. A. McConnell, K. J. Bastyr, and G. C. Lauchle, "Sound-speed determination in a fluid-filled elastic waveguide," J. Acoust. Soc. Am. **105**, 1143 (1999).
- ²⁴PCB Model W353B66 Accelerometer, PCB Piezotronics, Inc., 3425 Walden Ave., Depew, NY 14043.
- ²⁵G. S. K. Wong and T. F. W. Embleton, Eds., *AIP Handbook of Condenser Microphones: Theory, Calibration, and Measurements* (AIP Press, New York, 1995).
- ²⁶Acoustech Corporation, P.O. Box 139, State College, PA 16804.
- ²⁷Reson Model TC4013 Hydrophone, Reson, Inc., 300 Lopez Rd., Goleta, CA 93117.
- ²⁸1201 Low Noise Preamplifier, Ithaco Corp., 735 West Clinton St., Ithaca, NY 14850.
- ²⁹Servo 150 Studio Amplifier, Samson Technologies Corp., Syosset, NY 11791.
- ³⁰J. S. Bendat and A. G. Pierson, *Random Data: Analysis and Measurement Procedures*, 2nd ed. (Wiley, New York, 1986).

Low-frequency ambient sound in the North Pacific: Long time series observations

Keith R. Curtis, Bruce M. Howe, and James A. Mercer

Applied Physics Laboratory, College of Ocean and Fishery Sciences, University of Washington, Seattle, Washington 98105

(Received 27 May 1998; revised 28 May 1999; accepted 8 July 1999)

Long-term statistics of ambient sound in an ocean basin have been derived from 2 years of data collected on 13 widely distributed receivers in the North Pacific. The data consist of single hydrophone spectra (1–500 Hz in 1-Hz bands) averaged over 170 s and recorded at 5-min intervals. Cumulative probability distributions of the ambient sound level show that for the open-ocean arrays at 75 Hz, sound levels are 3 dB higher than the median level 10% of the time and 6 dB higher 1% of the time. For the coastal arrays, sound levels are 7 dB higher than the median level 10% of the time and 15 dB higher 1% of the time. The clearest feature in many of the spectrograms is a strong annual cycle in the 15–22 Hz band with peak signal levels up to 25 dB above the sound floor; this cycle is attributed to the presence and migration of blue and fin whales. On average, whales are detected 43% of the time. Ships are heard 31%–85% of the time on the coastal receivers and 19%–87% of the time on the open-ocean receivers, depending on the receiver. On average, ships are detected 55% of the time. The correlation coefficient between the sound level in the 200–400 Hz band and wind speed, determined from satellite and global meteorological analysis, is on average 0.56 for the coastal receivers and 0.79 for the open-ocean receivers. For some receivers, the sound level in the 12–15 Hz band is correlated with the sound level in the 200–400 Hz band, with a correlation coefficient of 0.5. © 1999 Acoustical Society of America. [S0001-4966(99)02512-6]

PACS numbers: 43.30.Nb, 43.30.Pc [DLB]

INTRODUCTION

Understanding the variability of ambient sound in the ocean is essential for investigating air–sea interface processes, such as wind, rainfall, snowfall, and breaking waves,^{1–4} as well as for monitoring submarine seismic events and ship traffic.^{5–7} Understanding ambient sound variability is necessary for many naval applications. The effect of ambient and anthropogenic sound, e.g., shipping traffic, on marine animals has recently been of concern.

Ambient sound has traditionally been characterized by short-term sound measurements at many locations.^{8–10} The goal of the present study is to use long time series of ambient sound at fixed locations to extend our understanding of the variability of the sound field and the causes thereof.

In the frequency band 5–500 Hz, the most common sources of sound are seismic events, whales, ships, and wind-generated breaking waves. Blue and fin whales vocalize primarily in the range 15–35 Hz.^{11–13} Humpback whale calls are in the range 50–1000 Hz;¹¹ only the lower-frequency (15–35 Hz) sounds are addressed here. The most distinctive characteristic of ship-generated sound is the acoustic energy from propeller cavitation at blade-rate harmonics (the fundamental frequency is typically around 8 Hz); additional energy, both broadband and tonal, is produced by machinery and shaft rotation. The sound from breaking waves covers a wide band (at least 0.1–20 kHz), but the peak frequency ranges from 200 Hz to 2000 Hz, depending on the type of breaking wave.¹⁴ It has been shown that the sound level in the range 200–500 Hz is primarily a function of wind speed and correlates well with the energy dissipated by breaking waves.^{1,15,3,16} Sound level is not related to wave height,¹⁷

probably because nonbreaking waves contribute to the wave-height estimate. If swell is removed from the wave-height estimate, then the correlation of sound level with wave height increases significantly.¹⁷

Most observations of low-frequency ambient sound have been sparse and isolated in time and space. Exceptions are a 1-year time series (0.4–30 Hz) from hydrophones near Wake Island,¹⁸ a 3-year time series of sonobuoy data (at frequencies of 50, 100, 200, 440, 1000, and 1700 Hz),¹⁹ and a 2–3-year study (10–400 Hz) that used several of the same receivers used here (the last data, taken by Wenz²⁰ in 1963–1966, have only recently become available to the general public). Another relevant observation where the results of acoustic modeling of the meteorological conditions were correlated with the ambient sound is a 1-year time series (at 200–300 Hz) from the Greenland Sea.²¹

The work reported here is a component of the Acoustic Thermometry of Ocean Climate (ATOC) project.²² Concern that the acoustic signals used for measuring ocean temperature change might adversely affect marine mammals led to obvious questions such as, what fraction of the time does the ambient sound level exceed a given level, and how does the ATOC signal level compare with ambient sound levels, levels that were known (in an almost anecdotal fashion) to fluctuate significantly depending on time and location? The original motivation for the present work is that the data needed to answer these simple questions did not exist or were not available. The ambient sound time series reported here may also be used to address other related questions about shipping and marine mammal behavior.

This paper is organized as follows. A brief description

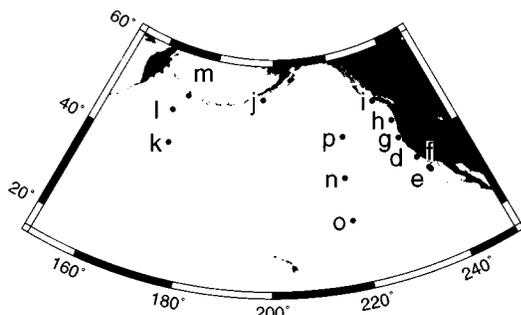


FIG. 1. Map showing the locations of receivers used in this study.

of the data acquisition and processing is given in Sec. I, followed in Sec. II by a general overview of the data. Spectrograms from various receivers and associated statistics of a general nature are presented. Representative spectra of ATOC receptions are compared with spectra of the background ambient sound. A description of whale and ship sound is given in Sec. III. A neural network is used to classify spectra according to the presence of each type of sound. This is followed by a discussion of wind data in Sec. IV. The sound in the integrated band 200–400 Hz is analyzed to determine the correlation with wind speed, using wind estimates based on satellite data from the Special Sensor Microwave Imager (SSM/I) and the National Center for Environmental Prediction (NCEP) for the Pacific Ocean. The final section discusses ramifications of these measurements and possible new avenues of study.

This paper is primarily descriptive in nature and focuses on the overall picture of ambient sound and its variability in time and space.

I. DATA ACQUISITION AND PROCESSING

Time series of ambient sound spectra have been collected using 13 widely distributed receivers, including U.S. Navy Sound Surveillance (SOSUS) arrays in the North Pacific (Fig. 1 and Table I). The measurement equipment at each site consists of two systems, the receiver (hydrophones, cable, and amplifier) and a signal-conditioning/data-acquisition system installed by the Applied Physics Laboratory, University of Washington, as part of ATOC.²²

The spectra have been corrected for the frequency response of the receiver using terminal sensitivity curves provided by the U.S. Navy. With one exception, no recent calibrations of the receiver systems exist. A calibration from 1 to 200 Hz of receiver *d* (Ref. 23) compares well with the original calibration. However, the measured absolute sound levels are obviously in error for many of the receivers. Therefore, all the spectra are plotted on a relative scale, where the median level at 75 Hz is arbitrarily shifted to 0 dB (i.e., the units are dB *re*: median level at 75 Hz). When comparisons

TABLE I. Locations of receivers *d*, *e*, and *f* (WGS84 coordinate system).

Receiver	Latitude	Longitude	Depth
<i>d</i> , Point Sur	36° 17.948' N	122° 23.631' W	1359 m
<i>e</i> , San Nicolas South	32° 54.913' N	120° 22.548' W	1106 m
<i>f</i> , San Nicolas North	33° 22.614' N	120° 36.913' W	1094 m

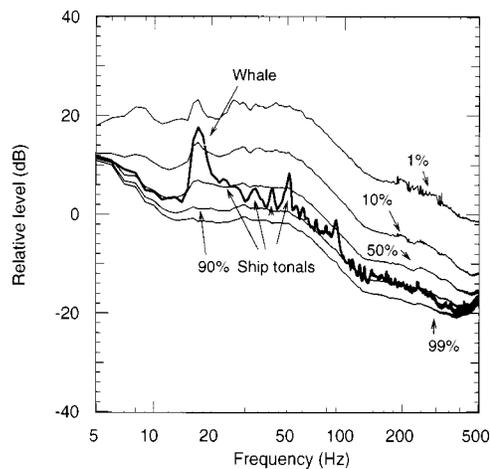


FIG. 2. Cumulative probability distribution function for receiver *d*, derived from all of the spectra from receiver *d*. The curves show the sound level, in decibels re median level at 75 Hz, that is exceeded 99%, 90%, 50%, 10%, and 1% of the time at a particular frequency. Overlaid on the cumulative probability distribution is a spectrum from October 1995 showing both a whale signal and a nearby-ship signal with tonals at intervals of roughly 8 Hz.

are made, residuals of the spectrum are used, by either subtracting out a median/mean level or removing a sound floor (as will be discussed in Sec. II).

The response of the receivers is given by the terminal sensitivity as a function of frequency. Signals are band-pass filtered between 2 Hz and 500 Hz, where the high-pass filter is due to ac coupling and the low-pass filter consists of a set of two filters. The first is a two-pole Butterworth filter with the -3 -dB point at 1000 Hz; the second is another two-pole Butterworth filter with the -3 -dB point at 500 Hz. Signal amplification and analog-to-digital conversion are independent of frequency from 5 to 1000 Hz.

At 5-min intervals, 170 s of data are sampled at 2000 points per second from three hydrophones. Each sound sample, or record, is subdivided into 10 groups of 32768 samples, for a total length of 163.84 s. Power spectra for each group are ensemble averaged and smoothed over 1-Hz bins from 0 to 500 Hz. The resulting 501-point spectrum is saved, and the original sound sample is discarded. Only data from a single hydrophone on each receiver are discussed here (in this context there is little difference between the data collected from the three hydrophones on a particular receiver). Outliers occur when hydrophones fail or when 60-Hz signals are present; the affected spectra are discarded. Large gaps occur in the time series because of equipment malfunctions on site or damage to the submarine cables connecting the hydrophones to shore. Small gaps occur during ATOC receptions and system administration tasks.

II. OBSERVATIONS—GENERAL OVERVIEW

One of the simplest ways to quantify variability is with a probability density function (PDF) and its integral, the cumulative probability distribution function (CPDF). The CPDF of 1-Hz-wide spectral levels (in decibels) for the data collected here was computed for the frequency range 5–500 Hz. The results for receiver *d* are given in Fig. 2. The curves

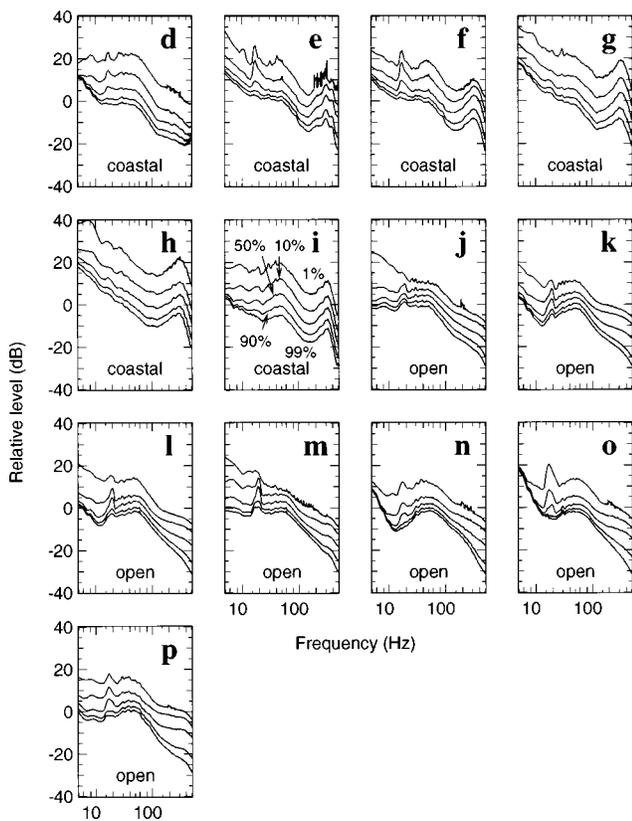


FIG. 3. Cumulative probability distribution function for every receiver, derived from the spectra for all 13 receivers. The curves show the sound level, in decibels re median level at 75 Hz, that is exceeded 99%, 90%, 50%, 10%, and 1% of the time at a particular frequency. Receivers *d*–*i* are coastal, and the other receivers are located in the open ocean. Since the time sampling was somewhat irregular, with record lengths between 1 and 2.7 years, there may be biases due to seasonal and other effects.

show the sound levels that are exceeded a certain fraction of time as a function of frequency relative to the median level at 75 Hz. The measured median levels at 75 Hz for receivers *j*, *k*, *l*, *m*, and *n* are 81.2, 80.5, 80.2, 82.6, and 83.7 dB, which are within a reasonable range of values for ocean ambient sound;¹⁰ the measured median levels for the other receivers differ from these by more than 10 dB and are considered unrealistic. The CPDFs for all the receivers are shown in Fig. 3. The “whale” peak near 17 Hz is evident even in the median curve; such a peak indicates that the sound level at 17 Hz is greater than the levels at other frequencies for the same percentage of time. The relative increase in spectral level in the 30–100 Hz band reflects the ubiquity of ship sound. The skewness of the probability distribution varies with frequency. For the lower frequencies, the distribution is positively skewed, similar to a Rayleigh distribution.²⁴ For the higher frequencies, the distribution is almost symmetrical.

The spectra for coastal receivers *e*–*i* have a peak at about 350 Hz (Fig. 3) which has a level similar to the band where ship sound is dominant, i.e., 30–100 Hz. Receiver *d* is also near the coast but shows no peak of similar strength. Interestingly, ambient sound spectra measured by Wenz²⁰ more than 30 years previously, using some of the same coastal receivers as used in this study, show that the ship sound peak near 50 Hz was about 10 dB higher than the level

at 300 Hz. Also sound due to ship traffic has increased on the long term.²⁵ At this time we do not have an explanation for the anomalous strength of the peaks at 350 Hz in the spectra for receivers *e*–*i*.

Since our primary concern is the variability in sound level with time, we first subtracted the “sound floor” (the site-dependent, frequency-dependent threshold that is exceeded 99% of the time) to make this variation clearer. When looking at the resulting spectrogram for receiver *o* (Fig. 4), the eye is first drawn to the feature at 15–22 Hz that has an annual cycle with a peak-to-peak amplitude of approximately 25 dB. This feature is attributed to the vocalizations of fin whales. The ridge is very clear and peaks roughly at 17 Hz. Although this feature covers 7 Hz, the signal is loudest at roughly 17 Hz. The variation in bandwidth is most likely due to the nature of the different whale calls. In spectrograms for receivers *d*, *e*, and *f*, harmonics of 17 Hz can also be seen which are characteristic of blue whales.^{12,26,27} The lack of harmonics at receiver *o* [Fig. 4(bottom)] suggests that only fin whales are vocalizing near this receiver. It is not known if these vocalizations are from many whales, a few whales vocalizing nearly continuously, or some combination thereof. Calls of other whale species are not prominent in these data, and no work has yet been done on identifying signatures of other species.

Another noticeable feature in Fig. 4 is intermittent periods of increased sound level that are highly correlated over 200–500 Hz and have time scales of 1 day. These high sound levels are caused by high winds associated with storms. To obtain a different perspective on the variability of ambient sound, we also computed the covariance matrix, $\mathbf{C} = \mathbf{C}(f_1, f_2)$, which gives the covariance of the spectral levels at frequencies f_1 and f_2 . For receiver *k* [Fig. 5(top)], the variance at 200–500 Hz is highly correlated over the whole band (correlation coefficient >0.8) and is also correlated with the energy in the 12–15 Hz band (correlation coefficient 0.6). It has been suggested^{18,20} that this low-frequency component is associated with distant wind events; using data from hydrophones near Wake Island, McCreery *et al.*¹⁸ obtained correlation coefficients of 0.66 and 0.7 between wind speed and the sound levels at 10–12 Hz and 12–14 Hz. While shipping noise is not very evident in the record for receiver *k*, the whale component at 17 Hz is clear, with no correlation with other frequencies. In contrast, the covariance for receiver *i* [Fig. 5(bottom)] is largely dominated by shipping, with the fundamental frequency (9 Hz), the higher harmonics, and the wide-band noise quite evident.

In the high-temporal-resolution spectrograms, the presence of tonals at ship blade rates shows up as bright lines, typically at the lower harmonics (8 Hz, 16 Hz, 24 Hz, etc.). The lines are too intermittent to show up in the long-term averaged spectrograms because the duration of a ship event is typically about an hour. In high-resolution spectrograms, one can see a correlation between the ship tonals and the broadband sound associated with them.

In the last few years there has been considerable interest in the effect of manmade sounds on marine mammals. One of the original motivations for this study was to compare the ATOC signal with ambient sound. ATOC signals are broad-

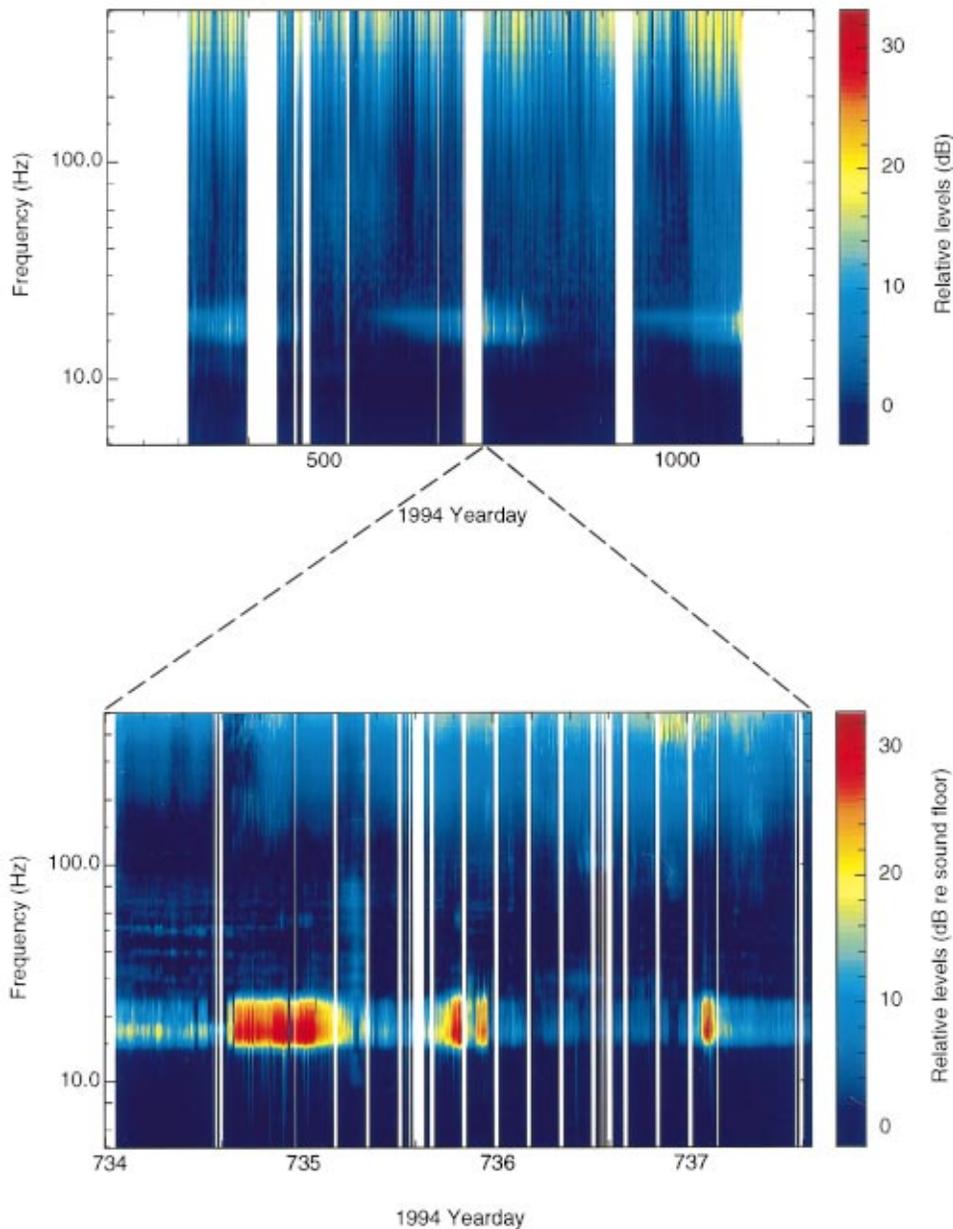


FIG. 4. Spectrogram measured on open-ocean receiver *o*. The spectrogram is relative to the sound-floor spectrum (i.e., the sound level that is exceeded 99% of the time in Fig. 3). (top) Spectrogram for November 1994–January 1997. Because of the amount of data involved, the data displayed are 1-day averages. (bottom) Spectrogram measured on receiver *o* for about 3 days in 1994, showing higher temporal resolution. Each individual 3-min spectrum is plotted.

band *m*-sequences²⁸ with a center frequency of 75 Hz, a bandwidth of 37.5 Hz, and a source level of 260 W, or 195 dB *re*: 1 μ Pa at 1 m. The range from the ATOC source on Pioneer seamount to receiver *d* is 148 km. Measurements of the ATOC signal at receiver *d* show that the ambient sound at receiver *d* is louder than the ATOC signal about 8% of the time, or 120 min per day (Fig. 6). For the particular spectrum shown, shipping and possibly fin or blue whale harmonics dominate the spectrum. At a range of 34 km (18 nautical miles) from the source, the ambient sound level would be greater than or equal to the ATOC signal level 2% of the time (assuming cylindrical spreading and the same ambient sound statistics as along the path between the source and receiver *d*). For an average ATOC source duty cycle of 2%, or 30 min per day, the ATOC signal level is not anomalous (i.e., higher than the range of ambient sound levels) until the range to the ATOC source is less than 34 km.

In the following sections, we discuss the statistics of ambient sound due to whales, ships, and wind in more detail.

III. WHALE AND SHIPPING SOUND

Fin whales produce series of pulses which, after averaging, show as a peak in the spectrum at 15–22 Hz.²⁹ Blue whales produce long patterned sequences of a part A (amplitude modulated series of pulse) and a part B (downsweep with strong harmonics) signal.^{30–34} This is evident as a peak at 15–22 Hz, as well as harmonics at 34, 51, and 68 Hz which can be quite visible in a spectrogram. Figure 7 shows receptions at 17 Hz, which was used to represent whale vocalizations. Some receivers (*n, o, p*) show large changes with time, of the order of 25 dB peak to peak, while others (*g, h, i*) show little or no changes. The northern coastal receivers (*h* and *i*) show a small seasonal variation at 17 Hz.

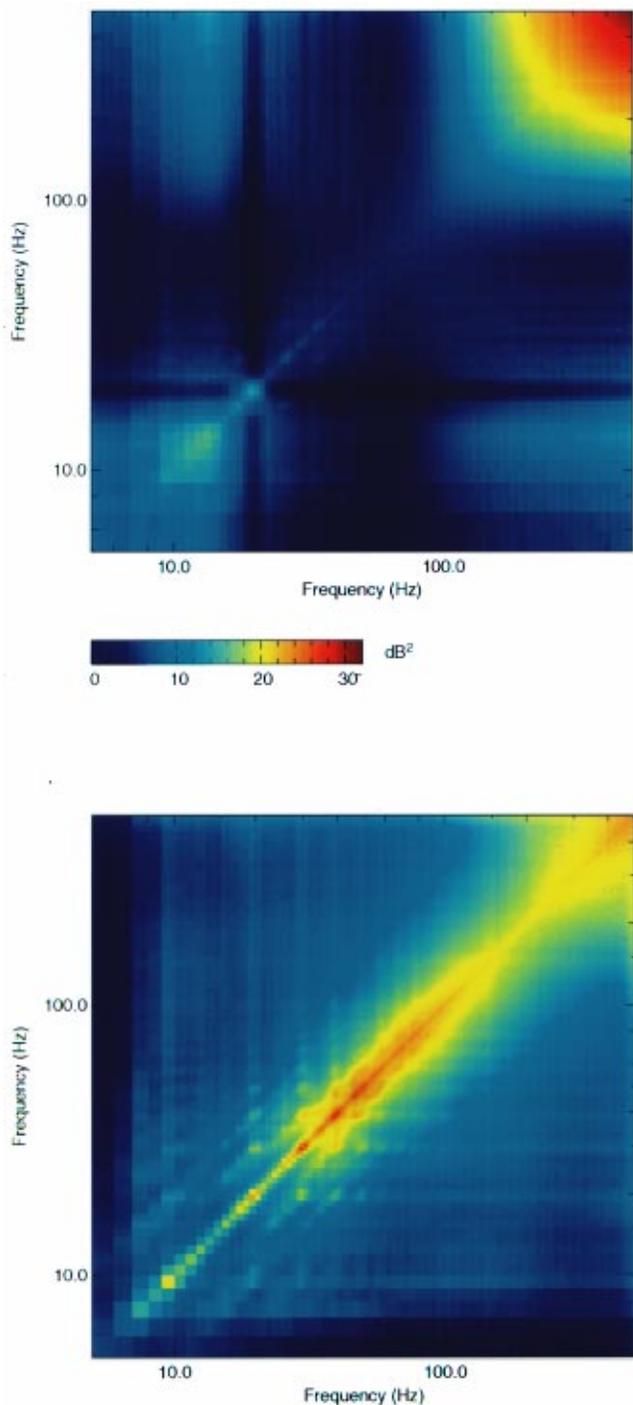


FIG. 5. Sound-level power-spectrum covariance matrices computed for (top) receiver *k* and (bottom) receiver *i* when using all the data. The color scale is in decibels squared.

Although not shown, only the coastal receivers off California (*d, e, f*) show evidence (i.e., harmonics) of a strong blue whale signature.

Preliminary work has been done on classifying spectra based on the presence of whale and shipping sound. In an experiment run on the Stuttgart Neural Network simulator,³⁵ we trained a simple neural network to classify whether whale or ship sound was present in a spectrum. The experiment utilized a two-layer feed-forward neural network with an input layer consisting of normalized spectral levels from 5 to

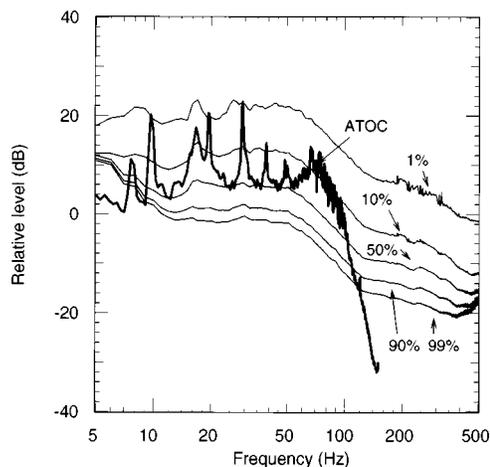


FIG. 6. Cumulative probability distribution function for ambient sound levels at receiver *d* (as in Fig. 3) along with the average spectra over a 20-min ATOC reception arriving from Pioneer Seamount 148-km distant. The ATOC data were low-pass filtered at 114 Hz.

104 Hz in 1-Hz increments. The average spectral level was removed, and the input values were then normalized with the peak value before being used. The training data were selected subjectively; spectra were identified by eye as to whether ship or whale (blue and/or fin) signatures were present or absent. Sample spectra were collected from the data sets for each receiver and then combined to form the training and validation sets. The network correctly classified 92% of the spectra used in training and 90% of the validation spectra that were held in reserve for evaluating the performance of the network (Fig. 8). The neural network produces a continuous measure of detection, $0 < D < 1$, with a 0 being a perfect nondetection and a 1 being a perfect detection. We take $D > 0.5$ to indicate a detection.

Because the training and validation data were selected by eye, it is quite possible that the subjective classifications were not always correct. An attempt was therefore made to avoid overfitting the network, i.e., to accept a less than perfect performance on the training data. By avoiding overfitting, it was hoped that the neural network would generalize to other spectra that it was not trained on.³⁶ As can be seen in Fig. 8, the resulting network can detect whether a spectrum contains ship or whale components.

Figure 9 shows the neural network response when clas-

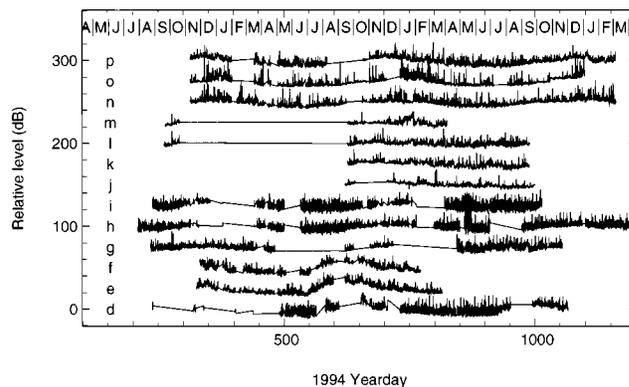


FIG. 7. Time series at 17 Hz, the peak of the whale band, for the various receivers. Each curve is shifted 25 dB.

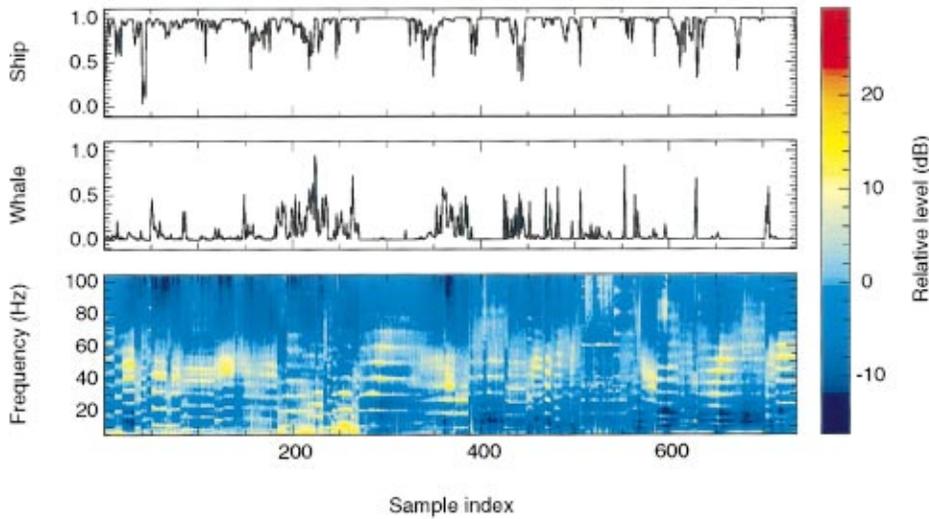
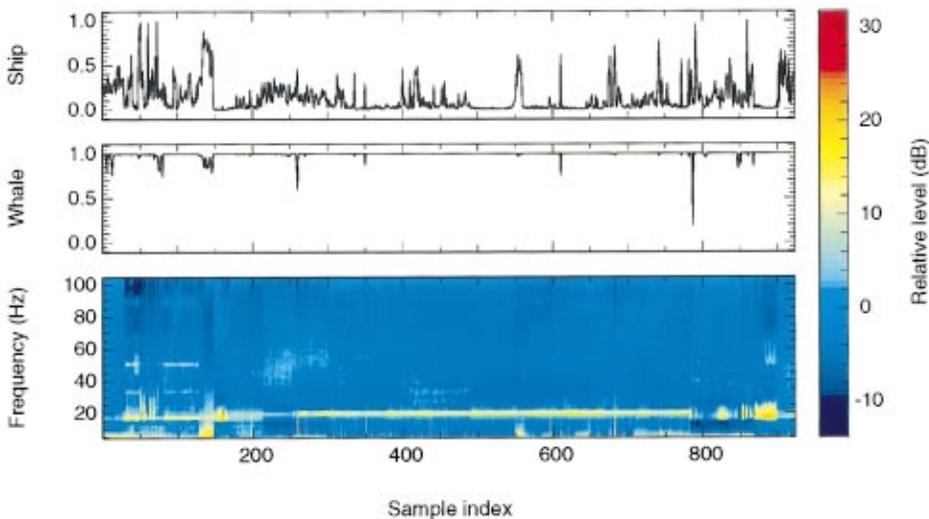


FIG. 8. Subjectively selected data (bottom panels) used to train a neural network to detect ship events (top panels) and whale events (middle panels). (top) Ships/no whales; (bottom) no ships/whales.



sifying whale receptions at receivers e and p . Receiver e shows a definite seasonal signal [Fig. 9(top)]. The “whale season” starts quite abruptly in mid-July and ends in January. For receiver p , the whale season starts in late August and ends in March [Fig. 9(bottom)]. Surprisingly, there appear to be more ship detections for receiver p , an open-ocean receiver, than for receiver e , a coastal receiver. Receiver p is near the United States–Far East shipping lane, whereas receiver e is seaward of San Nicolas Island and may be shielded from shipping lanes by local bathymetry.²⁰

The whale detections have a distinct bimodal appearance: in most cases, the algorithm is sure that whale vocalizations are or are not present [Fig. 10(a)]. In general, the whale season starts in the summer and ends in the winter; in most cases, the beginning and end are distinct events. On average, whales are detected 43% of the time (Table II). The average increase in level at 17 Hz varies from 2 dB at receiver g to 9 dB at receiver o .

The variation in the network’s response is much more

continuous for ship detections [Fig. 10(b)]. Perhaps this is not so surprising given the wider variety of ships compared to whales. On average, ships are detected 55% of the time; this varies from a low of 18% for receiver o to a high of 87% for j . As with receiver p , receiver j is probably hearing the shipping going between the United States and the Far East, whereas receiver o is an open-ocean receiver between California and Hawaii where little ship traffic is expected. A seasonal cycle may be present for receivers n and o [Fig. 10(b)].

IV. WIND SOUND

In Fig. 4 the sound due to high wind is evident as events several days long in the frequency range of 150–500 Hz. Since the wind speeds above the receivers vary seasonally, this suggests that the ambient sound level would also vary seasonally. A seasonal cycle is found in sonobuoy data (of the order of 10 dB peak to peak)¹⁹ and in the earlier work of

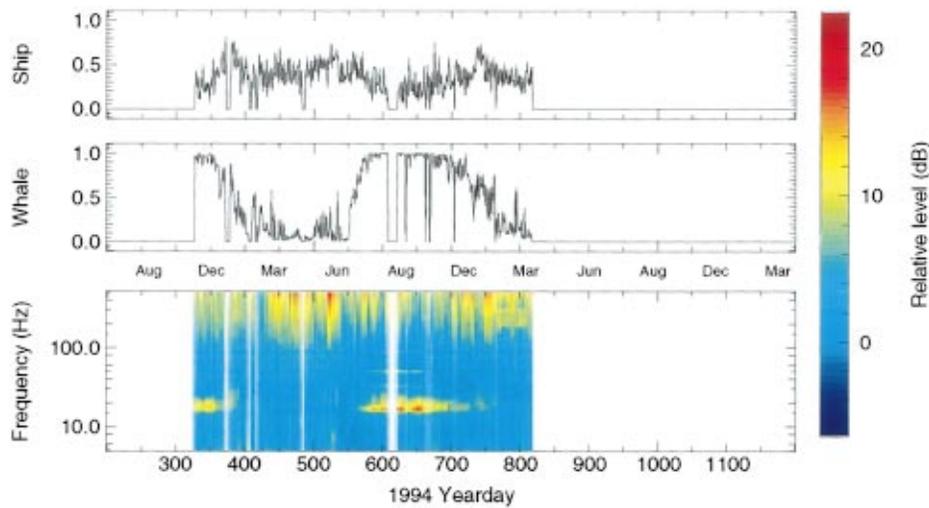
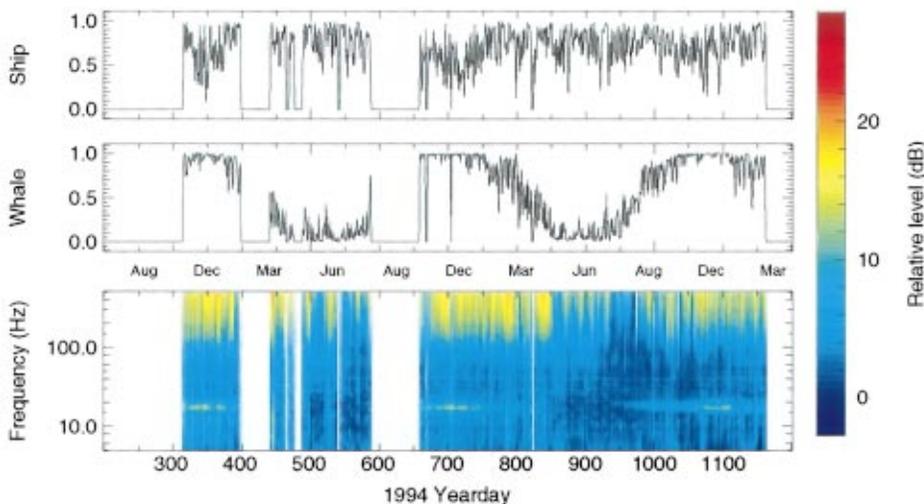


FIG. 9. Responses of neural network for whale and shipping detection along with corresponding spectrograms for (top) receiver *e* and (bottom) receiver *p*.



Wenz (5 dB peak to peak on some receivers).²⁰ However, a seasonal cycle is not evident in the data for most of the receivers (exceptions may be *n* and *p*), possibly because of sizable gaps in the time series.

In the following subsections we first describe a very simple model to predict the sound level due to wind. This model integrates the effect of the wind over the ocean surface. The variations of the model-predicted sound level are compared to the variations of the observed sound level in the 200–400 Hz band. We also compare the variation in sound level measurements in the 200–400 Hz band with independent estimates of wind speed.

A. Zeroth-order model

In the zeroth-order approximation, sound propagates in a straight line from a source on the surface to a receiver in the water. Thus any surface sound source, such as breaking waves, contributes to the signal received at a hydrophone located on the ocean floor. As the distance from the hydrophone increases, the signal received from any single source

decreases, but the area of possible contributing sources increases. Because of chemical absorption of sound, there is a range at which the presence or absence of sources ceases to be significant. We consider this range the effective listening radius (see the Appendix).

The effective listening radius imposed by absorption is an upper bound, as the sound also undergoes attenuation due to scattering by internal and surface waves. This upper bound could be used to truncate numerical integration in ocean acoustic propagation models. Thorp's expression for the attenuation coefficient³⁷ gives the effective listening radii for frequencies of 200 Hz and 400 Hz as 1650 km and 650 km, respectively (Fig. 11).

Because significant energy is contributed from sources at long ranges (until the energy is attenuated by chemical absorption), such sources should be considered when predicting the received level I_r . If sources with intensity I_i are specified on a grid, then a crude estimate of the intensity at a receiver is

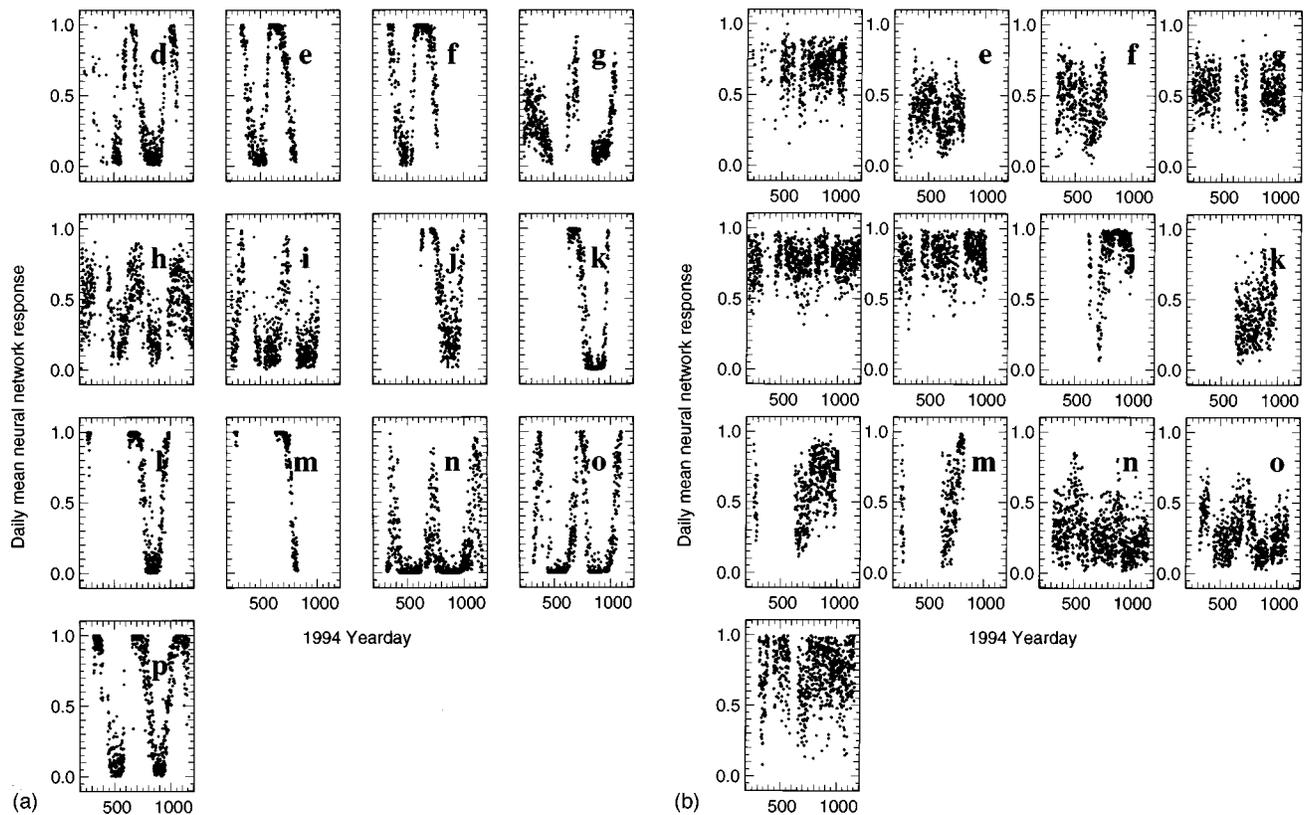


FIG. 10. Daily mean network response for (a) whale-produced sound and (b) ship-produced sound for each receiver.

$$I_r = \sum_i \alpha_i I_i, \quad (1)$$

where α_i is the attenuation due to spreading and absorption along a path connecting the receiver and grid point i . If the I_i values are known from a specified distribution of wind speed, then this method should be superior to simply using the intensity predicted from the wind speed above the receiver to obtain I_r .

TABLE II. Percentage of time that ships and whales were detected by the neural network.

Receiver	Ships (%)	Whales (%)
<i>d</i>	71.6	38.1
<i>e</i>	31.4	48.0
<i>f</i>	44.2	58.4
<i>g</i>	51.8	24.5
<i>h</i>	84.9	40.6
<i>i</i>	84.1	25.3
<i>j</i>	86.9	51.6
<i>k</i>	32.9	43.0
<i>l</i>	60.6	58.0
<i>m</i>	52.3	74.9
<i>n</i>	22.2	15.7
<i>o</i>	18.5	28.2
<i>p</i>	76.4	59.3
all	55	43

B. Long-term comparisons: SSM/I winds and NCEP winds

Wind fields have been obtained using data from two or three satellites equipped with Special Sensor Microwave Imagers (SSM/Is) and from a reanalysis project at the National Center for Environmental Prediction (NCEP) (where data gathered over the past few decades have been reanalyzed using one consistent methodology). The wind speeds determined from the SSM/I data cover a 50-km-wide swath and are available twice a day, but values are not available for pixels where there is rainfall. The satellite orbit is sun-

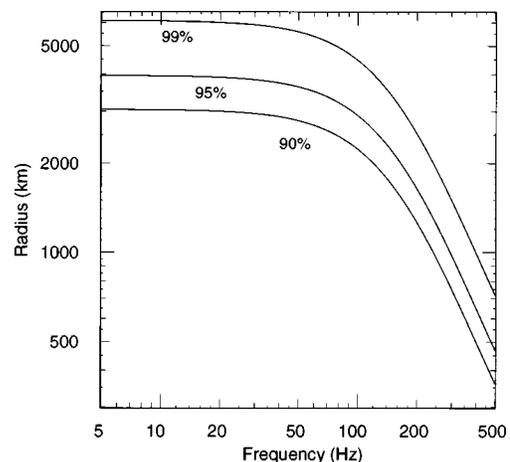


FIG. 11. Effective listening radius R_e plotted as a function of frequency; $(R_e \cdot f) = \beta I(\infty, f)$, where β has values of 0.90, 0.95, and 0.99. This assumes a homogeneous distribution of sound sources over the ocean surface.

TABLE III. Correlations of SSM/I and NCEP winds with observed sound levels.

Receiver	SSM/I			NCEP		
	<5 m/s	>5 m/s	+Model	<5 m/s	>5 m/s	+Model
<i>d</i>	0.11	0.34	0.29	0.01	0.35	0.29
<i>e</i>	0.30	0.61	0.72	0.35	0.62	0.73
<i>f</i>	0.23	0.63	0.68	0.40	0.64	0.72
<i>g</i>	0.23	0.49	0.49	0.06	0.71	0.65
<i>h</i>	-0.02	0.42	0.41	0.08	0.53	0.49
<i>i</i>	0.07	0.39	0.46	0.21	0.48	0.49
<i>j</i>	0.05	0.72	0.65	0.09	0.69	0.75
<i>k</i>	-0.01	0.82	0.66	-0.01	0.74	0.82
<i>l</i>	-0.06	0.82	0.60	0.05	0.75	0.74
<i>m</i>	0.07	0.84	0.60	0.05	0.71	0.75
<i>n</i>	0.15	0.79	0.73	0.20	0.83	0.86
<i>o</i>	0.18	0.73	0.69	0.25	0.66	0.79
<i>p</i>	0.38	0.83	0.72	0.20	0.75	0.84
all		0.65	0.59		0.65	0.69

synchronous with a period of 102 min. The SSM/I data files were processed using the Wentz algorithm³⁸ and formatted into quarter-degree pixels (courtesy of Remote Sensing Systems).

The NCEP wind data have a 12-h time step and are formatted into approximately 1° pixels. Unlike the SSM/I wind data, the NCEP data do not contain pixels with missing data. The NCEP data were thus easier to use, but in order to ensure that the same time scales were being compared, the acoustic data were low-passed filtered (over 12 h) before being compared with the NCEP data. The NCEP analysis does not include SSM/I data.

The observed sound levels, integrated over the band 200–400 Hz, were compared with the different wind products in two ways: The sound levels measured at each receiver were compared with wind estimates derived from the overhead pixel, and with the levels predicted when using a simple model incorporating the SSM/I or NCEP wind estimates. The correlations between the observed sound levels and the NCEP winds estimated from the overhead pixels are generally the same as the corresponding correlations for the SSM/I

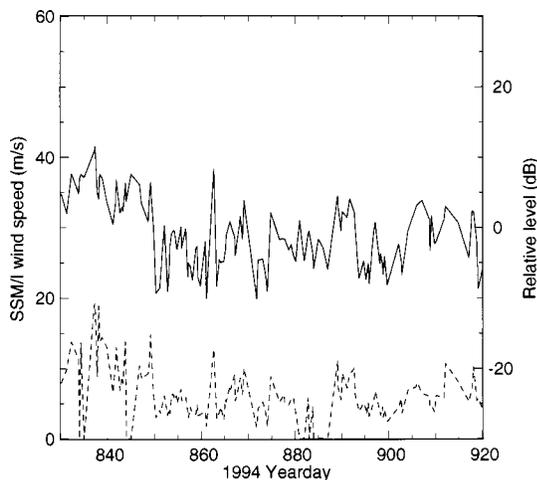


FIG. 12. SSM/I wind speed (dashed line) compared with observed sound level (solid line) at receiver *p* at times corresponding to the SSM/I measurements.

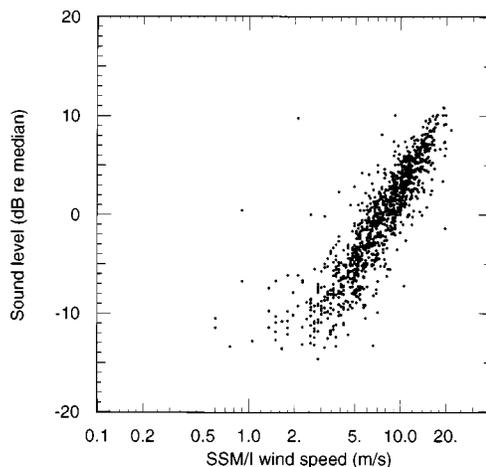


FIG. 13. Comparison of sound level received in the 200–400-Hz band at receiver *p* and SSM/I wind speed estimated from the pixel over the receiver.

winds (Table III). There is slightly more scatter in the relationship between NCEP wind speed versus sound level than in SSM/I wind speed versus sound level. For each receiver, correlations were computed for when the wind speeds were greater than 5 m/s and when they were less than 5 m/s. In general, the time series of wind speed in the overhead pixel of any receiver correlated well with the integrated sound level in the band 200–400 Hz (Fig. 12).

Figure 13 shows the relation between the sound level and the logarithm of the wind speed estimated from the SSM/I data. There is a clear difference in the slope of the relation for wind speeds less than 5 m/s and for wind speeds over 5 m/s. Below 5 m/s there is little if any correlation. For the SSM/I data, receivers *h*, *i*, *j*, *k*, *l*, and *m* had correlations less than 0.1 at low wind speeds (see Table III). Receivers *e* and *p* had the highest correlations at low wind speeds, 0.3 and 0.4, respectively. This suggests that for low wind speeds and frequencies of 200–400 Hz the source level could be considered a constant or at least unrelated to wind speed.

Above 5 m/s, receivers *d*, *g*, *h*, and *i* had low correlations with the SSM/I winds, less than or equal to 0.5; the balance of the receivers had correlations between 0.6 and 0.8. The correlations reached the 0.8 level for receivers *k*, *l*, *m*, *n*, and *p*. Using NCEP data, all the receivers except *d*, *h*, and *i* have correlations in the range 0.6–0.8.

Some of the overhead comparisons with SSM/I data are poor, specifically *d*, *g*, *h*, and *i*. Those receivers are near land, and the presence of nearby land is likely contaminating the SSM/I wind estimate. The NCEP winds at the land/sea boundary have larger errors also. For receiver *d*, the number of SSM/I data points is small, and that would limit the correlation as well.

When running the simple model to predict the received sound level due to wind distributed over the ocean surface, we used SSM/I pixels within 1000 km of a receiver to find wind speeds at several time periods during a day. The wind speeds were then averaged over all data available for that day for a particular pixel to produce a daily average for that pixel. Using the Chapman and Cornish¹⁵ relationship ($NL = B + 20\gamma \log v$, where NL is the sound level, v is the wind speed, and the other parameters (γ, B) are determined em-

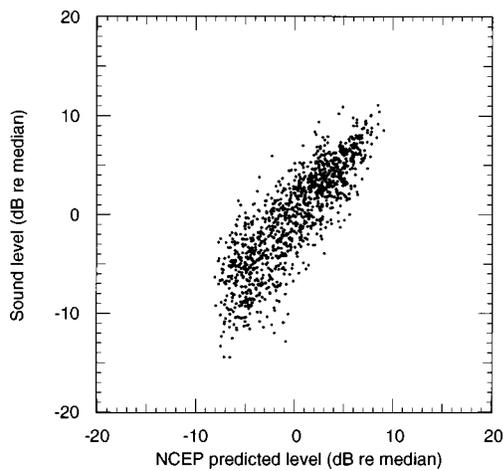


FIG. 14. Comparison of sound level received in the 200–400-Hz band at receiver p and sound level predicted using the simple model and NCEP wind speed fields at receiver p .

pirically) between wind speed and sound source level, we converted the daily averaged wind speed field to a sound intensity field and calculated the sound level at a receiver from Eq. (1). The time series of observed sound levels were low-passed filtered (1 day for SSM/I and 12 h for NCEP) and compared with the levels predicted from the SSM/I and NCEP wind fields. Numerical results are given in Table III, and the results for the NCEP winds and receiver p are shown in Fig. 14. The NCEP/model values correlate slightly better than the NCEP wind speeds in the overhead pixel with the observed levels; the average correlations are 0.69 and 0.65, respectively. However, the reverse is true for the corresponding SSM/I comparisons, with wind speed in the overhead pixel being correlated slightly better than the SSM/I model values; the average correlations are 0.65 and 0.59, respectively.

For all the receivers, the model predictions when using the NCEP wind fields are better than the corresponding SSM/I predictions. The model predictions when using the NCEP winds had better correlations than the predictions when using the SSM/I winds for receivers $e, f, i, j, n, o,$ and p , sometimes as much as 0.1 better. In some cases, though, the model/NCEP correlation was less than that for the other receivers, as much as 0.1 less.

According to the model, almost all the sound is from distant and not overhead sources. In fact, the depth of the receiver is almost irrelevant, as the ratio of the energy received from sources within 12 km (one SSM/I pixel) to the energy received from sources beyond 12 km is about 1:25 irrespective of any reasonable receiver depth. If the local sound sources were in-phase and the distant sources were out-of-phase, then the local sources might dominate the measured sound level. The correlations for the overhead data and for the model data are similar; this probably indicates that the correlation length of the wind field is large.

It is conceivable that high-frequency noise associated with shipping might affect the correlation estimates. Using the neural network described earlier, we removed the spectra in which shipping sound was detected ($D > 0.5$) from the wind data set before correlating the measured data sound

levels with the SSM/I wind time series in the pixel above the receiver. For receivers $j, k,$ and l , the presence of shipping was so high that the amount of data available for the correlation estimate was significantly reduced, and the correlation between the observed level and the SSM/I wind was smaller. Otherwise, the presence or absence of spectra with shipping noise did not affect the correlation between the observed sound level and wind speed.

V. DISCUSSION AND CONCLUSIONS

In this paper we have investigated the long-term statistics of whale, shipping, and wind-generated sound in the North Pacific Ocean using data from many geographically distributed, fixed receivers. The CPDFs that are now available are powerful reference data for designing acoustic experiments.

The whale component of ocean ambient sound is only now becoming more fully appreciated as one of the major sources of sound in the ocean. The fact that, on average, it is detected in 43% of the time in the receiver spectra (the range is 18%–59%, depending on the receiver) indicates its ubiquity. On some receivers the annual variation in level can be as much as 25 dB; on average, the difference is 2–9 dB. The distinct seasonality of the signal suggests questions such as, where are the whales during the quiet times (March–July)? Are they simply not vocalizing? The latter strikes an incongruous chord given their apparent social character. Ideally, there would be a global network of acoustic sensors to monitor migration patterns and to determine whether effects such as El Niño perturb the migrations.

If the sound of whales is ubiquitous, the sound of ships is even more so. On average, ships are detected in 55% of the time in the spectra (the range is 18%–87%, depending on the receiver). The levels of ship sounds received on single hydrophones are typically 10 dB above the background level [Fig. 8(a)], but the levels cover a wide range (Fig. 3).

The wind analysis given here is a large-scale comparison between global-scale wind products (satellite data and meteorological analyses) and measured acoustic sound levels. The analysis indicates that the forward problem of predicting the sound level at a receiver is working reasonably well. In fact, we consider the average correlation of 0.69 between the measured sound levels and the NCEP model results to be quite good. The NCEP product has no formal error bars associated with it, but errors of 1–2 m/s are expected over the open ocean and even larger errors are expected near the coast, especially since so few direct oceanic wind measurements are assimilated into the NCEP meteorological model. Also, the estimated errors for SSM/I winds are 1 m/s and are probably larger near the coast. The average NCEP model correlations are 0.56 for the coastal receivers and 0.79 for the open-ocean receivers.

A natural next step in this analysis will be to use a more sophisticated acoustic propagation model that includes frequency dependence to see if the results are affected. Furthermore it would be worthwhile to compare the ambient sound levels with a modeled surface wavefield as has been done previously with good results.²¹

The formula given in Eq. (1) for estimating the received intensity given the distribution of noise sources suggests that it could be inverted to obtain the surface intensity if there were a sufficient number of data points to make the problem “well” determined. We are investigating using beam data from each receiver as one possible way of obtaining sufficient data to reconstruct the distribution of the surface sound field. The distribution of vocalizing whales and shipping could be more successfully determined using the above scheme because of the lower attenuation at lower frequencies.

At the receiver closest to the ATOC source (receiver d , 148 km away), the ambient sound is louder than the ATOC signal, on average, 2 h per day; the ambient sound and the ATOC signal would be of comparable magnitude only at ranges less than 34 km from the transmitter. It is clear that ship sound is a major source of sound in the band 15–100 Hz, and yet the cumulative effect of years of such sound on marine life is unknown.

A failing of the present data set is the lack of absolute calibration of the system. Performing a direct calibration is costly and difficult. It may be worthwhile to investigate the use of the “Holu Spectrum” between 0.4 and 6 Hz to set the absolute level.¹⁸ However, additional questions would still remain; for instance, is the “hump” at 350 Hz shown in Fig. 3 for most of the coastal receivers real or an artifact of the present measurement system? The average spectral curves plotted by Wenz²⁰ start to increase with frequency at 200 Hz, not at 100 Hz as we observe, and show no abrupt peaks at 350 Hz such as we observe.

We did not address ambient sound of seismic origin here, as the data collected consist of average spectra which do not show the details of the typical low-frequency, transient seismic signal. Seismic signals in the ocean have been treated by other researchers.^{5,6} It is clear that in the future, as we try to split the spectrum further into its components, transient signals, seismic, and others, as well as signals from other whale species, will have to be included in the analysis.

Currently we do not have an explanation for the 12-Hz wind signal observed at receiver k . This is an area for investigation by future research.

With the present data it is difficult to extrapolate results away from the the point receivers. We expect that by using data from beamforming arrays, we will be able to improve our spatial resolution and perhaps even attempt to invert Eq. (1) to obtain the geographic distribution of sound sources. In this context, the 12-Hz wind signal and the 17-Hz whale vocalizations will play a role.

ACKNOWLEDGMENTS

We appreciate the help of the many people who contributed to the collection of these data, including S. Leach, D. Reddaway, and S. Weslander, APL-UW; K. Metzger, University of Michigan; C. Miller, Naval Post-graduate School; J. Peeples, RPI; and the staff at the Whidbey Island Naval Facility. Helpful discussions with C. Clark, K. Fristrup, A. Frankel, and D. Mellinger of Cornell University and J. Calambokidis of Cascadia Research are gratefully acknowledged. C. Clark, B. Dushaw, and three anonymous reviewers

provided helpful comments that lead to improvements of this manuscript. This work is supported by the Acoustic Thermometry of Ocean Climate program sponsored by the Strategic Environmental Research and Development Program (SERDP) through the Advanced Research Projects Agency, and by the Ocean Acoustics Program of the Office of Naval Research.

APPENDIX

Consider a very simplified ocean (no convergence zones, etc.), where the intensity per unit surface area is I_0 and a signal is attenuated by two possible mechanisms: loss due to spreading (spherical or cylindrical) and loss due to absorption. Although representing the surface sources as dipoles is more realistic and is necessary for a hydrophone near the surface,^{39,16} for simplicity we will ignore the ray angle from the source to the hydrophone and consider the surface sources to be monopoles. If we represent the loss mechanism as spherical spreading out to some transition radius r_T and as cylindrical spreading for greater radii, the expression I_r for the intensity received at the hydrophone is

$$I_r = I_0 \int_0^{2\pi} \int_d^{(r_T^2 + d^2)^{1/2}} \frac{1}{2s} \exp(-\alpha s) ds d\theta + a I_0 \int_0^{2\pi} \int_{(r_T^2 + d^2)^{1/2}}^{(R^2 + d^2)^{1/2}} \frac{1}{2} \exp(-\alpha s) ds d\theta, \quad (A1)$$

where $a = (r_T^2 + d^2)^{-1/2}$, s is the slant range ($s^2 = r^2 + d^2$), and α is the chemical absorption at the frequency f of interest.

The solution is then

$$I_r(R, f) = 2\pi I_0 \left\{ \left[\ln(\alpha s) + \sum_{k=1}^{\infty} \frac{(-\alpha s)^k}{k \cdot k!} \right]_{(r_T^2 + d^2)^{1/2}}^{(R^2 + d^2)^{1/2}} + \frac{a}{2} \left[-\frac{\exp(-\alpha s)}{\alpha} \right]_{(r_T^2 + d^2)^{1/2}}^{(R^2 + d^2)^{1/2}} \right\}. \quad (A2)$$

As the first term is a constant, the second term shows that $I_r(R, f)$ has a limit as $R \rightarrow \infty$. Thus it is possible to define an effective listen radius R_e such that $I_r(R_e, f) = 0.95 I_r(\infty, f)$.

¹S. Vagle, W. G. Large, and D. M. Farmer, “An evaluation of the WOTAN technique of inferring oceanic winds from underwater ambient sound,” *J. Atmos. Oceanic Technol.* **7**, 576–595 (1990).

²J. A. Nystuen, “Rainfall measurements using underwater ambient noise,” *J. Acoust. Soc. Am.* **79**, 972–982 (1986).

³J. A. Nystuen, C. C. McGlothlin, and M. S. Cook, “The underwater sound generated by heavy rainfall,” *J. Acoust. Soc. Am.* **93**, 3169–3177 (1993).

⁴S. McConnell, M. P. Schill, and J. G. Dworski, “Ambient noise measurements from 100 Hz to 80 kHz in an Alaskan fjord,” *J. Acoust. Soc. Am.* **91**, 1990–2003 (1992).

⁵C. G. Fox, R. P. Dziak, H. Matsumoto, and A. E. Schreiner, “Potential for monitoring low-level seismicity on the Juan de Fuca Ridge using military hydrophone arrays,” *Mar. Technol. Soc. J.* **74**, 22–30 (1994).

⁶C. E. Nishimura and D. M. Conlon, “IUS dual use: Monitoring whales and earthquakes using SOSUS,” *Mar. Technol. Soc. J.* **27**, 13–21 (1994).

⁷C. J. Bryan and C. E. Nishimura, “Monitoring oceanic earthquakes with SOSUS: An example from the Caribbean,” *Oceanography* **8**, 4–10 (1995).

⁸V. O. Knudsen, R. S. Alford, and J. W. Emling, “Underwater ambient noise,” *J. Marine Res.* **7**, 410–429 (1948).

- ⁹G. M. Wenz, "Acoustic ambient noise in the ocean: spectra and sources," *J. Acoust. Soc. Am.* **34**, 1936–1956 (1962).
- ¹⁰R. J. Urlick, "Ambient Noise in the Sea," Undersea Warfare Technology Office, Naval Sea Systems Command, Department of the Navy, Washington, D.C. 20362, 1984.
- ¹¹W. J. Richardson, C. R. Green, Jr., C. I. Malme, and D. H. Thomson, *Marine Mammals and Noise* (Academic, San Diego, CA, 1995).
- ¹²M. A. McDonald, J. A. Hildebrand, and S. C. Webb, "Blue and fin whales observed on a seafloor array in the Northeast Pacific," *J. Acoust. Soc. Am.* **98**, 712–721 (1995).
- ¹³P. O. Thompson and W. A. Friedl, "A long term study of low frequency sounds from several species of whales off Oahu, Hawaii," *Cetology* **19**, 1–19 (1982).
- ¹⁴A. R. Kolaini and L. A. Crum, "Observations of underwater sound from laboratory breaking waves and the implications concerning ambient noise in the ocean," *J. Acoust. Soc. Am.* **96**, 1755–1765 (1994).
- ¹⁵N. R. Chapman and J. W. Cornish, "Wind dependence of deep ocean ambient noise at low frequencies," *J. Acoust. Soc. Am.* **93**, 782–789 (1993).
- ¹⁶L. Ding and D. M. Farmer, "On the dipole acoustic source level of breaking waves," *J. Acoust. Soc. Am.* **96**, 3036–3044 (1994).
- ¹⁷F. C. Felizardo and W. K. Melville, "Correlations between ambient noise and the ocean surface wave field," *J. Phys. Oceanogr.* **25**, 513–532 (1995).
- ¹⁸C. S. McCreery, F. K. Duennebier, and G. H. Sutton, "Correlation of deep ocean noise (0.4–30 Hz) with wind, and the Holo Spectrum—A worldwide constant," *J. Acoust. Soc. Am.* **93**, 2639–2648 (1993).
- ¹⁹R. H. Bourke, T. H. Holt, and C. R. Dunlap, "Ambient noise levels in the Northeast Pacific Ocean as measured by aircraft-dropped sonobuoys," in *Underwater Ambient Noise, Proceedings of a Conference held at SAACLANTCEN on 11–14 May 1982* (1982).
- ²⁰G. M. Wenz, "Low-frequency deep-water ambient noise along the Pacific Coast of the United States," *U.S. Navy J. Underwater Acoust.* **19**, 423–444 (1969).
- ²¹J. F. Lynch, H. X. Wu, R. Pawlowicz, P. F. Worcester, R. E. Keenan, H. C. Graber, O. M. Johannessen, P. Wadhams, and R. A. Shuchman, "Ambient noise measurements in the 200–300-Hz band from the Greenland Sea tomography experiment," *J. Acoust. Soc. Am.* **94**, 1015–1033 (1993).
- ²²B. M. Howe, S. G. Anderson, A. Baggeroer, J. A. Colosi, K. R. Hardy, D. Horwitt, F. W. Karig, S. Leach, J. A. Mercer, K. Metzger, Jr., L. O. Olson, D. A. Peckham, D. A. Reddaway, R. R. Ryan, R. P. Stein, K. von der Heydt, J. D. Watson, S. L. Weslander, and P. F. Worcester, "Instrumentation for the Acoustic Thermometry of Ocean Climate (ATOC) prototype Pacific Ocean network," in *Oceans '95 Proceedings, San Diego, CA, 9–12 October 1995*, pp. 1483–1500, 1995.
- ²³Gerald D'Spain, personal communication (1998).
- ²⁴Louis Maisel, *Probability, Statistics, and Random Processes* (Simon & Schuster Trade, 1983).
- ²⁵D. Ross, "On ocean underwater ambient noise," *Acoust. Bull.* **18**, (1993).
- ²⁶K. Stafford, "Characterization of blue whale calls from the Northeast Pacific and development of a matched filter to locate blue whales on U.S. Navy SOSUS (SOund SURveillance System) arrays," Oregon State University, Corvallis, OR, 1995.
- ²⁷Christopher Clark, personal communication (1997).
- ²⁸R. C. Spindel, "An underwater acoustic pulse compression system," *IEEE Trans. Acoust., Speech, Signal Process.* **27**, 723–728 (1979).
- ²⁹P. O. Thompson, L. T. Findley, and O. Vidal, "20-Hz pulses and other vocalizations of fin whales, *balaenoptera physalus*, in the Gulf of California, Mexico," *J. Acoust. Soc. Am.* **92**, 3051–3057 (1992).
- ³⁰P. O. Thompson and L. T. Findley, "Underwater sounds of blue whales, *balaenoptera musculus*, in the Gulf of California, Mexico," *Mar. Mammal Sci.* **12**, 288–293 (1996).
- ³¹J. A. Rivers, "Blue whale, *balaenoptera musculus*, vocalizations from the waters off Central California," *Mar. Mammal Sci.* **13**, (1997).
- ³²C. W. Clark and K. M. Fristrup, "Whales '95: A combined visual and acoustic survey of blue and fin whales off Southern California," *Rep. Int. Whal. Commun.* **47**, 583–600 (1997).
- ³³Christopher Clark, personal communication (1998).
- ³⁴K. M. Stafford, C. G. Fox, and D. S. Clark, "Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean," *J. Acoust. Soc. Am.* **104**, 3616–3625 (1998).
- ³⁵A. Zell, G. Mamier, R. Hubner, N. Mache, M. Schmalzl, T. Sommer, and M. Vogt, "SNNS: an efficient simulator for neural nets," in *MASCOTS '93 International workshop on modeling, analysis and simulation of computer and telecommunications systems, San Diego, CA*, edited by H. Schwetman, J. Walrand, K. Bagchi, and D. DeGroot, pp. 343–346, 1993.
- ³⁶J. A. Hertz, A. S. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison Wesley Longman, 1991).
- ³⁷W. H. Thorp, "Deep ocean sound attenuation in the sub- and low-kilocycle per second region," *J. Acoust. Soc. Am.* **38**, 648 (1965).
- ³⁸F. J. Wentz, "A well-calibrated ocean algorithm for special sensor microwave/image," *J. Geophys. Res.* **102**, 8703 (1997).
- ³⁹D. M. Farmer and S. Vagle, "On the determination of breaking surface wave distribution," *J. Geophys. Res.* **93**, 3591–3600 (1988).

Ambient noise imaging in warm shallow seas; second-order moment and model-based imaging algorithms

John R. Potter^{a)} and Mandar Chitre

Acoustic Research Laboratory, Electrical Engineering Department, National University of Singapore,
10 Kent Ridge Crescent, Singapore 119260

(Received 27 August 1997; revised 15 June 1998; accepted 1 July 1999)

Ambient noise can be used to produce images of submerged objects using the mean intensity of the backscattered energy, a technique coined “acoustic daylight” because of its direct analogy to vision. It is suggested that there may be substantial additional information in higher moments of the data. At high frequencies (>10 kHz), absorption suppresses long-range propagation so that a received signal is largely dependent on the local geometry, source characteristics, and the scattering properties of interceding objects. It is shown that for snapping shrimp (*Cragnon*, *Alpheus*, and *Synalpheus*) illumination (the primary sources in warm shallow water above a few kHz), significant information is embodied in the second temporal moments of intensity. There is no visual analog to this concept, which suggests a broader imaging approach which may be termed ambient noise imaging (ANI). Another ANI technique explored is the use of spatial cross correlation, which works well and also has no visual analogy. A model-based processor (Kalman filter) is also applied to track targets subject to highly variable illumination such as provided by snapping shrimp. Examples are presented using data provided by Scripps Institution of Oceanography from the initial deployment of the Acoustic Daylight Ocean Noise Imaging System (ADONIS) in San Diego. © 1999 Acoustical Society of America. [S0001-4966(99)01711-7]

PACS numbers: 43.30.Pc, 43.20.Fn, 43.30.Wi, 43.30.Gv [DLB]

INTRODUCTION

The possibility of using ambient noise in the ocean as the sole source of illumination to form images of submerged objects emerged about a decade ago, beginning with precursory ideas about objects causing ambient noise shadows.¹ Professor Buckingham and colleagues developed this early idea into the more complete concept of “acoustic daylight” (AD),² which included the possibility of increased illumination from “front-lighting” in addition to silhouetting by “backlighting.” These ideas are based on an analogy with optical vision, where the visual clues are all spatial, the human eye response being too slow to detect temporal variability in optical illumination under normal circumstances. This analogy is overly restrictive for the ocean acoustic case, where significant variations in the acoustic illumination occur on a time scale long compared to the sensors’ response time and sampling interval. Tracking these variations, including a statistical description, offers additional possibilities for imaging. The optical analogy also breaks down when considering the angular resolution that can be achieved, the wavelengths of sound in water being much larger with respect to practical acoustic apertures. The human eye has an aperture about 10^4 times the wavelength of visible light, whereas acoustic apertures are typically 10^2 wavelengths or smaller. Similarly, diffraction is expected to play a greater role as scattering objects are smaller in terms of wavelengths than for visible light.

An “Acoustic Daylight Ocean Noise Imaging System” (ADONIS) was constructed at Scripps’ Institution of Ocean-

ography (SIO). ADONIS used a 3-m-diam spheroidal reflecting lens to beamform incoming energy over the aperture onto 126 receivers in the focal plane, each of which thus provided an output associated with a unique “look-direction.” A largely analog electronics package was developed to estimate the average acoustic amplitude received in 16 frequency bands from 8 to 80 kHz. This estimation was performed for each channel by a switched-capacitor filter whose center frequency was shifted to each of the 16 frequencies in turn, followed by a “fallow” settling period before beginning again at the lowest frequency. The pressure signal in each frequency band was rectified and integrated over a 2-ms time window, approximately 30 times a second. Of these 30 “frames” each second, an average of 24 were received and recorded by the surface data acquisition and display system as validated corruption-free data frames. The amplitude at each frequency was therefore observed and recorded for a total of some 48 ms each second, representing approximately 5% of the time. Since no phase information was available, and the ambient noise sources are many and presumed incoherent with one another, the recorded acoustic amplitudes have in all cases been squared to give power estimates before further processing. It is these acoustic power values that we shall treat as the raw input data. No raw time-series data are available from ADONIS.

The theoretical possibility of imaging using the AD technique became demonstrated fact with the first deployment of ADONIS in 1994 which yielded images of arrangements of 1×1 m² target panels placed at 38-m range.³ The AD images from that deployment generally display a signal-to-noise ratio of 2–4 dB. A high contrast is unlikely to be observed in AD images due to the reduced angular resolution

^{a)}Electronic mail: johnp@arl.nus.edu.sg

and diffraction effects which “fill-in” geometric shadows. In addition, optical images are often formed from viewing scattered illumination from a few very bright sources, whereas underwater ambient noise is anticipated to be more isotropic, generating milder contrasts. Finally, ambient noise is likely to be statistically complicated both in time and space, and arise from several possible source mechanisms. This poses a problem, since a variation in noise intensity could arise either from scattering from an object moving through the field of view, or by natural variations in the noise generating mechanisms. In the ADONIS deployment which provided the data for this work, the sources were snapping shrimp of unknown species (possibly *Cragnon*, *Alpheus*, or *Synalpheus*), a conclusion we will discuss later. These sources are known to inhabit areas offering some shelter on the bottom or on structures and to “snap” episodically,^{4,5} thus providing substantial source structure in both time and space.

Rather than consider temporal variations as “noise,” to be suppressed by averaging, in late 1995 the Acoustic Research Laboratory (ARL) of the National University of Singapore (NUS) began to explore the possibility that the variability itself could contain information. We show that image information resides in the variance of the acoustic intensity in each pixel, or look direction. Image information can also be extracted from spatial cross correlations of the acoustic intensity data. These are both second-order moments of the acoustic intensity. Furthermore, we demonstrate that a model-based processor is able to estimate the presence of targets more reliably in sporadic and unreliable illumination conditions than a linear estimator based on raw data.

I. AMBIENT NOISE SOURCES RELEVANT TO THE EXPERIMENT

Immediately on entering the water in preparation to deploy ADONIS, the experiment divers heard the characteristic “crackling” of snapping shrimp. This sound could also be heard during quiet periods from within the hull of the Offshore Research Buoy (ORB), the floating platform we used to support our surface equipment. The sound of echolocating dolphins from the nearby NOSC pens were also heard intermittently by our divers. ADONIS often received a marked increase in acoustic power at high frequencies if oriented to the right of the target range used in the experiment, towards the pens. The imaging experiments were all conducted along a line free of this interference, and in any event the dolphins became disinterested as time wore on and echolocated on our site less and less.

During the experiment the sea state was mostly 0, and at most 1, with little or no wind. No precipitation fell during data collection. An observer on ORB’s deck communicated by VHF to the ADONIS operators to cease data acquisition if any motorized vessels approached. Many of the datasets were taken late at night when boat traffic was at a minimum, offering more opportunity to collect boat-free data and without the possible interference of noise from crews working on vessels tied up or moored nearby.

Although no raw acoustic time series was recorded during the experiment, such data were taken at a subsequent ADONIS deployment at the same site and have generously

been made available to us by SIO. A representative excerpt of these data, spanning a period of 4 s, is shown in the upper panel of Fig. 1. The characteristic broadband impulsive “spikes” of the shrimp snaps are clearly seen, with the maximum pressure over the 1-s period exceeding the average level by over two orders of magnitude. A spectrogram of this data, shown in the lower panel of Fig. 1, confirms that the sharp impulsive transients dominate the spectral energy distribution over the entire bandwidth of ADONIS, 8–80 kHz. The data in Fig. 1 are similar to those taken from Scripps’ Pier, 14 km to the north, shown in Fig. 1 of Buckingham and Potter’s “Sea Surface Sound 94” paper.⁶ The high-frequency ambient noise at Scripps’ Pier is also dominated by snapping shrimp, and the individual snaps have been triangulated by short arrays of hydrophones (Jeff Nystuen, personal communication). Indeed, one expects any shallow (less than 60 m) warm (>11°C) coastal waters offering some sheltering materials to be populated by snapping shrimp.⁷

One suspects from the appearance of Fig. 1 that snapping shrimp noise, if considered as a random process, is non-Gaussian. Indeed, it fails the chi-squared test for normal distributions. It has been noted elsewhere⁸ that snapping shrimp noise from a variety of locations (including the experimental site) exhibit the robust property of a log-normal distribution of energy with time. As Fig. 1 illustrates, such noise is characterized by relatively quiet periods of many milliseconds, punctuated by energetic narrow pulses of shrimp clicking. A measure of the extreme transient pulse nature of snapping shrimp noise is provided by the scintillation index SI (normalized variance of intensity) given by

$$SI = \frac{\langle I^2 \rangle - \langle I \rangle^2}{\langle I \rangle^2}, \quad (1)$$

where I is the acoustic intensity (proportional to pressure squared) and $\langle \rangle$ indicates the ensemble average.⁹ The value of the SI increases as extreme values of acoustic pressure become more common. For a zero-mean normally distributed random pressure variable, the SI=2 irrespective of the mean power level. For the data from the ORB site in Fig. 1, the value of the SI exceeds 500.

Examining ADONIS data from pixels subsequently identified as being oriented towards “target” and “nontarget” directions reveals that both types display approximately log-normal statistics. The ORB data usually display “frontlit” target conditions, we suspect as a result of the geometry of the Marine Facilities Pier to which ORB was tied and which, therefore, was behind ADONIS. The Pier very likely harbored considerable colonies of snapping shrimp. The prevailing “nontarget” noise is also shrimp, likely scattered about other structures and on the sea bed, but of a generally lesser intensity.

II. PRINCIPLE OF HIGHER-ORDER TEMPORAL MOMENT IMAGING

We proceed to the reasoning behind the ideas of temporal higher-moment imaging, initially proposed in 1996.¹⁰ Consider the acoustic energy received in a narrow beam corresponding to one of the image pixel “look” directions of an

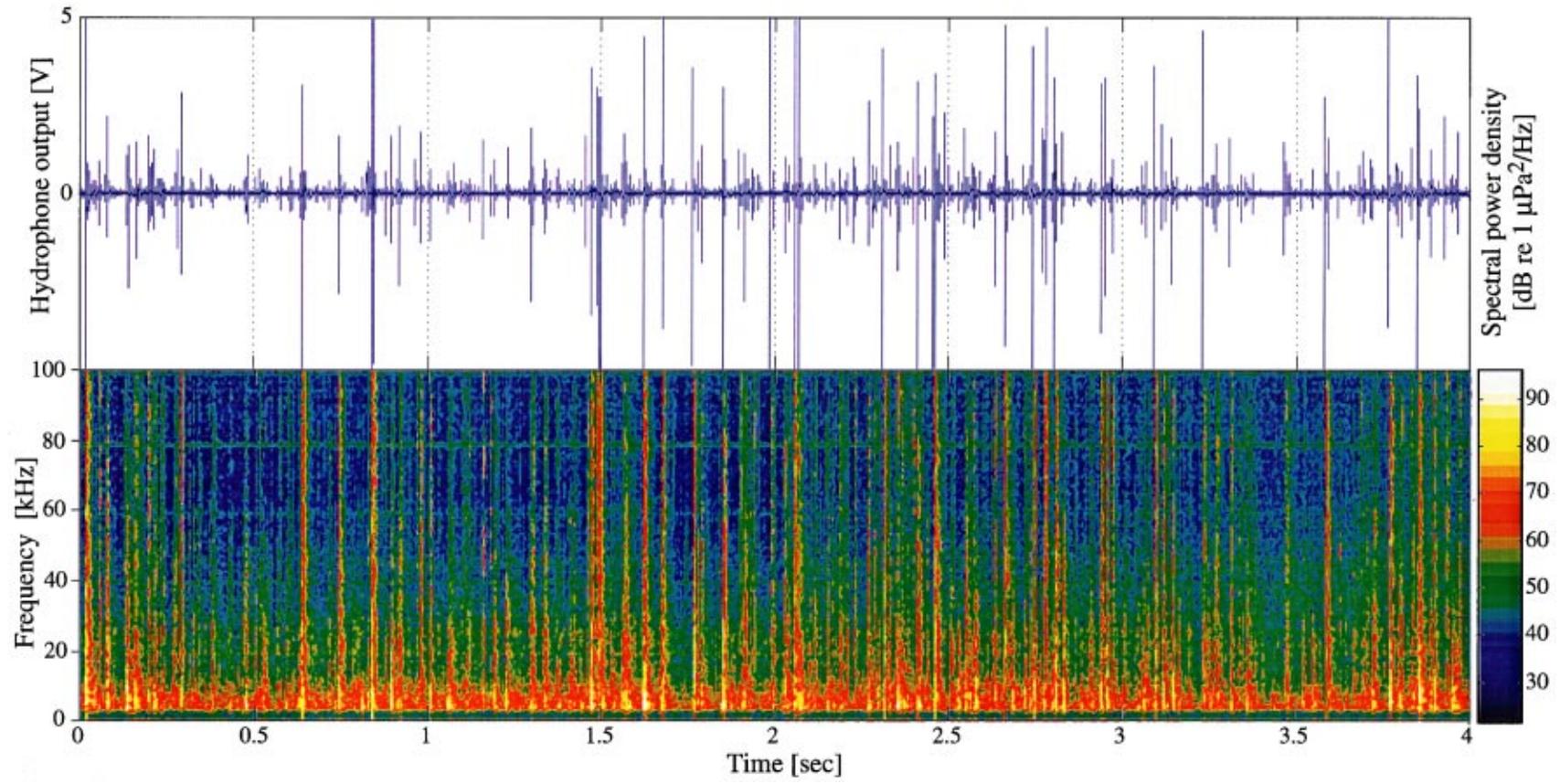


FIG. 1. Hydrophone voltage (proportional to acoustic pressure) versus time for 4 s of raw acoustic data from the experiment site (upper panel) and corresponding calibrated spectrogram evaluated by DWFFT (lower panel).

acoustic imaging system. This energy will be considered the basic data unit, although it is itself a second-order moment of acoustic pressure, taken at zero time and zero spatial lags. We consider a warm, shallow sea, where we anticipate snapping shrimp dominate the noise field at the high frequencies useful for acoustic imaging and an appropriate model for the ORB experiment. For systems that acquire data over time windows of several tens of milliseconds, as does ADONIS, we can therefore expect considerable variations in the illumination intensity from frame to frame. This variability, in particular the method of exploiting it to form images, has no visual analog.

In a time window Δt , the observed time-integrated energy received by a directional acoustic receiver will be one realization from a distribution whose characteristics are determined by the density of snapping shrimp in the beam and their distribution with range. If a target (even a perfectly reflecting one) is present in a receiving beam, this may change the received signal in a number of ways:

- (i) A target will scatter energy into the beam from other directions, in addition to partially obscuring energy that would otherwise have been received. If the source field is spatially anisotropic, this may result in a difference in the mean illumination over time, the AD imaging principle. The population density of snapping shrimp is inhomogeneous (they are known to prefer certain bottom types and to colonize surfaces which offer shelter⁵). Different directions will therefore be characterized by different shrimp populations.
- (ii) The solid angle of a beam's sensitivity will be different if the target scattering surface is curved. Concave targets will reduce the solid angle over which a receiver is sensitive, and convex targets (most commonly found due to hydrodynamic considerations) will increase the solid angle of a receiving beam. Contrary to one's first impression, this effect does not affect the mean received intensity. The change in solid angle is exactly balanced by the divergence of the acoustic pressure wave after scattering from the curved surface, so that the mean amplitude remains constant (conserving energy).¹¹ Nevertheless, the second-order temporal statistics are changed. The wider solid angle resulting from scattering from a convex surface reduces the variance, as the receiver is sensitive to a larger population of the (presumed random) sources. This provides one of the methods we refer to as ambient noise imaging (ANI).

There are thus at least two independent mechanisms by which a target is anticipated to change the received acoustic intensity statistics. The first is simply by reflecting energy from different directions. The second is subtler, leaving the mean acoustic amplitude unchanged, but affecting the variance. There is thus good reason to believe that imaging information should be contained in the higher temporal moments of the received acoustic time series, rather than just in the mean. In the (generally unfavorable) condition of isotropic noise, if the target has a curved scattering surface, then second-order moments should still yield images, even when

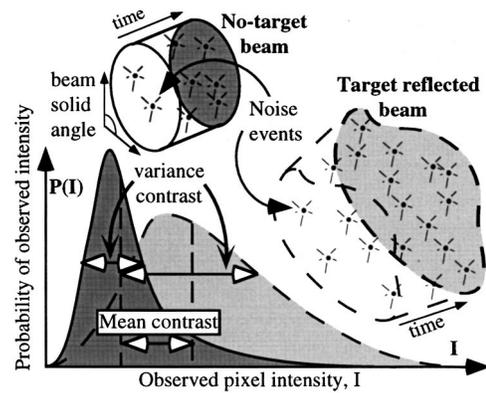


FIG. 2. Geometric basis for expecting differences in the target and nontarget intensity distributions. Distributions may differ not only in mean but also in higher moments, such as variance.

first-order estimates fail to do so. We have not been able to test this hypothesis, since the available data is all of flat targets and anisotropic noise.

Even in the case of flat targets, such as we have for the ORB experiment, the snapping shrimp distribution requires two parameters to describe it, and the mean provides only one. The other is obtained from the second-order moment of intensity. We thus anticipate that second-order measures will be generally useful in extracting imaging information, and moreso for data involving curved targets.

A schematic illustration of this principle is given in Fig. 2. The different contributing regions and modifications in the solid angle of an imaging beam on reflection from a target give rise to probability distributions for the received energy which are different in both the mean and variance. The cylinder delineated by solid lines depicts regular beam sensitivity, unscattered by targets. The space delineated by a dashed line depicts a beam scattered by a curved target which distorts and expands the solid angle within which source may contribute, in addition to redirecting it. Even if the mean source density is similar, many more sources will be observed in the scattered beam, at correspondingly lower individual intensities. The distributions of beam intensities estimated over many such frames may then appear as at the bottom of Fig. 2, where the solid line shows a distribution curve for unscattered beams, and the dashed line indicates target-scattered beams. An image may be formed by either considering the contrast in mean intensities of the two distributions, or a second-order temporal moment, such as the difference in variance of the two distributions.

III. ORB IMAGES USING FIRST- AND SECOND-ORDER MOMENTS

A disadvantage of the higher statistical moment method is that one generally requires increasingly longer time series to obtain reliable estimates for statistical moments of increasing order (see, for example, Ref. 12). We therefore consider only the first and second temporal moments in this paper, the mean and standard deviation (s.d.) of the received intensity in each beam. This permits us to calculate robust statistics over relatively few (250) frames. The ORB deployment of ADONIS concentrated on two types of target, the

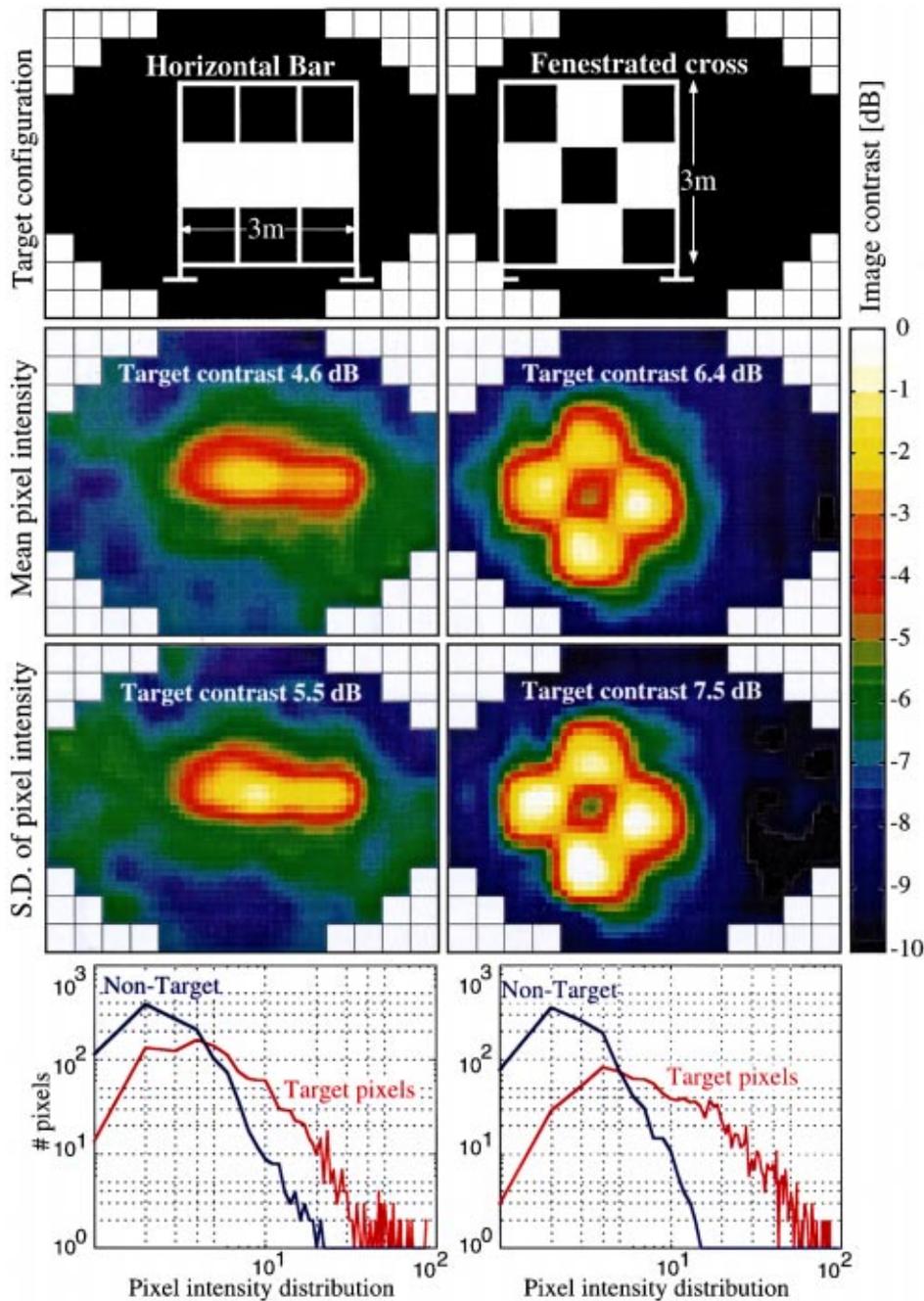


FIG. 3. “Horizontal bar” (left) and “fenestrated cross” (right) targets, shown schematically in the upper panels, with ambient noise images formed from first (mean) and second-order (s.d.) moments shown below, calculated from 250 frames. Lower panels show the observed probability distributions of target (red) and nontarget (blue) pixel energies for the two target configurations.

“horizontal bar” and the “fenestrated cross” at the maximum available range of 38 m.¹³ These targets were composed of flat $1 \times 1 \text{ m}^2$ reflective panels mounted on a “tic-tac-toe” frame. For the horizontal bar, three squares were placed in a horizontal row in the middle of the frame. For the “fenestrated cross,” additional panels were placed above and below the middle of the bar, and the center panel removed to form a “hole.” The $1 \times 1 \text{ m}^2$ “hole” subtends an angle of 1.5 degrees at the receiver, and the Rayleigh resolution of a 3-m aperture is expected to correspond to this value at approximately 20 kHz. Therefore, the lower part of the ADONIS bandwidth was expected to provide insufficient resolution to reveal the “hole,” and to seriously blur the other parts of the chosen targets. Diffraction of energy impinging on the targets, resulting in less backscattered energy

to the receiver, is also expected to reduce imaging contrasts. For this reason only frequency estimates above 26 kHz have been used in this analysis to form images, the average frequency across the bandwidth we have used being 48.3 kHz.

We proceed to show results from the mean intensity (AD) processor in comparison to second-order methods in both the time and spatial domains.

A. Second-order temporal statistical imaging

Some first- and second-order moment images are shown in Fig. 3, where the original receiving pixels have been spatially interpolated as a final processing step using bi-cubic splines, a process which considerably improves the eye’s ability to delineate important features. The top panel in

the left-hand column of Fig. 3 shows the configuration for the “horizontal bar” target. Below this, a mean intensity (AD) image has been displayed. The next panel shows an ANI image formed from the same 250 data frames using the standard deviation of the intensity. In the bottom panel, we show the probability distribution curves for target (red) and nontarget (blue) pixels, where six of each have been selected from the images above to form the statistics. It is clear that a substantial amount of image information is extracted from the standard deviation of intensity; not only sufficient to produce a very acceptable image, but one with a superior contrast to that obtained from the mean intensity. This is confirmed by the probability distribution curves, which show substantial differences in both the mean and variance of the two distributions.

The right-hand column shows the same panels as for the left, but using 250 frames of “fenestrated cross” target data. Again, the second statistical moment proves marginally superior in contrast, though both methods produce excellent images. The differences in the means and characteristic widths of the target and nontarget pixel intensity distributions are very clear, consistent with the conceptual scheme presented in Fig. 2. Despite presenting the more difficult task with regard to resolution, the “fenestrated cross” target is for some reason slightly better contrasted with its background than the “horizontal bar.” The distributions show that this occurs both because the nontarget pixel distribution is slightly narrower, and the target distribution slightly wider, than their counterparts for the “horizontal bar” target case. The reasons for this are unknown; it may simply arise from a small random change in the illumination and be of no particular significance.

Since the estimated mean and variance are independent parameters of the pixel intensity distributions (at least for normal or log-normal distributions), information extracted from these two parameters is independent, or orthogonal. Estimates from orthogonal processors can be combined to form images of greater contrast and statistical stability than using either processor alone.

B. Second-order spatial imaging

The temporal variance of a pixel’s energy is equivalent (within an additive constant) to its autocorrelation at zero time lag. A spatial extension of this idea is to consider the spatial cross correlations of pixel intensity at zero time lags, the off-diagonal elements of the correlation matrix formed by cross correlating all 126 receiver channels.

The underlying idea is that not only the mean intensity level and variance (which are statistical estimates), but also the actual observed time history of intensity variations will be correlated for imaging beams which receive energy from neighboring regions. Since reflecting targets are expected to cause very different source regions to illuminate the receiving beams than where there are no targets, it is anticipated that the cross correlations of the 126 channels will indicate how to divide the channels into two sets. One set would be associated with target-reflected look directions, the other with nontarget.

We proceed as follows: First, the (square) normalized correlation matrix of the 126 channels is calculated. The diagonal elements are unity and the (symmetric) nondiagonal elements are the 0-lag cross correlations between channels. Second, the matrix element with the smallest cross correlation is found. The two channels associated with this value are used as the seeds for the target and nontarget sets. The extent to which the other channels correlate to these two seed channels can be used to form the image. We do this by calculating the normalized distance of each of the remaining channels from the two seed channels in the correlation space and assigning a value between 0 and 1 which represents the relative distance to each of the two seeds.

Figure 4 shows images analogous to those in Fig. 3, except taken over the entire data file length for each of the two target types (2323 frames for the bar, and 934 frames for the fenestrated cross), so that image contrast is generally better than for the 250 frames of Fig. 3. Once again, bi-cubic spatial interpolation has been applied as a final step to improve the appearance of the images. The bottom panel of Fig. 4 shows images from the spatial cross-correlation method. The familiar forms are clear, confirming that the spatial cross correlation at zero time lag also contains substantial useful information. Indeed, the image contrast is slightly better both for the spatial and temporal second-moment methods than it is for the traditional mean intensity image. Note that the spatial cross correlation requires no information about the temporal statistics, and would form an image even if there were no information in the mean or second-order temporal statistics of intensity. The spatial cross-correlation method is therefore independent of the first two methods.

IV. MODEL-BASED PROCESSING

While the images of Figs. 3 and 4 are very gratifying, it should be remembered that these have been obtained from the best sequence of 250 frames in the case of Fig. 3 (equivalent to some 10 s in real time) or from long data sets (40–98 s) which suppress noise at the expense of responsiveness. For much of the data, no recognizable image is produced at all. Clearly, this has much to do with the variation in illumination and propagation conditions. There may have been episodic events of particularly serendipitous illumination, of just the right direction and strengths to best reveal the target. In any event, we cannot expect to produce good images on demand and at all times, subject as we are to the vagaries of ambient noise illumination. Nevertheless, targets can be expected to move in smooth, continuous ways through the field of view. Just because it is no longer lit appropriately does not mean that it no longer exists. When the human brain is confronted by sporadic bursts of useful information, it attempts to track the state of the external reality, ready to update its model of the outside world when new “useful” information becomes available. We should perhaps consider doing the same with our ANI systems. The usual term for this is “model-based” processing, where as much of the physical constraints which define the possible behavior of the system are encoded into a Kalman Filter, which then estimates the

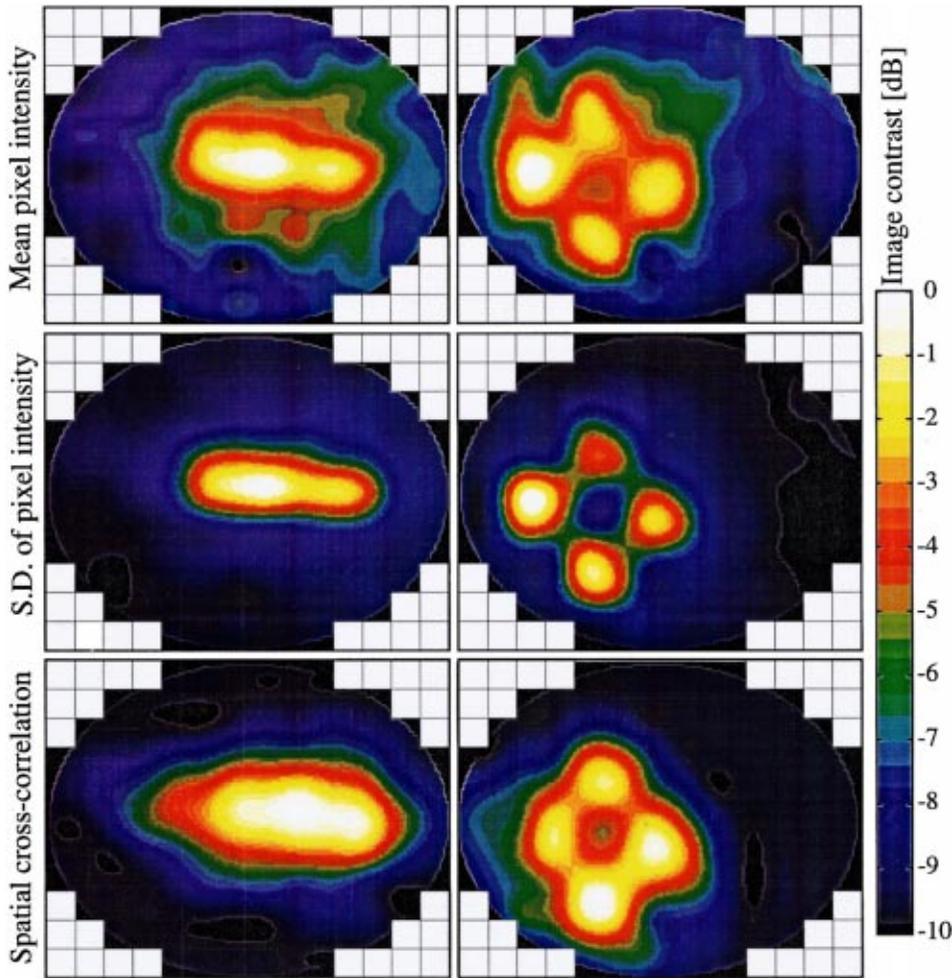


FIG. 4. Images of the horizontal bar and fenestrated cross targets as in Fig. 2, except formed over the entire length of the available datasets (2323 and 934 frames, respectively), and including spatial correlation images in the lowest panel.

current state of the system much more accurately than an uninformed one, even in the presence of substantial noise or total absence of recent useful information.

A. A modified Kalman filter

A Kalman filter is an optimal estimator (with respect to any quadratic function of estimation error) for linear systems perturbed by Gaussian white noise. That our noise is likely to be non-Gaussian should not necessarily deter us from trying this filter, unless an obviously superior processor presents itself. The Kalman filter permits one to encode the physical *a priori* information available about a system. These features make it an obvious candidate for estimating the beam intensity in ADONIS data as a preprocessor to (possibly nonlinear) imaging algorithms.

Our objective is to develop the simplest possible Kalman filter that demonstrates its usefulness by tracking the beam intensity better than simple averaging. To begin, we develop a Kalman filter to be individually applied to each pixel in the image. The state-space description for each pixel is written as

$$r(k) = H(k)x(k) + v(k), \quad (2)$$

where $r(k)$ is the observed energy at discrete time k in each beam at a given frequency. Here $r(k)$ is obtained from ADONIS data (though obviously calibrated and with all pos-

sible deterministic distortions removed) and $x(k)$ is the parameter to be estimated. It is the energy that propagates towards the sensing beam from the position of the target frame, i.e., before distortions due to propagation effect and pollution from nontarget-scattered energy entering the beam. The transform $H(k)$ embodies the propagation loss from target to receiver, which can reasonably be assumed constant. Since we are interested in contrasts between target and nontarget pixels rather than absolute values, we set $H(k)$ to unity. Here $v(k)$ is a superimposed zero-mean random component.

The evolution of the state-space over time can be modeled by a one-step auto-regressive process (Markov process):

$$x(k) = M(k-1)x(k-1) + w(k), \quad (3)$$

where M represents the evolution of the object space and $w(k)$ is a zero-mean random component. If we assume that the targets do not move, M can be set to unity. This reduces the state-space model to

$$r(k) = x(k) + v(k), \quad x(k) = x(k-1) + w(k). \quad (4)$$

We now require a physical model for how the illumination is generated and propagated into the beams. Consider the scenario in Fig. 5, which is the least complex model which still embodies some of the essential physical processes. The angular limits of a single pixel's beamwidth combined with the high degree of absorption at the frequencies employed (typically 48 kHz) limit the area in which the sources can be

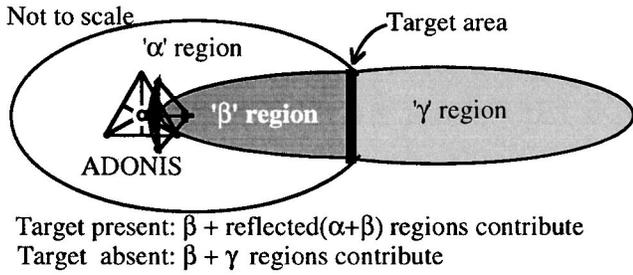


FIG. 5. Schematic of the simple geometric physical model of illumination distribution regions and propagation which contributes to received energy in a beam.

received in the absence of targets to a region which we have indicated consists of the union of areas β and γ . If a target lies in the beam, then energy from the γ region is shadowed, and that arising from some other region, termed α , may be reflected into the receiving beam. Allowing for the β region to contribute both by direct and target-reflected paths, we obtain expressions for the received energy (assumed summed incoherently),

$$r(k) = \Pi[A(k) + \alpha(k)] + (1 + \Pi/3)[B(k) + \beta(k)] \quad (5)$$

if there is a reflecting target in the beam, and

$$r(k) = [\Gamma(k) + \gamma(k)] + [B(k) + \beta(k)] \quad (6)$$

if there is none, where the source contributions from the regions α , β , γ are modeled as comprising of mean expected values A , B , Γ with superimposed zero-mean Gaussian perturbations α , β , γ . The factor Π is the reflection coefficient of the target, including geometrical divergence due to curvature. The dividing factor of 3 in the reflected B-region term in Eq. (5) arises from the increased path length associated with the scattered energy from the target compared to directly received energy.

For simplicity, we assume that the target is perfectly reflecting and flat, oriented towards ADONIS, appropriate for the ORB deployment. The absolute amplitudes of the noise terms will depend on the length of the time for which the energy is integrated, the beamwidth of the receiver, the spatial geometry and extent of each region, and the attenuation at the chosen frequency. Since we seek only to estimate the noise terms in a relative sense, these effects can be incorporated into the noise parameters.

Equations (5) and (6) show that any substantial difference in the average amplitude from the reflected α and directly propagating γ regions will reveal the target over a sufficiently long averaging period. This is indeed how the first images were produced. Additionally, even if Π , Γ , A , and B conspire to make the expressions in Eqs. (5) and (6) to be nearly equal, the target might still be revealed by the differences in the variance of the random contributions α , β , and γ .

From Eqs. (5) and (6) and our assumptions,

$$r(k) = x(k) + [B(k) + \beta(k)], \quad (7)$$

where

$$x(k) = x(k-1) + [[\alpha(k) - \alpha(k-1)] + 4/3[\beta(k) - \beta(k-1)]] \quad (8)$$

if there is a target in the beam, and

$$x(k) = x(k-1) + [[\gamma(k) - \gamma(k-1)] + [\beta(k) - \beta(k-1)]] \quad (9)$$

if there is not. Note that the noise term in Eq. (7) is not zero-mean as required unless we estimate B a priori and subtract it from the data before Kalman filtering. Since the β region is not expected to contribute as strongly as the other regions due to its small physical size, we anticipate $B \ll A, \Gamma$ and hence $r(k)$. Numerical experimentation gave results insensitive to the choice of B , so it was set to zero.

Even though two separate models are used for deriving the Kalman filter state-space equations, the state space Eq. (7) is the same for both. The second state space equation for the target and the nontarget scenario [Eqs. (8) and (9)] differ only in their zero-mean random noise term. Thus, the Kalman filter algorithm used for both target and nontarget cases is the same. The Kalman filter to estimate intensities $x(k)$ are then derived in the normal way.¹⁴

To compare the performance of the Kalman filter with simple averaging, Fig. 6 shows images obtained by the two methods side by side for an example period of some 18 s real time while viewing the fenestrated cross. Both the simple averaging and Kalman filter were initialized with frame 1 of the file. The simple averaging results are shown in the left panels, and the Kalman filter results on the right. The displayed frames proceed from no. 261 (shown at the top), where the image is clear to both algorithms, through no. 344, where it begins to deteriorate and on to frame no. 450, when only one of the four target panels is effectively illuminated. At this point, the simple averaging shows only that panel. The Kalman filter, however, recognizes that the other panels are still present, and that the lack of illumination does not imply lack of target. It interprets the data in the light of the physical propagation model. By frame no. 706, the incoming data is severely polluted by a strong source in the lower right of the field of view, so that the fenestrated cross image has been completely obliterated in the simple averaging image. The Kalman filter retains the target image and is only slightly affected by this noise pollution, which does not fit the state-space model and is therefore largely rejected.

The Kalman filter does not have a fixed averaging time. The effective averaging time is adaptive, becoming shorter when the data matches the model and becoming longer when the data is largely seen as noise. It is therefore not possible to estimate the performance improvement due to a Kalman filter by comparing the averaging time of the Kalman filter with that of simple averaging. However, the performance improvement can be estimated by initializing both the Kalman filter and the simple-averaging algorithm with noise. When actual data is presented to both the filters, they both discard the noisy initial estimate and pick out the target from the data at approximately the same time. This suggests that the performance advantage of a Kalman filter demonstrated

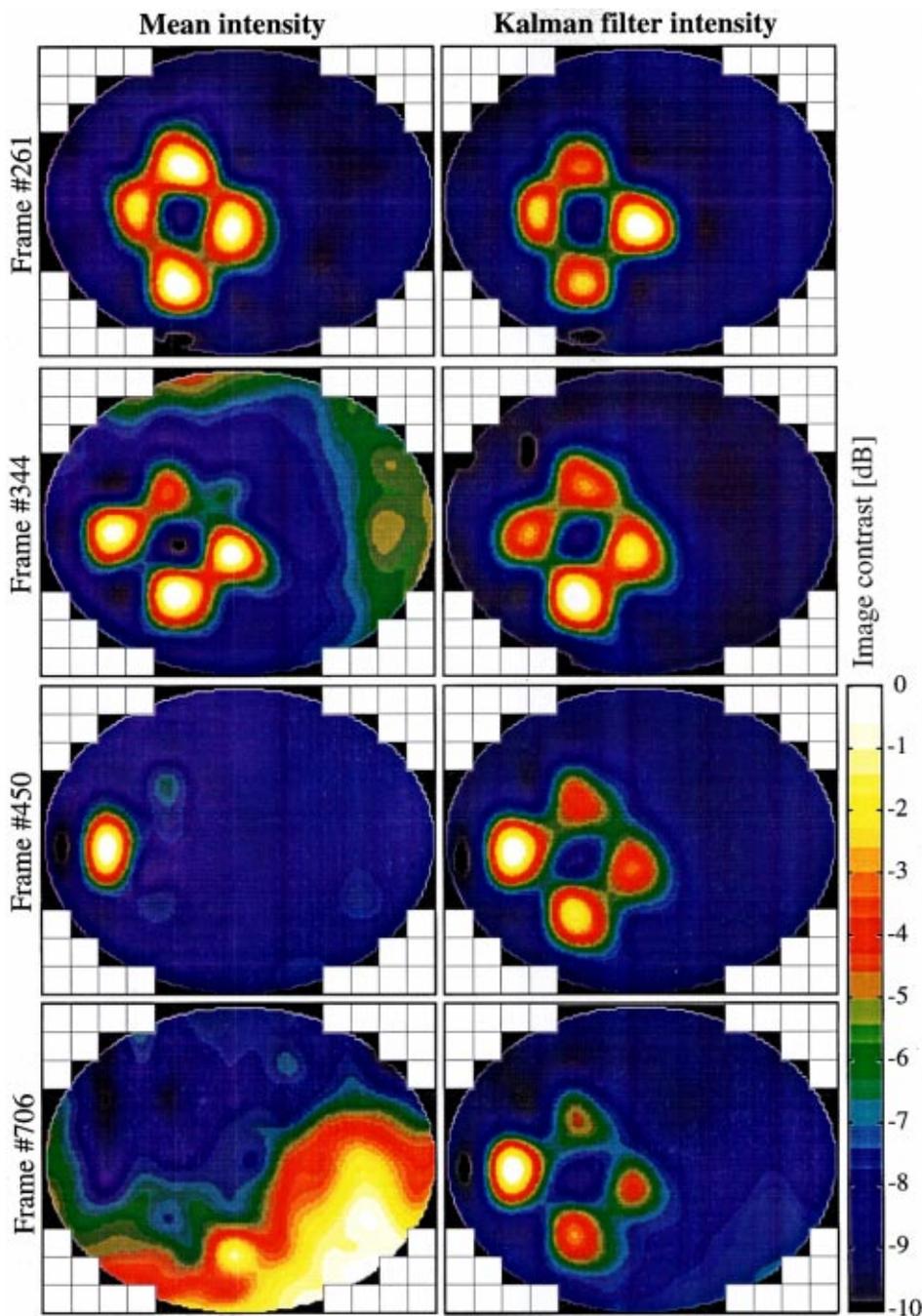


FIG. 6. A sequence of four selected frames over a period of 18 s showing the mean intensity (left) and Kalman filter estimate (right) at each frame.

in Fig. 6 is not due to a longer effective averaging time, but rather because of the adaptive nature of the Kalman filter.

The Kalman filter model developed here is a static model. It can easily be extended to provide the ability to image moving objects by a simple adaptive enhancement. When the innovation sequence of all Kalman filter channels shows that the model matches the data poorly, the estimated variance of each channel can be increased, thus allowing the filter to accept changes in the data due to movement of the target. Such a moving target scenario was tested successfully with actual data from the ORB deployment and seen to be nearly as stable as the traditional Kalman filter.

While the Kalman filter model presented here is a very basic one, it nevertheless demonstrates that the method has

great potential power to improve the tracking of targets illuminated by ambient noise.

V. DISCUSSION

Although far from all the data frames are satisfactory, acceptable ambient noise images can certainly often be formed by simple averaging of received intensities, which corresponds most closely to the visual analogy of Acoustic Daylight. Additionally, we have found that we can exploit at least two of the higher statistical moments, processes that have no visual analog. The statistical complexity of ocean ambient noise invites an improved analysis of the source statistics and underlines the potential gains of using a spectral estimation technique which uses all the incoming data,

such as a discrete windowed fast Fourier transform (DW-FFT), rather than the sequential stepped analog filter employed in ADONIS, which effectively discards some 94% of the data. It is not yet clear how much of the interframe variability is due to source variation, and how much is due to weak multiple scattering by inhomogeneities in the intervening medium, though if dominated by propagation effects we would not expect such clear images from the spatial cross correlation as shown in the lower panels of Fig. 4. Now that the basic ANI concept has been proven, these and related issues need to be investigated in order to design improved imaging hardware and signal processing approaches. The ANI principle is sufficiently novel (primarily due to its spatially diffuse, random, incoherent source field) that tools developed for optical and radar imaging are as likely to be useful as established marine acoustic propagation techniques.

Because high frequencies are attenuated more rapidly with range, received high-frequency sources are likely to be nearer to the target and receiver than received energy at lower frequencies. This improves the likelihood of observing a strong anisotropy in the ambient noise directionality at high frequencies, compared to noise generated by a spatially similar distribution of lower-frequency sources over a larger volume. As has been demonstrated by simulation, ANI imaging potential is strongly linked to the degree, if not so dependent on actual orientation, of the noise anisotropy.¹⁰ There is thus good reason to believe that in warm shallow waters, which constitute some 70% of the Earth's coastal regions, snapping shrimp will provide excellent ambient noise illumination for imaging.

VI. CONCLUSIONS

We have derived two independent processors (temporal second-moment and spatial second-moment) for ambient noise imaging (ANI) which produce as good or better contrast images compared to the mean intensity processing of the original acoustic daylight (AD) principle. The target information contained in the mean intensity is controlled by the spatial inhomogeneity of the sources. The target information in the temporal second moment of intensity is controlled both by the source distribution and by the target shape. The temporal mean and variance of the sources themselves are independent parameters of the source distribution, and hence orthogonal. The two methods can therefore be used in concert to improve dynamic contrast and robustness of the image. The spatial cross-correlation method requires no knowledge of the temporal statistics, and is therefore orthogonal to both methods. Indeed, all three processors (and perhaps others, too) would likely be run in parallel in an ideal ANI system.

We have also developed a simple physical model of the illumination and propagation environment which forms the foundation of a Kalman filter to estimate reflected intensities in each beam. Despite the crude assumptions in this physical model, the Kalman outperforms simple averaging, significantly reducing the problems associated with episodic illumination of the targets.

ACKNOWLEDGMENTS

This work was supported by DRD, Singapore and the National University of Singapore. ADONIS data and a raw time series of ambient noise taken from the experiment site were graciously provided by Scripps Institution of Oceanography, UCSD. Special thanks to Michael Buckingham, who first set us on the path, Grant Deane and Chad Epifanio, who contributed freely and substantially to the development of ADONIS and provided the raw time series, and to the many other players who strove to build ADONIS.

- ¹S. Flatté and W. Munk, "Submarine detection: Acoustic contrast versus Acoustic glow," JASON Report JSR-85-108 (MITRE, McLean, VA, 1985).
- ²M. J. Buckingham, B. V. Berkhout, and S. A. L. Glegg, "Imaging the ocean with ambient noise," *Nature (London)* **356**, 327–329 (1992).
- ³M. J. Buckingham, J. R. Potter, and C. L. Epifanio, "Seeing in the ocean with background noise," *Sci. Am.* **274**(2), 86–90 (1996).
- ⁴D. H. Cato and M. J. Bell, "Ultrasonic ambient noise in Australian shallow waters at frequencies up to 200 kHz," Materials Research Laboratory, DSTO, Australia Report MRL-TR-91-23 (1992).
- ⁵M. L. Readhead, "Snapping shrimp noise near Gladstone, Queensland," *J. Acoust. Soc. Am.* **101**, 1718–1722 (1997).
- ⁶M. J. Buckingham and J. R. Potter, "Observations, theory and simulation of anisotropy in oceanic ambient noise fields and its relevance to Acoustic Daylight imaging," in *Sea Surface Sound '94*, Lake Arrowhead, CA (World Scientific, Singapore, 1994), pp. 65–73.
- ⁷F. A. Everest, R. W. Young, and M. W. Johnson, "Acoustical characteristics of noise produced by snapping shrimp," *J. Acoust. Soc. Am.* **20**, (1948).
- ⁸J. R. Potter, T. W. Lim, and M. Chitre, "Ambient noise in shallow warm waters and implications for imaging," in *Sea Surface Sound '97*, Chilworth Manor, University of Southampton, UK (1997), pp. 45–54.
- ⁹B. J. Uscinski, *The Elements of Wave Propagation in Random Media* (McGraw-Hill, London, 1977), p. 62.
- ¹⁰J. R. Potter and M. Chitre, "ADONIS imaging with a Kalman filter and higher-order statistics," in Proceedings of the Third European Conference on Underwater Acoustics (Heraklion, Greece, 1996), pp. 349–354.
- ¹¹J. R. Potter, "Acoustic imaging using ambient noise: Some theory and simulation results," *J. Acoust. Soc. Am.* **95**, 21–33 (1994).
- ¹²J. S. Bendat and A. G. Piersol, *Random Data*, 2nd ed. (Wiley, New York, 1986), Chap. 8.
- ¹³C. L. Epifanio, J. R. Potter, G. B. Deane, M. Readhead, and M. J. Buckingham, "Imaging in the ocean with ambient noise: the ORB experiments," *J. Acoust. Soc. Am.* **106**, 3211–3225 (1999).
- ¹⁴M. S. Grewal and A. P. Andrews, in *Kalman Filtering: Theory and Practice* (Prentice-Hall, Englewood Cliffs, NJ, 1993), Chap. 4, Sec. 4.2.

Imaging in the ocean with ambient noise: the ORB experiments

Chad L. Epifanio,^{a)} John R. Potter,^{b)} Grant B. Deane, Mark L. Readhead,^{c)}
and Michael J. Buckingham^{d)}

*Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego,
9500 Gilman Drive, La Jolla, California 92093-0213*

(Received 16 April 1997; accepted for publication 8 July 1999)

Acoustic daylight imaging is a new technique that has been proposed for creating pictorial images of objects in the ocean from the ensonification provided by the incident ambient noise field. To investigate the feasibility of the technique, a series of experiments was performed from the research platform ORB, moored in San Diego Bay, Southern California. Central to these experiments was an acoustic receiver known as ADONIS (acoustic daylight ocean noise imaging system), which consists of a spherical reflector, 3 m in diameter, with an elliptical array of 130 hydrophones at the focal surface. This system, which is broadband, operating between 8 and 80 kHz, forms a total of 126 receive-only beams spanning the vertical and horizontal. The ambient noise power in each beam is mapped into a pixel on a VDU. Various types of targets were used in the experiments, including planar panels and cylindrical, polyethylene drums containing wet sand, seawater or syntactic foam (essentially air), and most of the experiments were conducted with the targets at ranges between 20 and 40 m. At the time of the experiments the noise field in the area was created primarily by snapping shrimp. Moving, color images of the object space were successfully created with ADONIS. Some representative static images from the moving sequences are presented and discussed in the paper. © 1999 Acoustical Society of America. [S0001-4966(99)04611-1]

PACS numbers: 43.30.Nb, 43.35.Sx, 43.60.Gk [DLB]

INTRODUCTION

Over recent years, anecdotal evidence has indicated that a submarine can be detected by the acoustic shadow it casts in the ambient noise field. In the mid-eighties, at the suggestion of Allen Ellinthorpe, Flatté and Munk¹ investigated the shadowing phenomenon theoretically and, at around the same time, Buckingham² independently introduced the idea that, as the ambient noise field is in many respects analogous to daylight in the atmosphere, it should be possible to create recognizable, pictorial images of objects in the ocean solely through the acoustic illumination provided by the ambient noise. Such “Acoustic Daylight” imaging, as Buckingham² designated the technique, would be similar to conventional photography, whereby an image of an object is created even though the object itself may not be luminous and no dedicated light source (e.g., a flashgun) is used for illumination. Ambient light (daylight) scattered from the object gives rise to the photographic image, whereas ambient sound would fulfill the same role in the ocean.

To test this idea, Buckingham *et al.*³ conducted a pilot experiment off Scripps Pier in 1991, the results of which indicated that, at ranges up to 12 m, a rectangular target can be detected solely from the modifications its presence intro-

duces into the ambient noise field. In effect, the targets created an acoustic contrast, that is, a difference in the noise level between the situations where the targets were present and absent. The acoustic detector used for focusing in this experiment was a parabolic dish of diameter 1.2 m, with a piezoelectric hydrophone located at the focal point. The surface of the dish was faced with closed-cell neoprene foam, which is an efficient acoustic reflector (essentially pressure-release) over the operating frequency band from 5 to 50 kHz. When the output of this receive-only, single-beam system is displayed as an intensity map, it forms what is in effect a single-pixel image of the object space. Since they are linked by a one-to-one correspondence, more pixels would require more beams.

The targets in the pilot experiment fully occupied the beam throughout most of the operating frequency range. Much of the ensonification was generated by snapping shrimp located on the pier pilings behind the parabolic dish, a situation analogous to front lighting produced when the sun is behind a (photographic) camera. The acoustic contrast observed in the experiment was around 3 dB, which is comparable with theoretical estimates by Buckingham⁴ of the contrast created by a spherical target embedded in plane-wave noise fields showing various degrees of anisotropy. Similar levels of contrast were also obtained by Potter⁵ from a numerical simulation of acoustic daylight imaging of volumetric targets in a shallow water channel.

In his wave-theoretic analysis of ambient noise imaging, Buckingham⁴ considered a linear array of hydrophones steered to endfire as the acoustic receiver. He found that the maximum contrast occurs when the angle subtended by the spherical target at the receiver matches the angular beam-

^{a)}Current address: Nautronix, 15150 Avenue of Science, San Diego, CA 92128-3416.

^{b)}Current address: Electrical Engineering Department, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260.

^{c)}Current address: Maritime Operations Division, Defence Science and Technology Organisation, P.O. Box 44, Pyrmont, New South Wales 2009, Australia.

^{d)}Also at: Institute of Sound and Vibration Research, The University, Southampton SO17 1BJ, England.

width (as measured between the -3 dB points) of the detection system. In this case, if the beamwidth is greater than the angle subtended by the target, background noise enters the beam, which reduces the contrast, while a beamwidth that is smaller than the subtended angle leads to rejection of scattered energy, again reducing the contrast. This observation provides a guide to imaging system requirements but should not be interpreted as a universally valid criterion of image quality: no such metric is known that is independent of the source distribution.⁶

Makris *et al.*⁷ have developed a theoretical model of ambient noise imaging in a shallow water waveguide. They considered a spherical target at mid-depth in the water column, with the focusing performed by a planar, billboard array. When the beamwidth (approximately) matched the angular width of the target, they found that the contrast in the resultant image was 3.6 dB (bottom left panel of their Fig. 12), which is comparable with the results of Buckingham⁴ and Potter.⁵ With a significantly broader beam, however, the contrast fell to 0.35 dB (bottom right panel of their Fig. 12), which is also consistent with the previous theoretical investigations.^{4,5}

The spatial scale that is resolved in an acoustic daylight image will, of course, be determined by the beamwidth of the acoustic detection system, in accord with the Rayleigh resolution criterion. To create a multiple-pixel image, some form of multi-beam or beam-scanning receiver system is required: the scattered acoustic energy from the object space must be sampled over a range of arrival angles. Clearly, the single-beam system used in the pilot experiment at Scripps Pier is not satisfactory for such an application.

To fulfill the imaging requirement, a performance specification for a prototype multi-beam receiver was drawn up. It was decided that the system should have 100 or more beams, a decade of bandwidth, a beamwidth at the highest operating frequency of less than 1 degree, corresponding to a spatial resolution of better than 1.75 m at a range of 100 m, and rear baffling to prevent unwanted noise from behind the receiver corrupting the response. All these requirements were incorporated into the acoustic daylight ocean noise imaging system (ADONIS), which was designed and built at Scripps over a period of about 2 years between 1992 and 1994.^{8,9} In this article, ADONIS is described briefly and a selection of ambient-noise images of planar and volumetric targets, obtained during two deployments in San Diego Bay, southern California, in August 1994 and October 1995, are discussed. These deployments have come to be known as the ORB experiments.¹⁰

I. ADONIS

Figure 1 shows a schematic of the ADONIS multi-beam receiver. Mechanically, the system consists of a spherical acoustic reflector with both a radius of curvature and a diameter of 3 m. The dish itself is a spherical fibreglass shell, the concave side of which is faced with closed-cell neoprene foam, selected because it is an almost perfect (pressure-release) acoustically reflecting material. A steel framework provides the dish with structural rigidity. The dish assembly is mounted on a vertical mast, which stands on a horizontal,

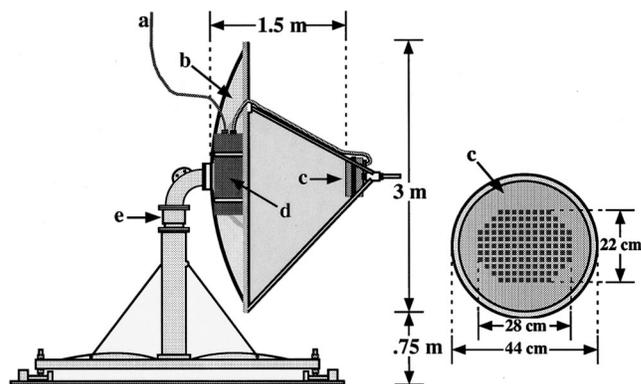


FIG. 1. Schematic of ADONIS. (a) Umbilical cable which provides power to the system and transmits data to a desk-top computer at the surface. (b) Spherical reflector faced with low density, closed cell neoprene foam. (c) 130-element hydrophone array, showing the elliptical configuration of the sensors, which have a center to center spacing of 2 cm. (d) Electronics canister where filtering is performed and frames are constructed. (e) Supporting mast containing a coaxial hydraulic motor for rotating the dish in azimuth.

triangular base frame. At the corners of the base frame are height-adjustable legs, which divers use for leveling the system after it has been deployed on the seabed. Inside the mast is a remotely controlled, coaxial hydraulic motor, which is capable of rotating the dish around a full 360 degrees in the horizontal. This facility is useful for panning the dish across the object space, and for monitoring the horizontal directionality of the ambient noise field.

The acoustic sensing is performed by an array of 130 piezoelectric hydrophones arranged in an approximately elliptical configuration, with the major (horizontal) and minor (vertical) axes containing 14 and 11 sensors, respectively. Each of the sensing elements has a square cross section with center-to-center spacing of 2 cm. The face of the array of sensors is slightly convex to match the curvature of the focal region of the spherical reflector. This and many other design details of the dish and the sensor array were established with the aid of a suite of simulation software packages that were developed at the outset of the project.

A multi-element reflector is an unusual design for an acoustic receiver. From the point of view of ambient noise imaging, it has the advantage that the beam forming is performed as a natural consequence of the geometry of the system: no phase or time delays are necessary. Sound incident on the dish from a given direction is focused onto a particular hydrophone, or, conversely, a given sensor has a unique "look" direction, which is governed by its position in the array head. Thus, the 130 hydrophones in the array provide a total of 130 receive-only beams, which are distributed in the vertical and horizontal (although only 126 beams are used for forming images). From the symmetry of the system, the beam widths in the vertical and horizontal are essentially the same. Incidentally, all the hydrophones in the array are off-axis, and the system suffers some degree of aberration arising from scattering by the rim of the dish. A spherical reflector was selected because the aberrations associated with the off-axis sensors are rather less pronounced than those encountered with a parabolic dish. (Of course, with an on-axis

sensor a parabolic dish provides ideal focusing and is thus the preferred choice. Such a system was used in the pilot experiment off Scripps Pier.³⁾

ADONIS operates over one decade of bandwidth (cf. one octave for the human eye), extending from 8 to 80 kHz. At the highest frequency the beamwidth between the -3 dB points is 0.75 degrees, which increases to 3.5 degrees at the lowest frequency. The beam centers are independent of frequency, providing an overall angle of view of approximately 10 degrees in the horizontal and 8 degrees in the vertical. In the lower frequency ranges the beams overlap, as a result of which the information they contain is no longer independent. The top frequency of 80 kHz was selected on the grounds that, beyond 80 kHz, the dominant noise component in the ocean was expected to be thermal noise,¹¹ which contains no imaging information. It is interesting to note, however, that the theoretical limit of around 80 kHz for the onset of thermal noise applies only on the assumption of a true point receiver. In reality, a receiver of finite size will have the effect of reducing the thermal noise fluctuations through surface averaging,¹² thereby increasing the frequency at which thermal noise becomes dominant. This offers the prospect of increasing the upper frequency limit to around 150 kHz in future ambient noise imaging systems.

The geometrical focusing performed by the dish provides a gain of approximately 22 dB at the highest frequency of 80 kHz, which is important in amplifying the very low-level signals scattered from the target. The gain of the dish is defined as the intensity of the acoustic field in the focal region divided by the intensity of the direct-path arrival at the same position. As with any lens system, the position of the focal point relative to the dish varies with the range to the target. Focusing is achieved by racking the hydrophone array in and out relative to the center of the dish in much the same way as the focusing on an optical camera is performed by moving the lens back and forth relative to the film plane.

The broad bandwidth of ADONIS offers the prospect of exploiting “acoustic color” for the classification of targets. For example, an object that scatters low frequencies better than high could appear as red in the final image, whereas a good high-frequency scatterer could appear as blue. Thus, it may be possible to distinguish, say, a hollow shell from a solid object of similar shape simply from the different colors of their images.

To show acoustic color in the images, a spectral analysis of the signal in each channel must be performed. Rather than applying a computationally demanding digital fast Fourier transform (FFT) to each channel, an analog technique is used in ADONIS in which a bandpass filter with a constant Q of 4 is swept over the full bandwidth 25 times per second (the frame rate). This task is performed by a bank of 130 switched-capacitor analog filters (one for each sensor channel), which sample the level of the signal at 16 frequency points, or bins, uniformly spaced logarithmically across the frequency band. Figure 2 shows the filter frequency responses; the center frequencies are listed in Table I.

A given frequency bin is sampled simultaneously in all channels. The sampling across the frequency bins is performed sequentially in time, and at each frequency the signal

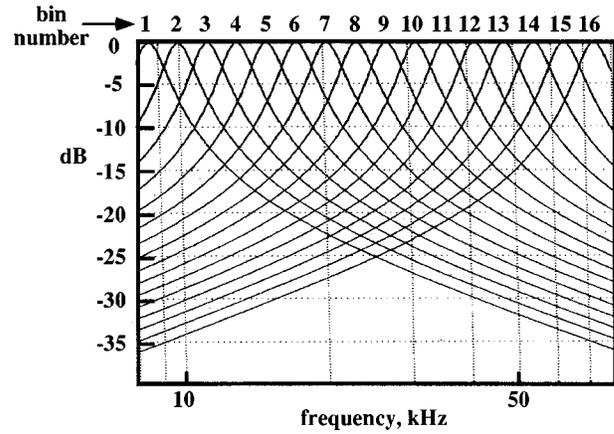


FIG. 2. Frequency responses of the sixteen switched capacitor filters, each normalized to the maximum response.

is allowed to settle before its average is taken over an interval of approximately 1 ms. The time for the complete 16-bin sweep is 40 ms, which sets the 25-Hz frame rate. Thus, each spectral estimate is obtained over an interval that is $\frac{1}{40}$ of the frame acquisition time. From a weighted sum of the resultant 16 intensities, a color is selected from a color palette and applied to the appropriate pixel on a monitor screen. At any instant, the 126 pixels on the screen constitute one image, or frame, in a sequence of continually refreshed static images. The frame rate of 25 Hz is sufficiently rapid to provide smoothly flowing movement in the final videolike output.

It became apparent during the initial ORB deployments that, because of an impedance mismatch, the switched capacitor filters had trouble driving the following stage when operating in the upper 60% of their dynamic range. As a result, the system shows a nonlinear response, which manifests itself as a reduction in contrast between pixels. However, only 25% of the data lie in the upper 60% of the dynamic range, where the effect is most severe. In most of the examples presented below, the nonlinear data were removed prior to creating the images.

The raw (uninterpolated) images are produced in real time on a desktop computer, which also handles data acqui-

TABLE I. Center frequencies at which the intensity of the noise is sampled in each channel, 25 times per second.

Bin	Frequency (kHz)
1	8.5
2	10.0
3	11.7
4	13.8
5	16.0
6	18.6
7	21.3
8	24.6
9	28.3
10	32.6
11	37.5
12	43.1
13	49.5
14	57
15	64.4
16	75.0

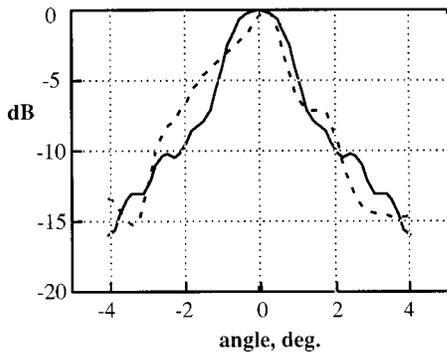


FIG. 3. Example of the measured (dashed) and theoretical (solid) in-air beam pattern taken at a frequency of 8.5 kHz (corresponding to 37.5 kHz in water).

sition and storage. As part of the image processing, a number of refinements are available, including various averaging and normalizing schemes for enhancing image stability. In post-processing, the raw data are interpolated by a factor of 5, that is, five additional points are included between data points, which leads to a significant improvement in the visual quality of the images.

II. BEAM PATTERNS OF ADONIS IN AIR

Before deploying ADONIS in the ocean, the beam patterns of the dish were measured at several frequencies in air on Scripps Pier. The frequency was scaled down appropriately to compensate for the difference between the speed of sound in water (≈ 1500 m/s) and in air (≈ 340 m/s). A microphone was mounted in the focal region on the axis of the dish, and the system was panned past a harmonic acoustic source at a range of approximately 30 m. The closed-cell neoprene foam facing the dish was absent during these in-air tests, thus exposing the underlying spherical fiberglass shell, which is essentially rigid.

Figure 3 shows an example of the measured in-air beam pattern at a frequency of 8.5 kHz, corresponding to a wavelength of 4 cm and an equivalent in-water frequency of 37 kHz. The slight asymmetry about the beam center in the measured curve is due largely to a breeze that was blowing across the axis of the dish. For comparison with the measurement, a theoretical beam pattern, derived from an analytical model developed by Deane¹³ specifically for this purpose, is also shown in Fig. 3. It can be seen that the theory and measurement show reasonably good agreement. Notice that at this frequency the beam width at the -3 dB points is only 2 degrees. A similar level of agreement was observed at the remaining frequencies for which measurements were performed.

III. BEAM PATTERNS OF ADONIS IN WATER

The beam patterns of ADONIS were measured in water by placing a spherical acoustic source at a range of 40 m and panning the dish across it in both directions, from left to right and right to left. For these tests the pressure-release, closed-cell, neoprene foam facing was present on the dish, as was the case for all the underwater imaging experiments. The

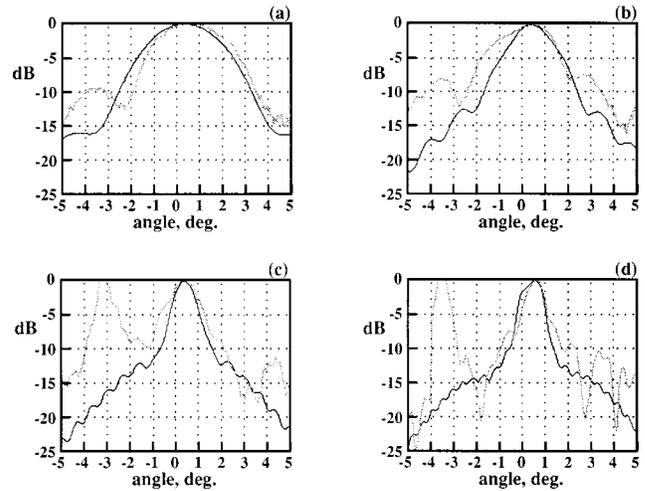


FIG. 4. Measured (gray) and theoretical (black) beam patterns of an array element close to the axis of the dish. (a) 10 kHz; (b) 25 kHz; (c) 50 kHz; and (d) 75 kHz. The array was focused at 25 m, as in the imaging experiments, and the source was at 40 m. The peak to the left of the main lobe in experimental traces (c) and (d) is due to boat noise.

dish was focused to 25 m, with a depth of field extending from 15 to 40 m, which was also the setup used for imaging with ADONIS.

Figure 4 shows examples of the measured beam patterns, at frequencies of 10, 25, 50, and 75 kHz, of a single element located near the geometrical center of the array. For comparison, the theoretically predicted beam patterns¹² are included in the figure. It is evident that the agreement between the theoretical and measured main lobes is reasonable at all four frequencies. Notice that the observed beamwidth at the -3 dB points is approximately 1 degree at 75 kHz, increasing to nearly 3.5 degrees at 10 kHz. In Fig. 4(c) and (d), the large peak to the left of the main lobe is the result of boats passing behind the testing area.

The operating feature of paramount importance in ADONIS is the shape of the beams. At the design stage, it was considered essential to predict the beam patterns accurately, in order to estimate imaging performance. The good agreement between theory and experiment in Fig. 4 provides confidence that the system does indeed perform as intended.

IV. TARGETS

Three different types of target were used in the imaging experiments: square, planar panels of area 1×1 m²; cylindrical, polyethylene drums with a capacity of 113 l (30 U.S. gallons), 0.76 m high, 0.5 m in diameter, with a wall thickness of 0.5 cm; and a hollow, titanium sphere with a diameter of 70 cm and wall thickness of 1.5 cm. A swimming diver with a closed breathing system was also used as a target and, as reported elsewhere,¹⁰ acoustic daylight images were obtained as he moved across the field of view.

The panels were mounted on 3×3 m² square frame resembling a tic-tac-toe board, which was placed vertically on the seabed (Fig. 5) in front of ADONIS at a range of 40 m. The face of the frame was approximately normal to the axis of the dish. Most of the panels were flat, and constructed of

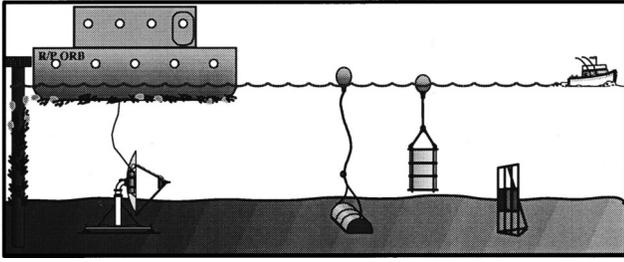


FIG. 5. Schematic of the ADONIS deployments below R. P. ORB. In practice, the various types of target illustrated in the figure were not deployed simultaneously. To the left of ORB is one of the pier pilings.

wood, fiberglass, aluminum, or aluminum faced with neoprene foam. One panel was of corrugated steel 3.2 mm thick.

The cylindrical drums were filled with wet sand (density $\approx 1900 \text{ kg/m}^3$), sea water (density $\approx 1000 \text{ kg/m}^3$), or syntactic foam (density $\approx 290 \text{ kg/m}^3$), referred to hereafter as the sand drum, the water drum, and the foam drum, respectively. Each drum was clamped with 6.4-mm-diam connecting rods between steel endplates, 6.4 mm thick, and weights were attached to the base of the foam drum to make it negatively buoyant. These drums were deployed in the water column, suspended from surface flotation units, and also partially submerged in the silty sediment forming the seabed (Fig. 5). In all cases the axis of each drum was perpendicular to the axis of the dish.

In a number of the water column deployments, several drums were suspended simultaneously from a 4-m-long floating wooden beam, an arrangement which allowed constant spacing between targets to be maintained. Both ends of the beam were anchored to fix orientation and distance from ADONIS. The depth of the targets was set so that at mid-tide the center of mass of each drum was 2 m above the seafloor, placing it in the middle of ADONIS' field of view. In the bottom deployments, divers arranged the drums at a range of approximately 15 m from ADONIS, which was tilted downwards by 10 degrees to keep the target in the field of view. From the point of view of imaging, these seabed deployments were particularly challenging, since the drums were partially (30%–50%) buried in the silty sediment.

The spherical target was supported in a metal cage, weighted to make it negatively buoyant, and suspended from surface floats (Fig. 5) at a depth of 2 m above the seafloor at mid-tide. As with the suspended drums, the sphere at mid-tide was then in the center of ADONIS' field of view.

V. DEPLOYMENT PROCEDURES

ADONIS was deployed through a moonpool in the rectangular, annular barge R. P. ORB, which was moored at the end of the Marine Facilities (MarFac) Pier, San Diego Bay. The water depth in this sheltered location is a nominal 7 m. Once it was in place on the seabed, the clearance between the top of the dish and the sea surface was about 3 m. Figure 5 shows the deployment configuration, with ADONIS beneath ORB looking towards the targets.

For many of the trials, a video camera was used to record conditions on the surface, particularly boat traffic and

swell. Beneath the surface, as a precautionary measure in the early experiments, a programmable, high-frequency acoustic source was mounted on the target frame. This source was activated intermittently (usually at the beginning and end of each data collection period) to confirm the alignment of the receiver with the targets. Thus, the source and hence the target positions within an image were localized to within one pixel. Incidentally, this same source was used for determining the in-water beam patterns of ADONIS.

As an additional check, the dish was periodically panned over the object space, causing the target to migrate back and forth across the image plane. This is an important test of the integrity of the imaging system, since it ensures that the target seen in an image is not an artifact associated with errors in calibration or equalization of the channels. Considerable attention was paid to balancing the 126 channels because the acoustic contrast in most acoustic daylight images is generally less than 4 dB, implying that even small errors in equalization could have resulted in false detection.

VI. AMBIENT NOISE SOURCES IN THE EXPERIMENTAL AREA

Within the calm, shallow waters of San Diego Bay, the three dominant sources of acoustic energy within the frequency range of ADONIS are snapping shrimp, industrial activity, and boat traffic. Marine mammals in pens several hundred meters distant from the experimental site also make a minor contribution to the background noise field, but this had little effect on the acoustic daylight experiments. Of the three main sources, the pulses from snapping shrimp feature most prominently in the time series of the noise. Several species of shrimp were collected near the experimental site, including *Synalpheus lockingtoni*, *Alpheus clamator*, the related genus *A. californiensis*, and *A. bellimanus*. Typically, these creatures are the size of a thumb nail, yet they are capable of producing extremely energetic, very brief pulses of sound with a broad bandwidth extending up to 200 kHz or beyond. Source levels may be as high as 190 dB *re* 1 μPa^2 at 1 m with a duration of the order of 5 μs or less.^{14–16}

In general, snapping shrimp do not swim well, and they tend to cluster in colonies around pier pilings, outcrops of rock, kelp holdfasts, and similar habitats which offer the animals shelter.¹⁷ No seasonal variations in shrimp activity have been reported, although a slight diurnal fluctuation in intensity is sometimes observed, with the shrimp noise marginally louder at dawn and dusk.¹⁸ On the basis of these observations, we expected the noise created by snapping shrimp during the ORB imaging experiments to originate from fixed directions and to be more or less continuous throughout the day and night. The evidence from our observations with ADONIS is consistent with this picture of the shrimp noise.

Industrial noise in the experimental area also exhibits a fixed spatial pattern, but shows strong diurnal fluctuations which correlate with human activities around the shoreline. The noise produced by dockside machinery and shore activity generally contains more lower frequency acoustic energy than the pulses from the snapping shrimp. This was exemplified in some of the early ORB experiments, when the targets were inadvertently aligned with a naval loading dock

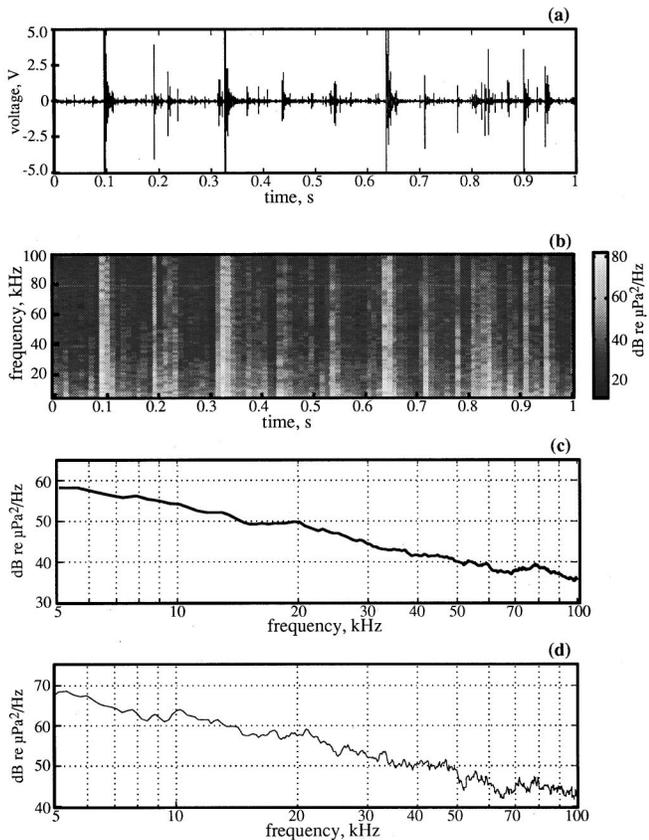


FIG. 6. (a) One second time series of (omnidirectional) ambient noise showing snapping shrimp pulses. (b) Spectrogram of the time series in (a) with the gray scale representing power spectral density. (c) Power spectral density of the noise, averaged over 30 s. (d) Power spectral density averaged over 15 shrimp pulses.

across the bay. Occasionally, the low-frequency noise generated around the dock area was more intense than the scattered noise from the targets, with the result that the targets appeared in silhouette at the lower frequencies in the acoustic daylight images.

Shipping and surface traffic produce a broad spectrum of sound, extending from below 100 Hz up to several kHz.¹⁹ At the lower frequencies, tonals appear associated with shaft and propeller blade rates, while cavitation produces a continuous spectrum extending to the higher frequencies. As a ship's speed increases, the radiated noise levels rise, implying that sound from vessels sufficiently close to the targets, such that the higher frequencies are not significantly attenuated, will contribute to the imaging process. Unlike shrimp and industrial noise, the component of the noise field generated by surface traffic shows a highly variable directionality. Accordingly, during daylight hours, when a high density of recreational and naval traffic is present, the ambient noise field in San Diego Bay is very variable, especially towards the lower end of the frequency band used for acoustic daylight imaging.

Examples of the omni-directional, broadband noise data collected with a low-self-noise hydrophone mounted on the rim of the ADONIS reflector (outside the focal region) are shown in Fig. 6. The data were recorded during late evening hours, when little boat or industrial noise was present. Figure

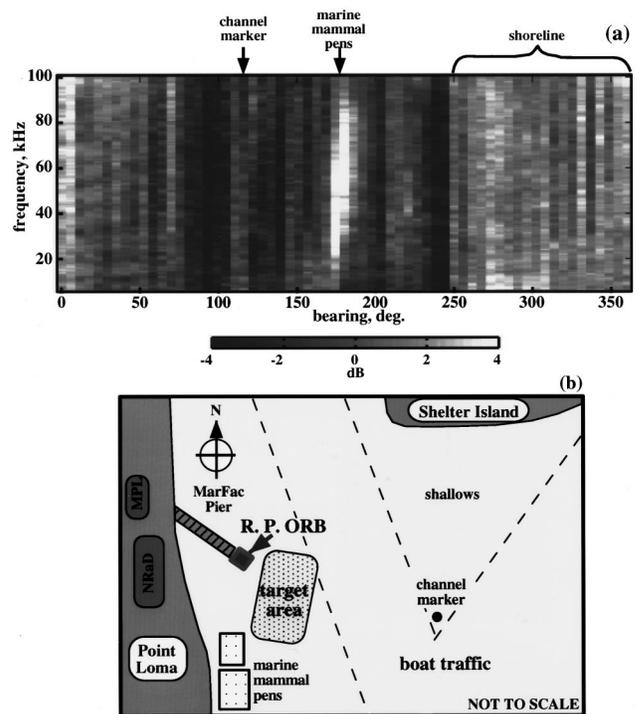


FIG. 7. (a) Frequency-azimuth plot of the noise, obtained by rotating ADONIS through 360 degrees in the horizontal. In a given frequency cell, the data are referenced to the noise level averaged over all bearing angles. (b) Plan view of the experimental site showing the test area and local features that affected the ambient noise environment.

6(a) shows a 1 s time series in which the noise is dominated by brief, energetic bursts of sound from local snapping shrimp. Counting only those pulses whose amplitude exceeds 10% of the highest peak level, the average period between snaps is 30 ms, which provides a useful measure of the averaging time required to obtain a temporally stable acoustic daylight image with this type of ensonifying field. Incidentally, the width of the pulses in Fig. 6(a) is about $8 \mu\text{s}$, which is actually the temporal resolution of the system, rather than the true width of a snapping shrimp pulse.

Figure 6(b) shows a spectrogram of the time series in Fig. 6(a), in which each vertical stripe represents an average power spectrum computed from 12 individual spectra obtained from 10 ms of data, with 70% overlap, using 512-point FFTs. It is clear from a casual comparison of the time series with the spectrogram that the pulses exhibit a broad spectrum extending up to at least 100 kHz, which is approximately the bandwidth of the measurement system. Note that the noise power varies by as much as 65 dB over a time scale as brief as 10 ms.

The spectrum of a longer time series is shown in Fig. 6(c). In this case, the data were split into 30 1-s segments, each of which was Fourier transformed in a standard way: Hanning windowing, 2048-point FFTs with 50% overlap, providing 220 spectra in the 1-s average. All 30 of the resultant spectra were then further averaged to obtain the spectrum in Fig. 6(c), which has a spectral resolution of approximately 100 Hz. The spectral gradient is -18.5 dB/decade , which is slightly steeper than the slope of wind-generated noise from breaking waves.²⁰

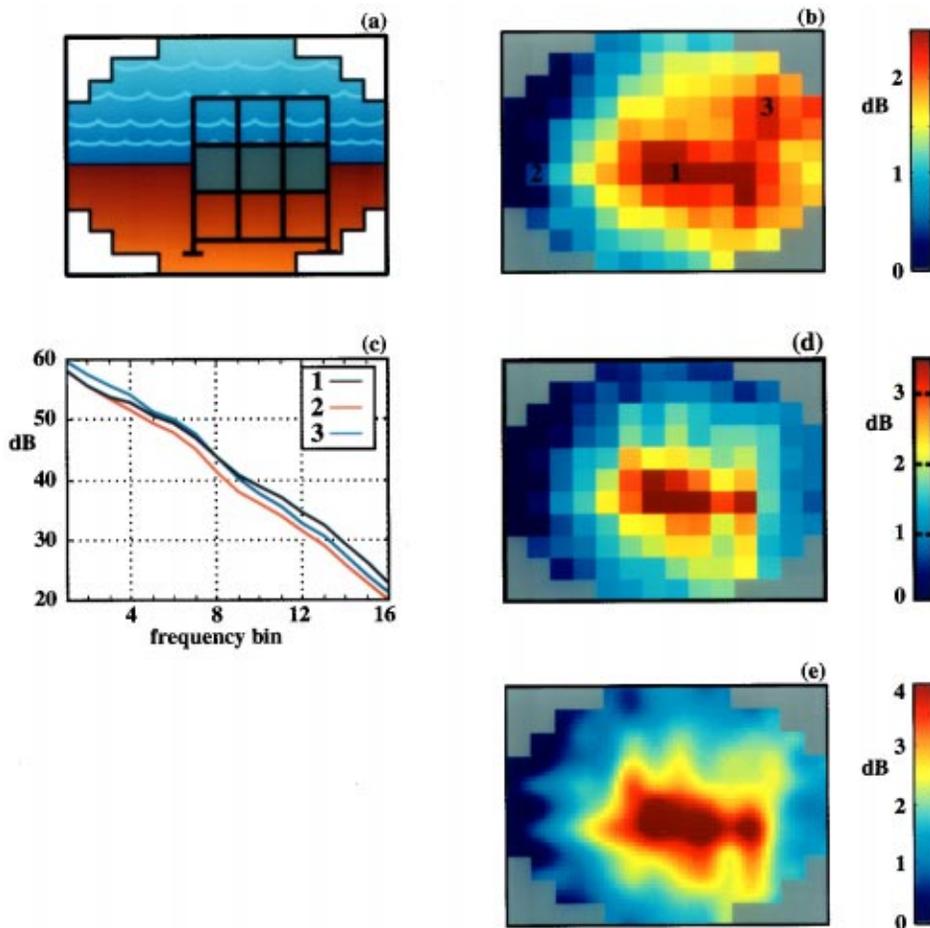


FIG. 8. Bar target at a range of 38 m. (a) Schematic of the scene falling within the field of view of ADONIS. (b) Uninterpolated, broadband, intensity mapped image formed by linearly averaging data from all sixteen frequency bins (8.5 to 75 kHz). (c) Spectrum of the noise in pixels on-target in dB *re* $1 \mu\text{Pa}^2/\text{Hz}$. (1), left of target (2), and above target (3). (d) Uninterpolated intensity mapped image formed by averaging over the top three frequency bins (57 to 75 kHz). (e) Interpolated version of the image in (d). The images in (b), (d), and (e) were formed from data taken at the same time.

To investigate the spectral content of the snapping shrimp pulses, 15 of the more energetic events were selected from the time series, Fourier analyzed, and averaged. The resultant spectrum is shown in Fig. 6(c). On comparison with the spectrum of the continuous time series in Fig. 6(c), it can be seen that, although the level of the shrimp spectrum is higher by about 8 dB, the spectral gradient is the same: -18.5 dB/decade. Although not conclusive, this observation suggests that the overwhelming contribution to the spectrum of the long time series in Fig. 6(c) is energy from the snapping shrimp pulses.

VII. HORIZONTAL ANISOTROPY OF THE NOISE

By rotating ADONIS in the horizontal using the hydraulic motor mounted within the mast, it was possible to measure the azimuthal dependence of the noise field through a full 360 degrees. A low-self-noise hydrophone mounted to the right of the imaging array, but within the focal region, was used to acquire the noise data. The beam pattern of this phone was horizontal, with essentially the same shape as that of the receivers in the main array head. Figure 7(a) shows the spectral and angular dependence of the noise, as observed in the early hours of the morning when no boats were present. The data set spans 10 min, with each vertical stripe computed from 10-s segments of the time series, corresponding to an angular scan of 5.5 degrees. Each segment was Fourier analyzed with 2048-point FFTs applied to time-windows

with 50% overlap, providing 2200 terms in the average and a spectral resolution of approximately 100 Hz. The spectra were then normalized by the average spectrum taken over the 360 degree scan.

The azimuthal variation of the noise spectral density in Fig. 7(a) amounts to about 8 dB. To interpret the origin of the anisotropy, the spectrogram may be compared with the plan view of the area around the experimental site in Fig. 7(b). A prominent feature in the spectrogram is the high energy burst around 180 degrees, which aligns with the marine mammal pens, and is attributed primarily to dolphin vocalizations. Between 250 and 360 degrees is a broad swath of noise, which arises mainly from snapping shrimp inhabiting the pier pilings and rocks along the shore. The noise signature at 110 degrees is also due to snapping shrimp, in this case a colony occupying the channel marker piling close to Shelter Island.

It is apparent from Figs. 6 and 7 that the ambient noise field around the MarFac Pier in San Diego Bay is highly variable, azimuthally nonuniform, and dominated by snapping shrimp noise. As far as acoustic daylight imaging is concerned, the shrimp noise provides excellent ensonification because of its broad bandwidth and high source level. Much of the shrimp noise originates on the pier behind ADONIS, giving rise to front ensonification of the targets. Any additional noise from passing vessels will also affect the imaging, in a way that depends on range and bearing. As the boats are mostly confined to channels in front of ADONIS,

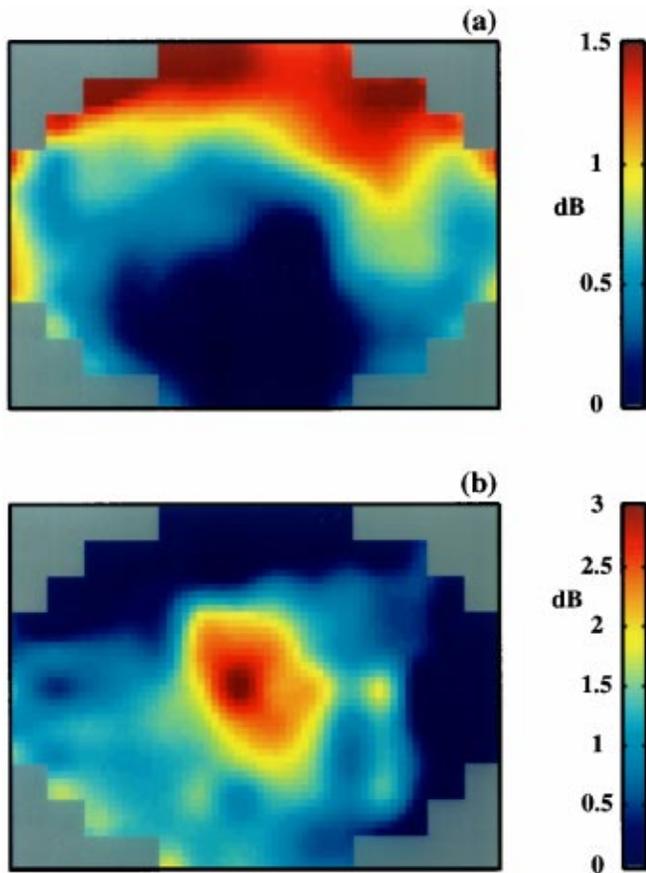


FIG. 9. Interpolated, intensity mapped images of the bar target at a range of 38 m. The two images were formed from data taken at the same time. (a) Low-frequency silhouette, formed by averaging over frequency bins 2, 3, and 4 (10 to 13.8 kHz). (b) High-frequency, front-ensonified image, formed by averaging over frequency bins 14, 15, and 16 (57 to 75 kHz).

their contribution to the noise field often increases the intensity of the noise from behind the target, tending to create a silhouette effect.

If snapping shrimp noise were unique to the San Diego area, then its utility in connection with acoustic daylight imaging would be limited, but it appears that this is not the case. The available evidence indicates that snapping shrimp are ubiquitous in temperate and tropical coastal waters, wherever appropriate habitats exist.^{15,21-24}

VIII. IMAGE FORMATION BY INTENSITY MAPPING

To create the acoustic daylight images, the intensity in each beam is mapped into a corresponding pixel on the screen of a computer monitor. Each image is rectangular, containing 14×11 pixels, with those pixels in the corners containing no information since there are no receivers at these locations. Thus, the useful image is approximately elliptical, having the same shape as the hydrophone array at the focal surface of the dish. The pixels show color, produced from the ‘jet’ colormap, with blue and red, respectively, representing low and high acoustic intensity over a band of frequencies. (In a few images, intensity mapping is not used but instead color represents the intensity at different frequencies, that is to say, ‘acoustic color.’) In all the in-

tensity mapped images presented below, the acoustic contrast between pixels is referenced, in dB, to the lowest intensity pixel.

The raw data from ADONIS yields images that are rather coarse grained, which often makes it difficult to recognize a target. A significant improvement in image quality is achieved by interpolating between data points. The interpolation algorithm used is a bi-cubic, by a factor of 5 in each direction. This inserts four additional points between each pair of original data points, resulting in a new 51×66 image matrix instead of the original 14×11 .

The intermittent character of the snapping shrimp noise [Fig. 6(a)] leads to a strong frame-to-frame variation in image intensity. The fluctuations are very evident in the sequences of moving images in which the screen is refreshed once every 40 ms. To alleviate the problem, time averaging is performed over a certain number of frames. In the moving sequences, either boxcar or exponential averaging is used, to provide a running average over a number of previous frames. Boxcar averaging is used to produce the still images. The time averaging, performed over 10 s (i.e., 250 frames) in all the images presented below (except in Fig. 15) stabilizes the images significantly.

Since the intensities at each frequency are estimated sequentially, the time window for each frequency bin within a frame is only 1.5 ms, with a period of 40 ms separating successive estimates in the same frequency bin. Thus, for 250 frames, the total window time for any particular bin is only 375 ms, or less than 4% of the total averaging time of 10 s. According to Fig. 6(a), there are on average 33 recognizable snapping shrimp pulses per second, with each pulse sufficiently brief to fall completely within the time window of a frequency bin. On average, only 13 of these pulses will occupy a given frequency bin in the 10-s interval over which the averaging was performed. One way of increasing the number of shrimp pulses contributing to an image is to average over several frequency bins, which is the approach adopted here. By averaging over three neighboring frequency bins, the effective windowing time is extended from 0.375 to 1.13 s, corresponding to 12% of the collected data.

In the moving sequences and most of the stills, the images are normalized by dividing the intensity in each pixel by the average intensity in all the pixels. This procedure leads to some improvement in frame-to-frame stability, but generally the effect is less pronounced than that due to time averaging.

IX. IMAGES OF THE PLANAR TARGETS

A. The horizontal bar

Figure 8 shows acoustic daylight images of a rectangular target formed from three square aluminum panels each 3.2 mm thick and faced with closed-cell neoprene foam 6.4 mm thick. The foam side of the panels faced towards ADONIS and the range to the target was 38 m. The scene within the field of view of ADONIS is illustrated in Fig. 8(a). In Fig. 8(b), a raw (uninterpolated) broadband image is shown, formed by averaging all sixteen frequency bins (8.5 to 75 kHz) over 250 frames, corresponding to 10 s of data. At the

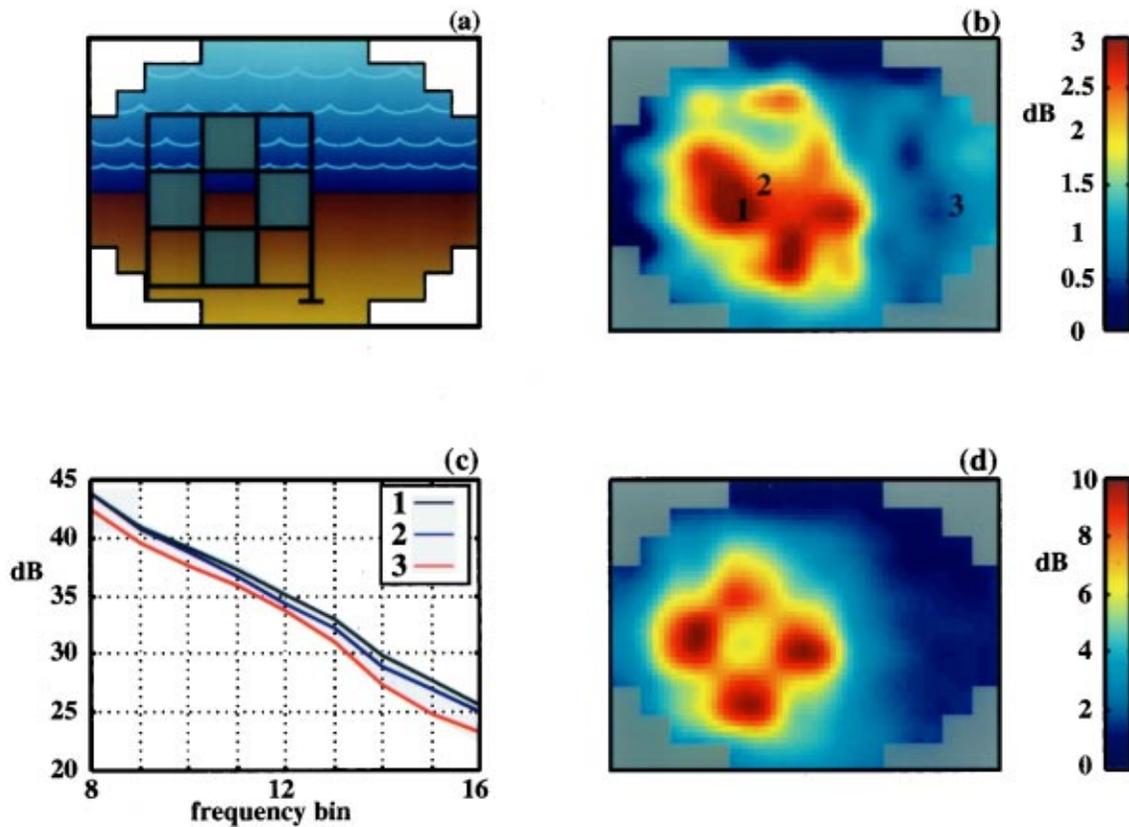


FIG. 10. The fenestrated cross at a range of 38 m. (a) Schematic of the scene falling within the field of view of ADONIS. (b) Example of a poor quality, high-frequency interpolated, intensity mapped image formed by averaging data over the top three frequency bins (57 to 75 kHz). (c) Spectrum, in dB *re* $1 \mu\text{Pa}^2/\text{Hz}$, of the noise in pixels, as identified in (b), on a target panel (1), in the central window (2), and off the target (3). (d) Example of a good quality, high-frequency, interpolated, intensity mapped image formed by averaging data from the top three frequency bins (57 to 75 kHz).

highest frequency, the beamwidth of 0.75 degrees gives an areal coverage of approximately $0.5 \times 0.5 \text{ m}^2$ at the target range, or in other words four beams intersect each $1 \times 1 \text{ m}^2$ target panel. The noise spectra in three pixels, one on and the other two off the target, are illustrated in Fig. 8(c), where it can be seen that the acoustic contrast, represented by the difference between the curves, tends to increase with increasing frequency. Figure 8(d) shows another raw image, formed in this case from an average of the top three frequency bins (57 to 75 kHz). The improvement in resolution obtained at the higher frequencies is quite evident on comparing Fig. 8(b) and (d). When interpolation is applied to the data in Fig. 8(d), the image in Fig. 8(e) is obtained. At these higher frequencies, the target is front ensonified from snapping shrimp on the pier pilings.

Although the target is visible in the raw, broadband data shown in Fig. 8(b), the low resolution resulting from the lower frequencies, combined with the granular pixel structure, give rise to a rather poor quality image. A distinct improvement in resolution can be seen in Fig. 8(d), where the low frequencies have been removed. Visually, the smoothing introduced by interpolation leads to a further improvement in image quality, as evident in Fig. 8(e), where the elongated shape of the bar is recognizable.

As the directionality of noise varies, for whatever reason, the appearance of the images changes accordingly.

Thus, there is no typical acoustic daylight image of a particular target, in the same way that there is no typical photograph of a given scene in daylight. In both cases, front and back illumination, for example, create different shadowing structures and hence different visual effects. The point is well illustrated by the low- and high-frequency, interpolated images of the bar target shown in Fig. 9. The low-frequency image in Fig. 9 is an admittedly crude silhouette, formed by back ensonification from the naval loading dock on the far side of San Diego Harbor. In the higher-frequency image the bar is front ensonified, probably from local snapping shrimp on the MarFac Pier pilings. Although the images in Fig. 9 are both of rather poor quality, the low-frequency silhouette can be seen to align with the high-frequency front-lit image of the bar. It is interesting that the effect seen in these two images is rarely observed in photography, namely, front illumination in one spectral band and, simultaneously, back illumination in another.

B. The fenestrated cross

Figure 10 shows another planar target, in the form of a fenestrated cross [Fig. 10(a)]. The panels, aluminum faced with neoprene foam, were the same as those forming the bar target in Fig. 8. Again, the range was 38 m and the foam side of the panels was facing ADONIS. At the time of the deploy-

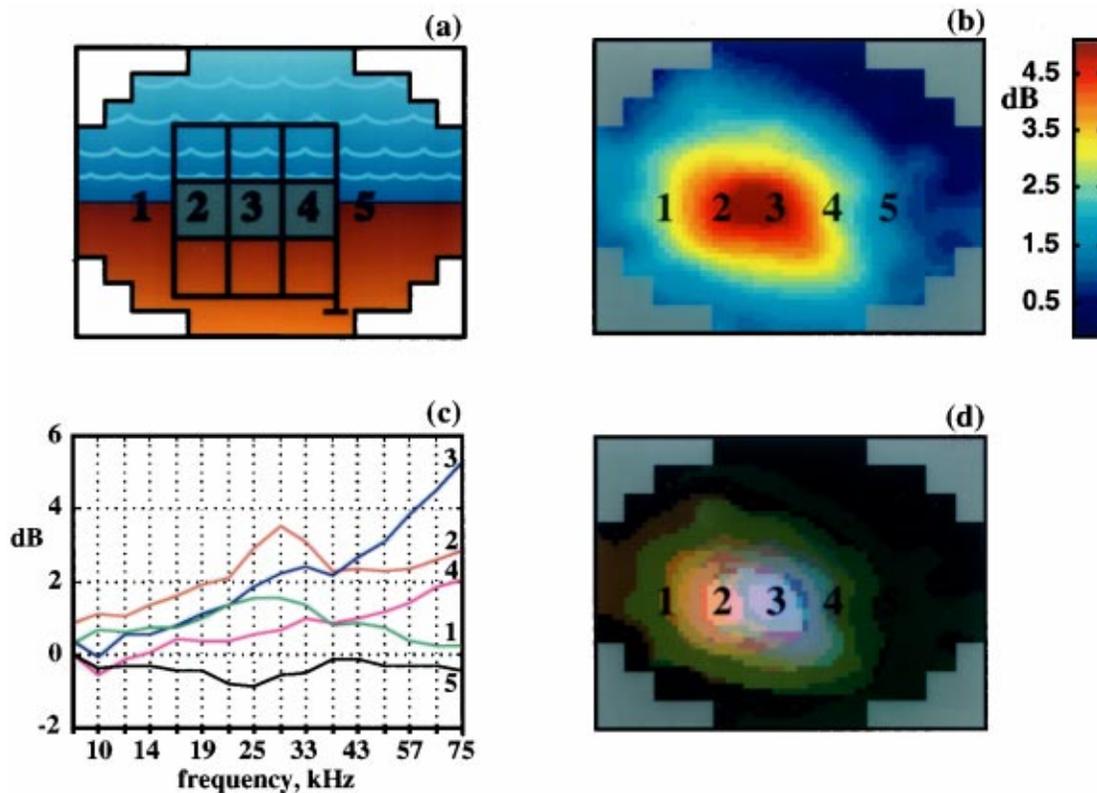


FIG. 11. Three metal panel targets at a range of 40 m. (a) Schematic of the scene falling within the field of view of ADONIS. Panel (2) on the left is 3.2-mm-thick aluminum faced with 6.4-mm-thick neoprene foam, with the metal side facing ADONIS. In the center, panel (3) is of 6.4-mm-thick aluminum. At the right, panel (4) is of 3.2-mm-thick corrugated steel. (b) Intensity mapped, interpolated image averaged over frequencies 24.6 to 75 kHz. (c) Selected spectra from pixels in (b), normalized to the image average. (d) RGB interpolated image formed from the same data as in (b) but with low frequencies mapped into red, middle frequencies into green and high frequencies into blue.

ment it was uncertain whether the central window would be resolved because, at the highest frequency of 75 kHz, the window was just within the limits of resolution of ADONIS (four beams fall within the window). However, the window has been resolved, although the quality of the image varies from frame to frame due to the variability of the ensonifying noise field. Figure 10(b) and (d), respectively, shows examples of a poor- and a high-quality interpolated image, averaged over the top three frequency bins (57 to 75 kHz) in both cases. Note that in Fig. 10(d), the contrast between the panels and the background is about 10 dB, which is significantly higher than expected based on the anisotropy of the noise field that usually prevailed in the area. This was one of the first deployments of ADONIS for which little environmental information is available, but it is known that at around the time the data were taken an extremely noisy angle grinder was being used on the MarFac Pier. The grinder would have provided very strong frontal ensonification, which is consistent with the high contrast in the image.

Even in the case of the poorer quality image [Fig. 10(b)] the central window is just distinguishable, and the contrast across the image is in the region of 3.5 dB. Three spectra from this image are shown in Fig. 10(c), one from a pixel on a panel, one from the background, and the third from the central window. By comparison, the window in the high-quality image [Fig. 10(d)] is clearly resolved and easy to distinguish.

Figure 11 shows a bar target [Fig. 11(a)] formed from

three square panels made of different materials, placed at a range of 40 m. The left panel (labeled 2) was 3.2-mm-thick aluminum sheet faced with 6.4-mm-thick neoprene foam, with the metal side facing ADONIS. In the center (labeled 3) was a 6.4-mm-thick aluminum sheet; and on the right (labeled 4) a 3.2-mm-thick steel corrugated panel of the type often used on sheds and storage buildings. In this deployment the target frame had been moved by approximately 10 degrees to the left of its original position, to avoid back ensonification from the naval loading dock on the far side of the bay.

Figure 11(b) shows an intensity-mapped image averaged over frequency bins 8 to 16 (24.6 to 75 kHz). The metal and foam panel (2) and the metal panel (3) are clearly visible with a contrast around 4.5 dB, while the corrugated panel (4) shows up but with a lesser contrast of 2.5 dB. However, the spectral content of all three panels differs, as can be seen in Fig. 11(c). At the lower frequencies, below 37.5 kHz, panel (2) is the best scatterer of acoustic energy, at higher frequencies panel (3) dominates, and across the full spectral range the corrugated panel (4) is less efficient at scattering than either of the other two panels.

C. RGB mapping

To utilize the spectral differences between the panels, different frequencies may be mapped into separate color components in the image. In this type of representation, the

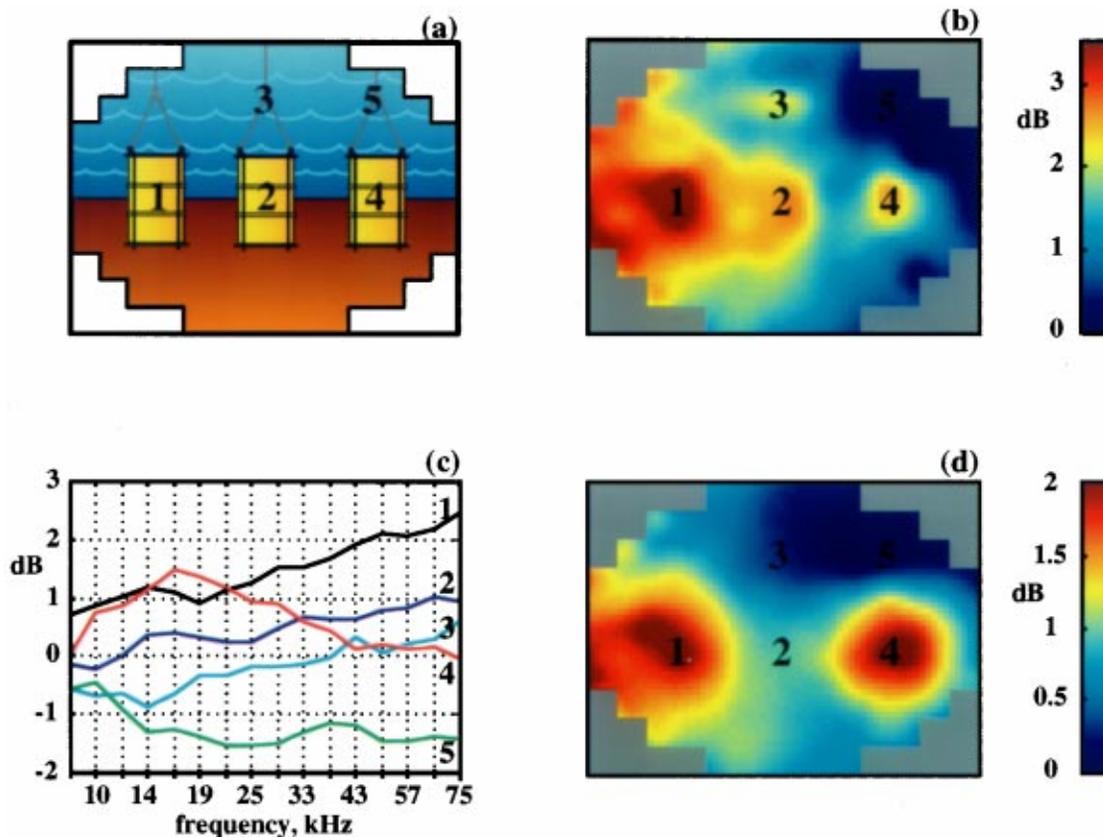


FIG. 12. Suspended drum targets at a range of 20 m. (a) Schematic of the scene falling within the field of view of ADONIS. Left, foam drum (1), center, water drum (2), and right, sand drum (4). (b) High-frequency, intensity mapped, interpolated image formed by averaging over the top three frequency bins (57 to 75 kHz). (c) Selected spectra of pixels in the image, normalized to the image average. (d) Low-frequency, intensity mapped image formed from the same data, averaged over frequencies 10 to 13.8 kHz.

hue and saturation are an indication of frequency content and the luminosity is an indicator of overall intensity. Thus, this type of imaging incorporates frequency and intensity information and the resultant “acoustic color” in the image can be exploited to discriminate between targets of similar shape but differing composition.

To generate Fig. 11(d) the frequency band was divided into three equal parts, with the lower range mapped into red, the middle into green, and the top into blue. Visually, it can be seen in this RGB image that the color of the panels changes from reddish-pink on panel (2), to light pinkish-blue on panel (3), to dark greenish-blue on panel (4). Thus, all three panels can be distinguished from one another simply from their different coloration.

Intensity mapping [Fig. 11(b)] and RGB mapping [Fig. 11(d)] in combination provide extensive information about the composition of the object being imaged. It follows that these mappings show promise for classifying targets. However, much needs to be learned about the acoustic scattering properties of various targets and target materials before the classification potential of the mapping schemes can be fully exploited.

X. IMAGES OF THE CYLINDRICAL TARGETS

A. Drums in the water column

The panel targets discussed above presented a large surface area normal to the axis of the ADONIS dish, making

them relatively easy to image, particularly when front-sonified by the snapping shrimp on the MarFac Pier. Because of their curvature, the cylindrical targets were more challenging, since they presented only a narrow vertical stripe normal to the axis of the dish. Three polyethylene drums were used in the experiments, one filled with syntactic foam (essentially air), a second with seawater, and the third with wet sand.

Figure 12 shows two examples of intensity-mapped, interpolated images obtained when the drums were suspended from the sea surface [Fig. 12(a)] at a horizontal range of 20 m. From left to right the drums contain foam, water, and sand. The high-frequency image in Fig. 12(b) was averaged over the top three frequency bins (57 to 75 kHz), and the lower-frequency image in Fig. 12(d) was averaged over bins 2 to 4 (10 to 13.8 kHz).

In the high-frequency image of Fig. 12(b), all three drums are visible with acoustic contrasts of approximately 4 dB (foam) and 2.5 dB (both water and sand). Additionally, the supporting shackle is visible above the water drum. The picture is different at the lower frequencies [Fig. 12(d)], where the foam and sand drums show a contrast of 2 dB, while the water drum and shackle are not visible at all. These visual features in the two images are consistent with the spectra shown in Fig. 12(c).

Perhaps the most interesting question concerning Fig. 12 is why the water drum is visible in the high-frequency image, given that its acoustic impedance was much the same as that

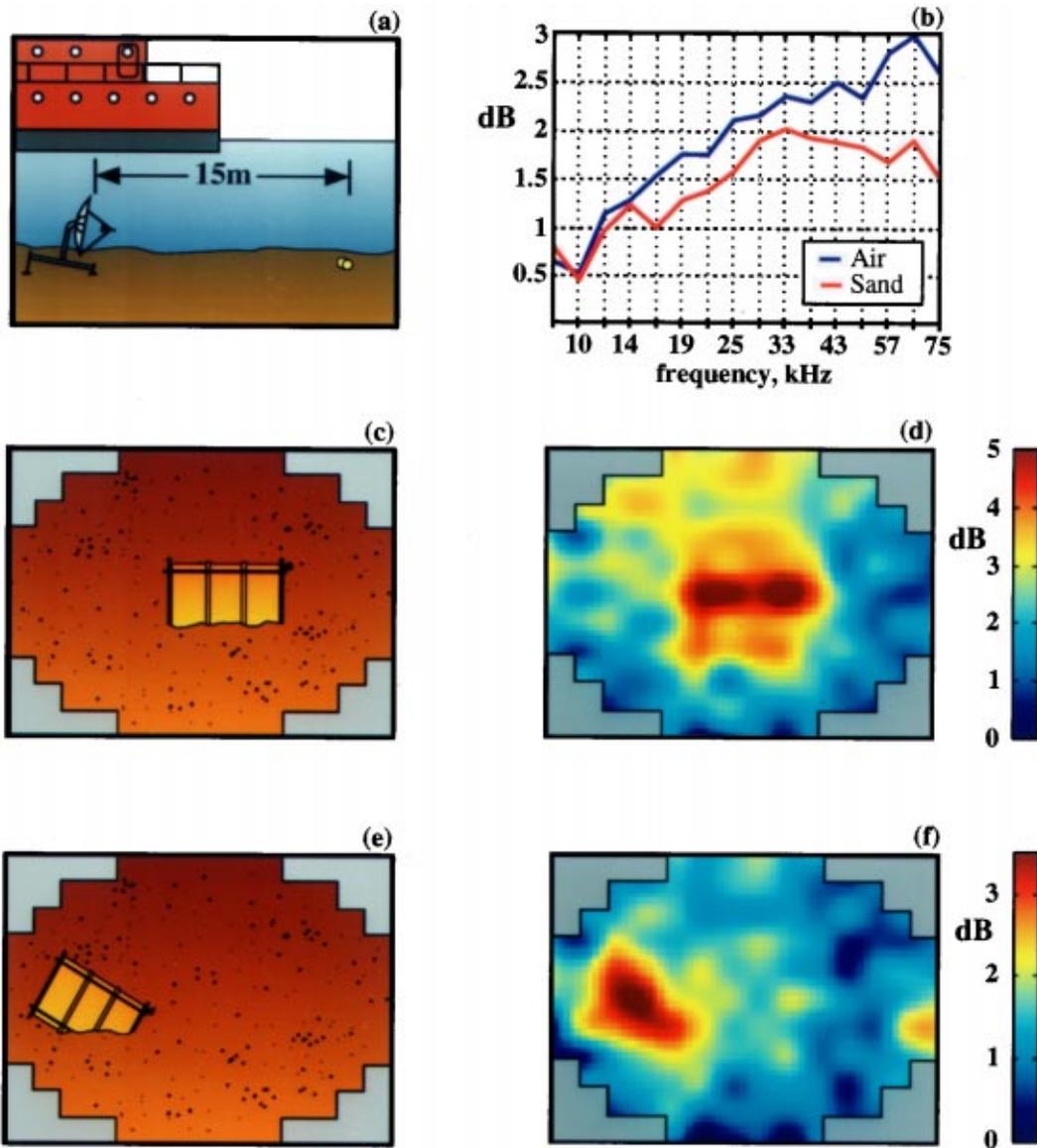


FIG. 13. Drums on the seabed at a range of approximately 15 m. (a) Experimental arrangement, showing ADONIS tilted forward by 10 degrees to point towards the bottom. (b) Spectra of pixels on the foam and sand drums, normalized to the image average. (c) Schematic of ADONIS' view of the foam drum. (d) Intensity mapped, interpolated image of the foam drum, formed by averaging over the top three frequency bins (57 to 75 kHz). (e) Schematic of ADONIS' view of the sand drum. (f) Intensity mapped, interpolated image of the sand drum, formed by averaging over the top three frequency bins (57 to 75 kHz). Note that in the image the axis can be seen to be tilted by approximately 45 degrees.

of the background water. The answer to this may lie in the way the drums were deployed. The sand drum was negatively buoyant by at least 135 kg and the foam drum was weighted to be negatively buoyant by 55 kg. The water drum, however, was weighted only with the 10 kg of its supporting cage. In a series of images created with shorter averaging times, the water drum can be seen to sway back and forth, a movement which causes the rusty metal components of the shackle and drum cage to rub together and generate noise. Spectra 2 and 3 in Fig. 12(c) are similar in shape, consistent with both being generated by the same mechanism. Moreover, the noise from the water drum and shackle was observed to rise when large boat wakes passed over the targets, while the noise from the foam and sand drums remained essentially unchanged. Based on this evidence, it seems that the water drum and shackle are visible in the

high-frequency image through their own self-noise rather than through scattering of the ambient noise field.

B. Drums on the bottom

To investigate the imaging performance of ADONIS against targets on the seabed, the foam and sand drums were held horizontally and dropped from the surface onto the bottom at ranges of 13 m and 15 m, respectively, from the dish. ADONIS was tilted forward by 10 degrees [Fig. 13(a)] so the drums would be in the center of the field of view. On impact, the drums became partially buried in the soft mud bottom, with 30%–50% of their volume below the interface. The sand drum landed slightly askew, with its axis angled down into the mud at approximately 45 degrees [Fig. 13(e)]. The axis of the foam drum was essentially horizontal [Fig. 13(c)],

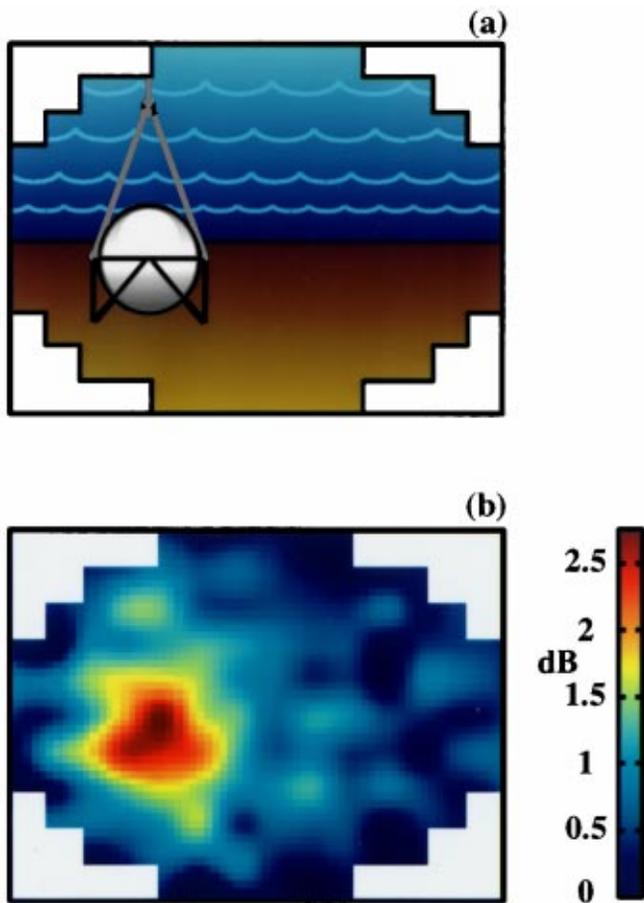


FIG. 14. Spherical titanium shell at a range of 25 m. (a) Schematic of the scene falling within the field of view of ADONIS. (b) Intensity mapped, interpolated image formed by averaging over the top three frequency bins (57 to 75 kHz).

and in both cases the axes were perpendicular to the axis of the dish.

The acoustic daylight image of the foam drum, formed from an average of the top three frequency bins (57 to 75 kHz), is shown in Fig. 13(d). Although the background is now the sediment, the acoustic contrast in the image is approximately 4 dB, which is essentially the same as when the drum was suspended in the water column [Fig. 12(b)]. An image of the sand drum, also averaged over the top three frequencies, is shown in Fig. 13(f). The contrast in this case is approximately 3 dB, which is similar to the contrast observed when the drum was suspended in the water column. By comparing Fig. 13(d) and (f), it is apparent from the images that the axis of the foam drum is horizontal and that of the sand drum is tilted at about 45 degrees. Figure 13(b) shows the normalized spectra of pixels on each of the bottom deployed drums. These spectra show much the same trends as those of the suspended drums in Fig. 12(c), which is consistent with the similar contrast in the images of the drums on the bottom and in the water column.

XI. IMAGING THE SPHERICAL TARGET

Geometrically, the titanium sphere represents a more challenging target than either the planar panels or the cylin-

drical drums. With a surface showing curvature in two orthogonal directions, the sphere presents only a small, spot-like area of surface normal to the axis of ADONIS, compared with a stripe in the case of the drums and a square for each of the panels. This difficulty is offset by the fact that the sphere is actually an air-filled shell with an acoustic impedance significantly different from that of water. Like the foam drum, it may therefore be expected to act as a good target for acoustic daylight imaging. As with the drums, the sphere was suspended from surface floats at a horizontal range of 25 m from the dish [Fig. 14(a)]. Deployment ropes were tied directly to the cage supporting the sphere, avoiding the use of a shackle, to alleviate the problem of self-noise in the system.

Figure 14(b) shows the interpolated acoustic daylight image of the sphere formed by averaging the top three frequency bins (57 to 75 kHz). The equator and upper hemisphere are clearly visible in the image, with a maximum contrast close to 2.5 dB. Below the equator, the lower hemisphere is difficult to distinguish, indicating that any energy scattered from this region must be several dB less than that from the upper hemisphere. In several respects, this image resembles Potter's⁵ numerical simulation of a sphere ensonified by surface sources in shallow water.

At the time the data for Fig. 14 were collected, however, the surface was calm and there were few surface sources. Most of the noise was generated by snapping shrimp on the pilings of the MarFac Pier, which provided horizontal, front ensonification. This raises the question as to why the top hemisphere and the equator should show a significantly stronger contrast than the lower portion of the sphere. One possibility is that the incident, horizontally traveling sound was deflected upwards and downwards by the upper and lower hemispheres, respectively. The upward deflected sound then encountered the highly reflective sea surface at near normal incidence, turned around and re-ensonified the top of the sphere, where it was scattered a second time, back towards ADONIS, to yield the high contrast in the image. The downward deflected sound behaved similarly except that it was reflected from the heavily attenuating bottom, a fine silty mud. There was, therefore, little energy available to ensonify the lower hemisphere from below and the acoustic contrast in this region of the image is correspondingly reduced.

As a consequence of the highly variable nature of the snapping shrimp noise (Fig. 6), successive frames, which are separated by intervals of only 40 ms, in general show significant variations in intensity. This is illustrated in Fig. 15, where a sequence of 12 successive frames of the spherical target is shown, with no temporal averaging. The sequence of frames was created from the top frequency, bin 16 (75 kHz), of the first 0.5 s of data used to form the image in Fig. 14(b). The sphere is difficult to identify in most of the images because of the high dynamic range of 20 dB, although it does stand out clearly in frame 10. Epifanio¹⁰ presents further sequences of images, which indicate that at least 1 in 25 frames contains a bright spot associated with a target (see Fig. 7.18 in Ref. 10); he also discusses the effect of averaging times on image quality.

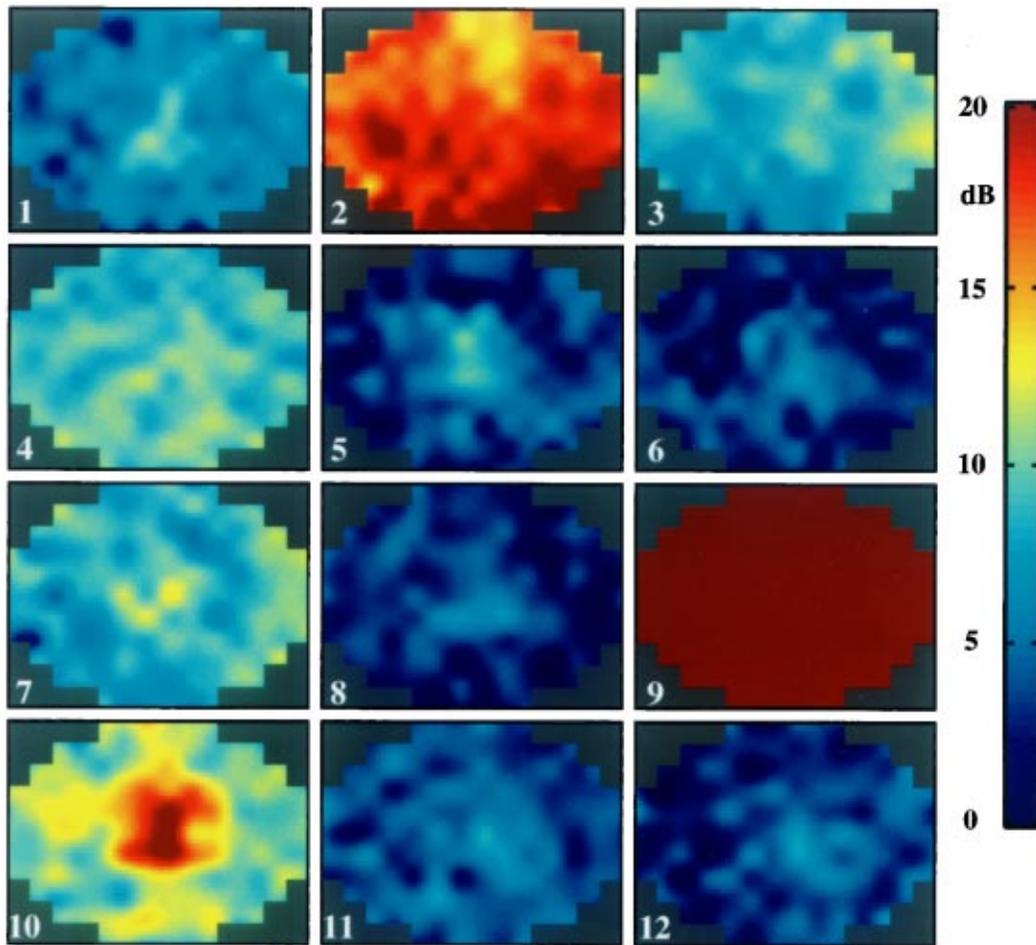


FIG. 15. Sequence of successive, intensity mapped interpolated images of the titanium sphere, formed from the top frequency bin 16 (75 kHz). The duration of this sequence of frames is approximately 0.5 s.

XII. CONCLUDING REMARKS

The images in this article are the first to have been obtained of objects in the ocean using naturally generated sound as the ensonification. In these early ORB experiments, the imaging system (ADONIS) was deployed close to shore in San Diego Bay, at the end of the MarFac Pier, and the principal source of noise was snapping shrimp inhabiting the pier pilings. A casual examination of the images reveals that the resolution falls far short of optical imaging, simply because of the limited aperture of the acoustic reflector used in the experiments: the pupil of the human eye is about 20 000 optical wavelengths across, whereas the diameter of the ADONIS dish is only 150 acoustic wavelengths at the highest operating frequency. Nevertheless, the images are recognizable in most cases, especially at the higher frequencies where the angular resolution is reasonable (approximately 0.6 m at a range of 50 m).

As with conventional photography, the appearance of a given scene in an acoustic daylight image shows considerable variability, depending particularly on the anisotropy of the ambient noise field at the time the data were collected. In other words, the shadowing in the images depends strongly on the directionality of the ensonification. During the experiments, as much environmental information as possible was recorded, including a measure of the horizontal directionality

of the noise. It was not feasible, however, to measure the horizontal anisotropy of the noise every time that imaging data were taken, making interpretation of some of the detailed, shadowing features in the images uncertain. However, considerable care was taken over the alignment of the ADONIS receiver and the targets, and checks were performed repeatedly to ensure that regions of enhanced contrast in the images are not artifacts of the system. The essential conclusion is that the acoustic daylight images obtained in the ORB experiments are genuine representations of the object space within the field of view of ADONIS.

ACKNOWLEDGMENTS

ADONIS was designed and constructed, and the first images were obtained in August 1994, under the direction of John Potter. Subsequently, as part of his Ph.D. research, Chad Epifanio introduced several improvements into the system and obtained many more images in October 1995. Nicholas Carbone, Mark Brentnall, David Sigurdson, Gene King, and Fernando Simonet contributed significantly to the fabrication of ADONIS. Eric Slater, Lloyd Green, and Michael Goldin generously provided assistance with a variety of engineering problems. The participation of Dr. Mark Readhead was made possible by the award of the Royal Australian Navy Science Scholarship for 1995. Sustained sup-

port for the acoustic daylight research project has been provided by the Office of Naval Research under Grant No. N00014-93-1-0054, for which we are grateful.

- ¹S. Flatté and W. Munk, "Submarine detection: Acoustic contrast versus acoustic glow," Report number JSR-85-108, MITRE Corporation, McLean, Virginia (1985).
- ²M. J. Buckingham, "Ambient noise as a source of acoustic illumination in the ocean," Technical memorandum (April 1987). (Accepted by the United States Patent and Trademark Office, Serial No. 08/012,894, February 3, 1993 in support of Patent Number 5,333,129 on Acoustic Imaging in the Ocean using Ambient Noise, assigned July 26, 1994 to the Regents of the University of California, inventor, Michael J. Buckingham.)
- ³M. J. Buckingham, B. V. Berkhout, and S. A. L. Glegg, "Imaging the ocean with ambient noise," *Nature (London)* **356**, 327–329 (1992).
- ⁴M. J. Buckingham, "Theory of acoustic imaging in the ocean with ambient noise," *J. Comput. Acoust.* **1**, 117–140 (1993).
- ⁵J. R. Potter, "Acoustic imaging using ambient noise: Some theory and simulation results," *J. Acoust. Soc. Am.* **95**, 21–33 (1994).
- ⁶B. D. Steinberg and H. M. Subbaram, *Microwave Imaging Techniques* (Wiley, New York, 1991).
- ⁷N. C. Makris, F. Ingenito, and W. A. Kuperman, "Detection of a submerged object insonified by surface noise in an ocean waveguide," *J. Acoust. Soc. Am.* **96**, 1703–1724 (1994).
- ⁸M. J. Buckingham and J. R. Potter, "Acoustic daylight imaging: vision in the ocean," *GSA Today* **4**(4) (1994).
- ⁹M. J. Buckingham, J. R. Potter, and C. L. Epifanio, "Seeing underwater with background noise," *Sci. Am.* **274**, 86–90 (1996).
- ¹⁰C. L. Epifanio, "Acoustic Daylight: Passive Acoustic Imaging using Ambient Noise," Ph.D. thesis, Scripps Institution of Oceanography, University of California, San Diego, 1997.
- ¹¹R. H. Mellen, "The thermal-noise limit in the detection of underwater acoustic signals," *J. Acoust. Soc. Am.* **24**, 478–480 (1952).
- ¹²H. B. Callen and T. A. Welton, "Irreversibility and generalized noise," *Phys. Rev.* **83**, 34–40 (1951).
- ¹³G. B. Deane, "The beam forming properties of a concave spherical reflector with an on axis receiver," *J. Acoust. Soc. Am.* **106**, 1255–1261 (1999).
- ¹⁴A. L. Anderson, "Ambient-noise measurements at 30, 90, and 150 kHz in five ports," *J. Acoust. Soc. Am.* **49**, 928–930 (1971).
- ¹⁵F. A. Everest, R. W. Young, and M. W. Johnson, "Acoustical characteristics of noise produced by snapping shrimp," *J. Acoust. Soc. Am.* **20**, 137–142 (1948).
- ¹⁶D. H. Cato and M. J. Bell, "Ultrasonic ambient noise in Australian shallow waters at frequencies up to 200 kHz," Report No. MRL-TR-91-23, Defence Science and Technology Organisation, Materials Research Laboratory, Sydney (1992).
- ¹⁷G. C. Jensen, "Pacific Coast Crabs and Shrimps" (Sea Challengers, Monterey, 1995).
- ¹⁸M. W. Johnson, F. A. Everest, and R. W. Young, "The role of snapping shrimp (*Crangon* and *Synalpheus*) in the production of underwater noise in the sea," *Biological Bulletin* **93**, 122–138 (1947).
- ¹⁹D. Ross, *Mechanics of Underwater Noise* (Pergamon, New York, 1976).
- ²⁰D. M. Farmer and S. Vagle, "Observations of high frequency ambient sound generated by wind," in *Sea Surface Sound* (Kluwer, Dordrecht, 1988), pp. 403–416.
- ²¹Naval Oceanographic Office, "Shallow water ambient noise levels in the Tongue of the Ocean, Bahamas, Fall of 1965 and Summer of 1966," Report No. IR No. 69–57 (1969).
- ²²K. S. N. Namboodiripad, "Acoustical ambient noise of the Cochin Backwaters," *Indian J. Marine Sci.* **7**, 75–78 (1978).
- ²³J. Ahn, E. Hamada, and K. Saito, "Study on the positioning of snapping shrimp and source levels of their sounds," *Tokyo Univ. Fisheries* **80**, 75–81 (1993).
- ²⁴M. Readhead, "The distribution of snapping shrimp noise near Gladstone, Queensland," Report number DSTO-TR-0047, Defence Science and Technology Organisation, Sydney (1994).

Automatic matched-field tracking with table lookup

Homer Bucker^{a)} and Paul A. Baxley

SPAWAR Systems Center, San Diego, California 92152-5001

(Received 18 September 1998; revised 14 September 1999; accepted 15 September 1999)

In matched-field processing, the signals measured at sensors in an array are correlated with the expected signals for a target at a specified point in the environment. This tends to be a computer intensive process because there are a large number of possible target locations and the calculation of the sound field may be complicated. A search by an operator for the target involves examination of various cuts through a multi-dimensional search space. In this paper it is shown that use of table lookup simplifies the modeling problem. Also, by use of matched-field tracking, the data can be processed in large time (~ 3 to 10 min) blocks. Using these methods, the processing can be done in real time with little, or no, operator intervention. An example of automatic matched-field tracking is presented using data collected during a shallow water test near San Clemente Island. Finally, a simple sensitivity study is made. The actual sound speed profile is replaced by one modified by low-pass filtering and, in a second case, by a constant speed profile. The results show that matched-field tracking, although degraded, still provides useful tracking data. © 1999 Acoustical Society of America. [S0001-4966(99)07212-4]

PACS numbers: 43.30.Bp [SAC-B]

INTRODUCTION

In matched-field processing,¹ the signals measured at sensors in an array are correlated with the expected signals for a target at a specified point in the environment. This tends to be a computer intensive process because there are a large number of possible target locations and the calculation of the sound field may be complicated. A search by an operator for the target involves examination of various cuts through a multi-dimensional search space.

In this paper we show that use of table lookup simplifies the modeling problem. Also, by use of matched-field tracking,² the data can be processed in large time (~ 3 to 10 min) blocks. Using these methods, the processing can be done in real time with little, or no, operator intervention. An example of automatic matched-field tracking will be presented using data collected during a shallow water test near San Clemente Island.³

Consider an array consisting of J sensors. Let F_j represent the signal received at a typical sensor. One method of forming a matched-field correlation value is to first form the matched-field beam $B(x,y,z)$, where

$$B(x,y,z) = \sum_j \hat{F}_j(x,y,z) F_j^* \quad (1)$$

In Eq. (1), F_j is the complex value of a Fourier coefficient with bin center at frequency f , $\hat{F}_j(x,y,z)$ is the expected signal at sensor j if the target has frequency f and is at location x,y,z , and the symbol $*$ indicates complex conjugate. In general, the vector F_j will not have a frequency exactly equal to f so there will be a rotation of B that is stabilized by processing the beam with a square-law detector. Therefore, if

the source does not move too far over a small time period, the matched-field correlation function can be written as $C(x,y,z)$ where

$$C(x,y,z) = \langle BB^* \rangle \quad (2)$$

In Eq. (2) the symbols $\langle \rangle$ indicate a time average.

As the sound source moves during larger time periods, the peak value of C can be tracked if the propagation model used to calculate F_j is accurate. One method of finding the track is to use various forms of a Kalman filter.⁴ The method of matched-field tracking used in this paper² is to concentrate on the end points (A and B) of a track segment.

I. MATCHED-FIELD TRACKING

In matched-field tracking, sensor data is collected over a 3- to 10-min time period. The process is to focus on the end points (A and B) of a potential track and to calculate a correlation value $C_T(A,B)$ for the track:

$$C_T(A,B) = \sum_t \sum_f C(x,y,z,f,t) \quad (3)$$

In Eq. (3) the sums are over frequency f and time t . Values of x , y , and z are calculated as functions of t using the end points A and B and t . Doppler corrections can be made, if significant, to \hat{F}_j because the velocity of the target can be calculated at each time. An alternate method is to calculate C_T as the match between measured and calculated off-diagonal elements of the covariance matrix. This method resulted in similar results.

II. RANGE DEMODULATION

In later sections, matched-field tracks will be calculated for a source tow in a shallow water test. Figure 1 shows the real and imaginary parts of the calculated sound field as a function of range for a typical source depth, receiver depth,

^{a)}Electronic mail: bucker@nosc.mil

Complex Pressure vs Range, Tgt. Depth = 60 m, Rcv. Depth = 100 m, Freq = 112 Hz

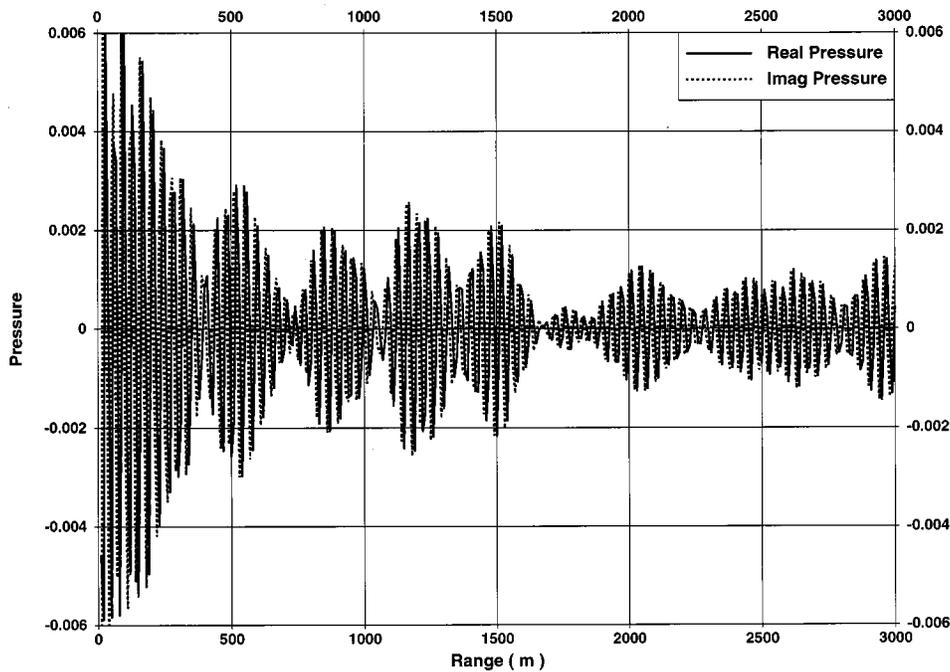


FIG. 1. Complex sound pressure as a function of range. Source depth = 60 m. Sensor depth = 100 m. Frequency = 112 Hz.

and frequency. A range-independent normal mode model was used for the calculation and the pressure is in arbitrary units. It is seen that there is a high-frequency modulation of the pressure with range. To reduce the modulation, $\exp(ik_0r)$ is factored from the pressure, where k_0 is a reference wave number. The result is shown in Fig. 2. It is clear that these demodulated values can be stored in a lookup table of reasonable size. To generate Fig. 2 a value of $k_0 = 2\pi f/1510$ was used, but this value is not critical. For the reference function \hat{F}_j used in Eq. (1) we will use

$$\hat{F}_j(r) = \exp(ik_0r)\tilde{F}_j(r), \quad (4)$$

where $\tilde{F}_j(r)$ is obtained by linear interpolation of values from a table derived by decimation of data shown in Fig. 2.

III. TEST RESULTS

Data was collected during track s5 of the SWellEX-96³ test. For this paper we used signals from six hydrophones of a nearly vertical line array in 216 m of water. Signals from

Range Modulated Pressure, Tgt. Depth = 60 m, Rcv. Depth = 100 m, Freq = 112 Hz

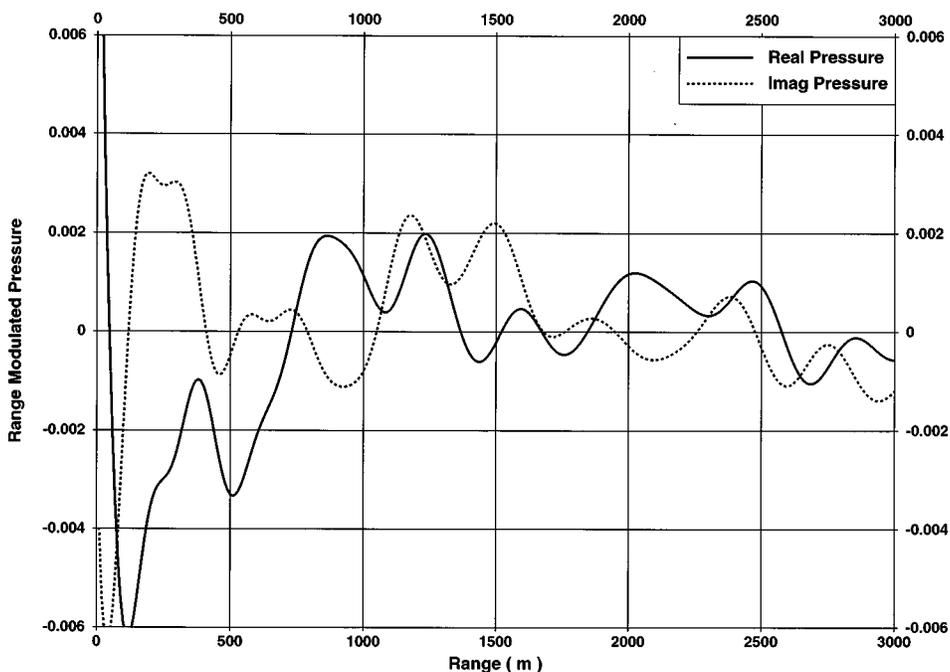


FIG. 2. Range-demodulated sound pressure as a function of range. Source depth = 60 m. Sensor depth = 100 m. Frequency = 112 Hz.

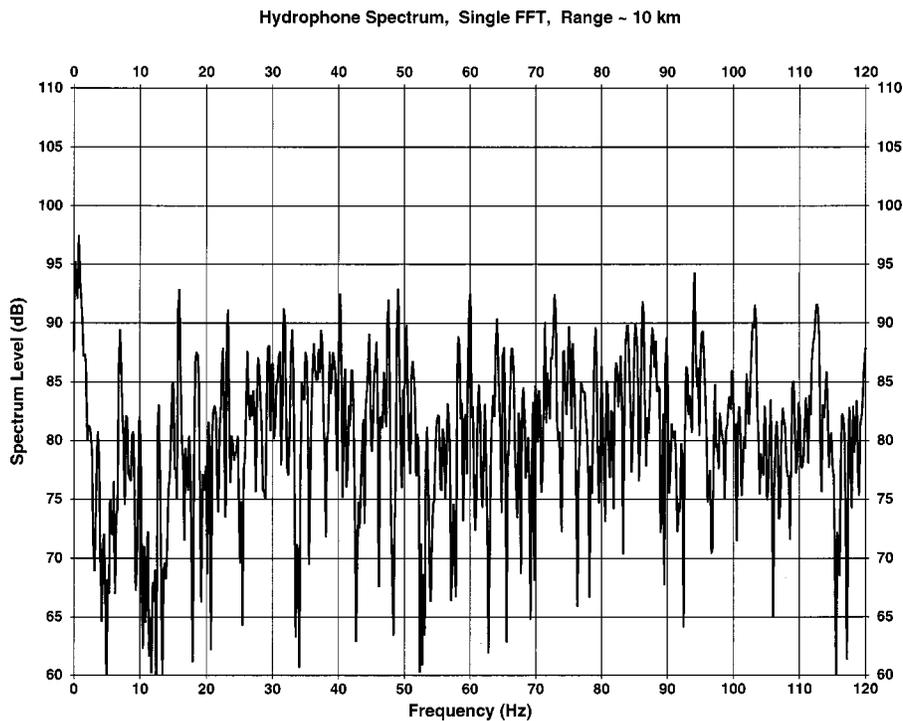


FIG. 3. Spectrum level ($\text{dB}/\mu\text{Pa}^2/\text{Hz}$) of signal recorded on a sensor near the center of the array at a range of ~ 10 km.

the six sensors, uniformly spaced from 97.6- to 210.1-m depths, were processed. The tow was along a straight line, except at the end of the tow, with a closest point of approach (CPA) of ~ 1 km. The nominal depth of the towed source was ~ 62 m. During most of the 85-min tow a number of tonals were transmitted. At several points along the tracks the tonals were turned off for a 1-min period. This occurred at ranges of ~ 8.5 , 6.0, 5.4, 3.0, and 1.1 km before the CPA and at ~ 1.1 and 3.0 km afterwards.

The spectrum of the signal recorded at a typical hydrophone when the target was at a range of $\times 10$ km is shown in Fig. 3. The Fourier bins have a bandwidth of ~ 0.18 Hz. For the matched-field tracking we used five Fourier bins at approximate frequencies of 49, 64, 79, 94, and 112 Hz and processed the data in ~ 6 -min segments.

A typical sound speed profile, measured *in situ*, is shown in Fig. 4. The bottom model was a 30-m sediment layer lying over basement. Sediment density was 1.76 g/cm^3 . At the top of the layer, $v_c = 1572.4 \text{ m/s}$, $\alpha_c = 0.2 \text{ dB/m_kHz}$, $v_s = 112.6 \text{ m/s}$, and $\alpha_s = 17.3 \text{ dB/m_kHz}$. At the bottom of the layer, $v_c = 1593.0 \text{ m/s}$, $\alpha_c = 0.2 \text{ dB/m_kHz}$, $v_s = 238.0 \text{ m/s}$, and $\alpha_s = 17.3 \text{ dB/m_kHz}$. In a normal mode model, developed by the first author, the sediment layer was represented by 40 homogeneous visco-elastic micro-layers. For the basement, the density was 2.66 g/cm^3 , $v_c = 5200.0 \text{ m/s}$, $\alpha_c = 0.02 \text{ dB/m_kHz}$, $v_s = 2550 \text{ m/s}$, and $\alpha_s = 3.4 \text{ dB/m_kHz}$. Figure 5 is a plot of bottom depth versus range measured from the beginning of the track. Although there is significant depth variation along the track, it was possible to use a range-independent propagation model because of the short ranges in the test. The processing was in two stages. In the first stage, it was assumed that the array was vertical and that $y_A = y_B$ and $z_A = z_B$ so that the search space had four dimen-

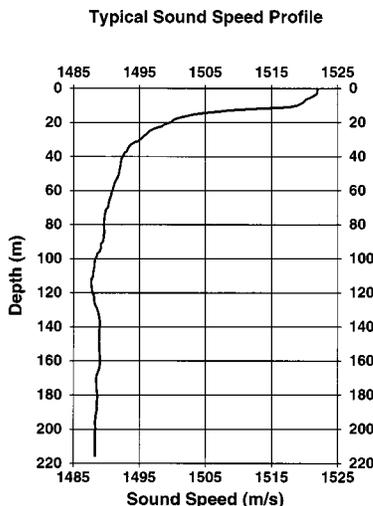


FIG. 4. Typical sound speed profile measured during the test.

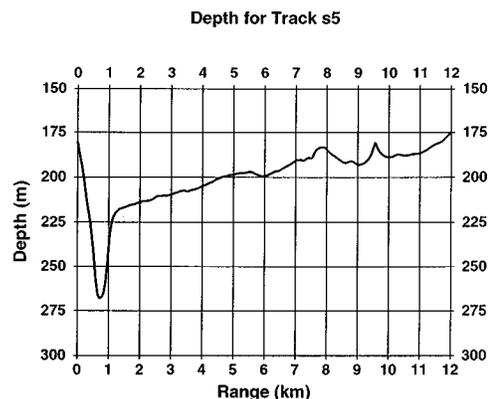


FIG. 5. Bottom depth versus range measured from the beginning of the track.

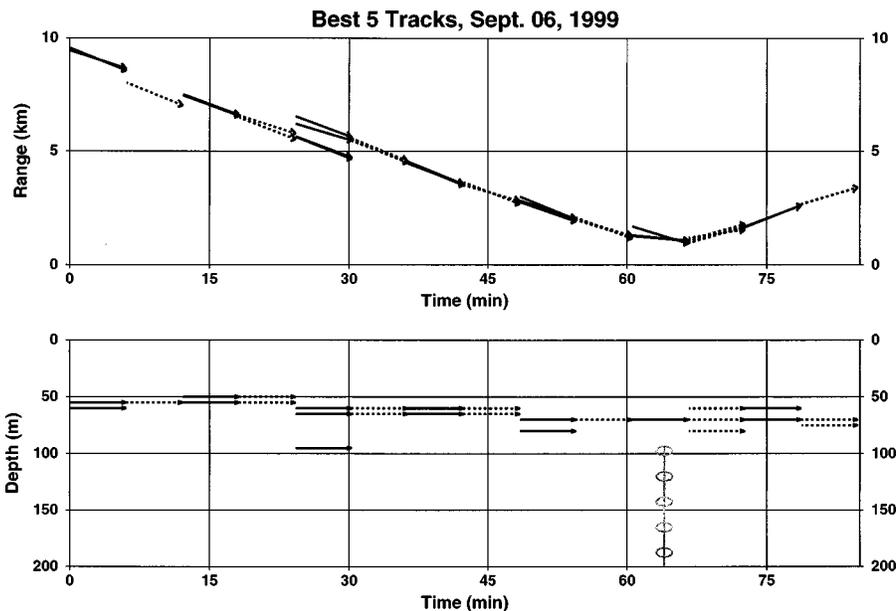


FIG. 6. Target tracks for source tow s5, SWellEX-96 test.

sions. Values of $C(A,B)$ were calculated for a rough grid of end points that satisfy a target speed >3 knots and <10 knots. Possible values of x_1 and x_2 were $-10.0, -9.5, -9.0, \dots, +3.5$ km. Values of y were $0.0, 0.2, 0.4, \dots, 2.0$ km, and values of z were $30, 40, 50, \dots, 120$ m. In all, 5610 possible tracks were examined. The five tracks with the largest correlation values were selected for further processing. In the second stage, the search space was increased to seven dimensions and each parameter was adjusted for maximum correlation. The new variables are an independent y_B and a_1 and a_2 , coefficients of a power series that defines the shape of the array. Also, the value of z was allowed to change in 5-m steps.

The following was used to calculate the sensor locations. Let u_j be the distance of sensor j above the bottom. For convenience, consider the sensors to be in the (x,u) plane. Let $x(u) = a_1x^1 + a_2x^2 + \dots$. Let s_j be the distance along the array, from the bottom, to sensor j . Then if the array coefficients and s_j are known, we can determine x_j and u_j by a numerical solution of u_j such that s_j is the value of the integral of ds evaluated from $u=0$ to $u=u_j$, and $(ds)^2 = (dx)^2 + (du)^2$.

Processing results are shown in Fig. 6. The best five tracks in each ~ 6 -min segment are represented by arrows. In the odd-numbered segments the arrows are solid. The arrows are dashed in the even-numbered segments. It is seen that the track depth is less than the actual depth of the source (~ 64 m) at the beginning of the run and is slightly greater at the end. This is consistent with the changing depth of the shallow water channel as discussed by D'Spain.⁵

IV. SENSITIVITY OF THE MODEL

In matched-field processing it is obvious that the quality of results depends on the accuracy of the propagation model. In matched-field tracking this constraint may not be as severe. For example, a 30% error in track segment parameters might be acceptable if the track holds together. For a simple look at this problem, we will use a smoothing filter to distort

the sound speed profile. First, the profile is reflected about the surface and reflected about the bottom. Next a box-car filter spanning 81 m was used to smooth the profile. The modified profile is shown in Fig. 7. The corresponding tracks are shown in Fig. 8. Although the track estimates are degraded, the track motion can be discerned.

Finally, new lookup tables are calculated for an isospeed profile where the sound speed is the average speed in the channel ($c = 1492$ m/s). Note that the bottom interaction will be different now because the sound speed in the water at the bottom is changed by ~ 4 m/s. The tracking results are shown in Fig. 9. Now there are significant errors in range and depth at the long ranges. Still, there is useful target information.

V. CONCLUSION

An example of automatic matched-field tracking has been shown. By use of table lookup the algorithm will oper-

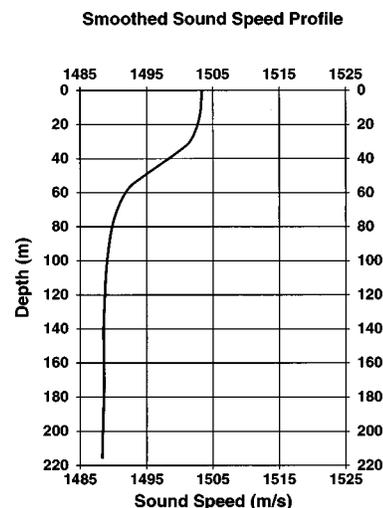


FIG. 7. Modified sound speed profile.

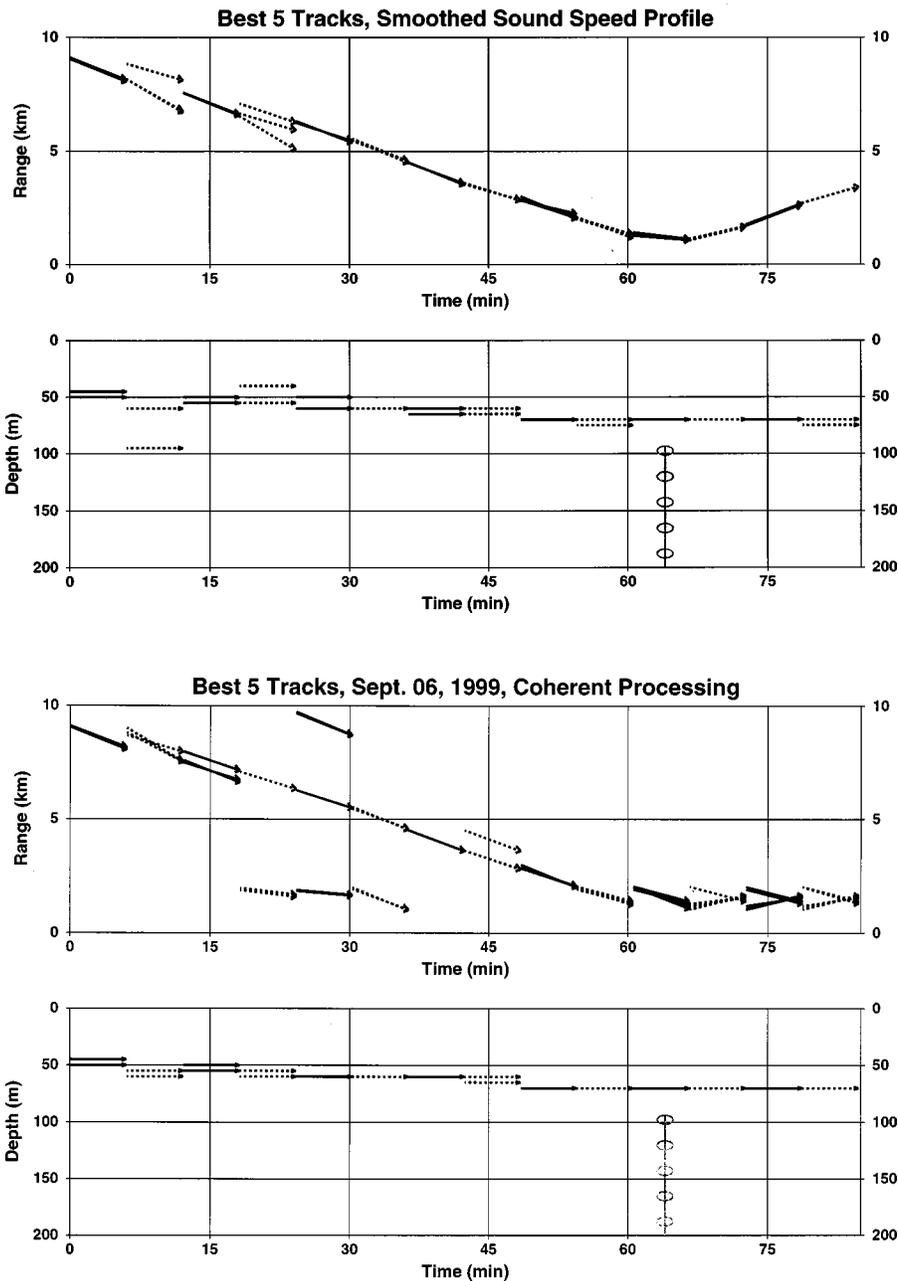


FIG. 8. Target tracks using modified/smoothed sound speed profile.

FIG. 9. Target tracks using constant sound speed profile, $c = 1492$ m/s.

ate in real time on a typical PC computer. A simple sensitivity example shows that useful tracks can be obtained using degraded propagation models.

ACKNOWLEDGMENTS

The authors would like to acknowledge help from Dr. Newell Booth, who directed the SWellEX-96 experiment, Phillip Schey, who generated the FFT tapes, and Richard Bachman, who provided the bottom sediment data. Support for this work was provided by the Office of Naval Research, Code 321-SS, and by the Space and Naval Warfare Systems Center IR Program.

¹A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).

²H. Bucker, "Matched-field tracking in shallow water," *J. Acoust. Soc. Am.* **96**, 3809–3811 (1994).

³N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D'Spain, and J. J. Murray, "Source localization with broad-band matched-field processing in shallow water," *IEEE J. Ocean Engr.* **21**, 402–412 (1996).

⁴J. C. Hassab, *Underwater Signal and Data Processing* (CRC, Boca Raton, FL, 1989).

⁵G. L. D'Spain, J. J. Murray, W. S. Hodgkiss, and N. O. Booth, "Predicting the broadband matched field processing results for the arc event during SWellEX-3," MPL Technical Memorandum 442 (October 1994).

A three-dimensional propagation algorithm using finite azimuthal aperture

Kevin B. Smith^{a)}

Department of Physics, Naval Postgraduate School, Monterey, California 93943

(Received 1 October 1998; revised 13 August 1999; accepted 17 August 1999)

The use of three-dimensional (3-D) propagation models is becoming increasingly common in current underwater acoustics applications. Such models typically treat the propagation as a function of range, depth, and bearing, consistent with previous multibearing two-dimensional ($N \times 2$ -D) models. However, to obtain accurate 3-D solutions at long ranges, many bearings must be computed in order to maintain the necessary cross-range resolution between bearings. In this paper, a 3-D parabolic equation (PE) model is developed with a marching algorithm based on the split-step Fourier technique in both depth and azimuthal bearing. The algorithm includes a scheme to compute the solution for only a finite azimuthal aperture rather than the full 360° of bearing. The success of this algorithm allows for increased cross-range resolution at long range without increasing the number of bearings needed in the calculation. Results of this technique are compared to results from a suggested benchmark case. [S0001-4966(99)02712-5]

PACS numbers: 43.30.Bp [SAC-B]

INTRODUCTION

In recent years, our ability to accurately and efficiently compute acoustic propagation in ocean waveguides has continued to improve. Much of the effort has focused on two-dimensional (2-D) approximations to the wave equation producing solutions of the acoustic field in range and depth. Such solutions neglect out-of-plane scattering and refraction by making an uncoupled azimuth (UNCA) approximation. In cases with azimuthal symmetry, such approximations are exact. For many environments and/or applications, this approximation remains valid in the presence of small azimuthal variations.

For some environments, however, the azimuthal variations may significantly influence the propagation, especially when propagating long ranges such as ocean basin scales.¹ The influence of 3-D fluctuations in shallow water environments is currently being addressed by several researchers.²⁻⁴ In order to deal with such cases, various models have been adapted (or are being developed) that include, at least to some level of approximation, the influence of azimuthal coupling. Although continually being upgraded and developed, some of these are now currently available to the general acoustics community.⁵⁻⁷

Because of the increased computational load of such models, algorithms need to be as efficient as possible while maintaining accuracy. It is natural then to consider only the outgoing propagation from a localized source distribution rather than attempt to compute the full, two-way (outgoing and incoming) field solution. It is also natural to cast the problem into the form of solutions along outgoing radials in a cylindrical coordinate system rather than less convenient Cartesian coordinates. (The effects of Earth curvature will not be addressed.) However, one quickly runs into problems when considering the issue of environmental resolution be-

tween bearings. As the range from the source increases, the arc distance between bearings of fixed azimuthal separation grows linearly. If this distance grows too large, the environmental azimuthal variability may be grossly undersampled. Even for smooth cross-range variations, undersampling the field in azimuth may limit the model's ability to correctly predict the horizontal refracted nature of the acoustic paths. As an analogy, consider the influence of grossly undersampling the acoustic field in depth and the relationship with vertical propagation angles. To counter this problem, greater numbers of azimuths must be computed. Obviously, this problem can become overwhelming at longer ranges.

In this paper, a model algorithm is developed that attempts to alleviate some of the problems associated with the need for large numbers of azimuths by computing only a limited aperture of azimuths at longer ranges. In Sec. I, the general theoretical development of the 3-D acoustic model is given. The numerical implementation of the limited azimuthal aperture algorithm is described in Sec. II. Results of this algorithm are then compared with other solutions and a suggested benchmark solution in Sec. III. In Sec. IV we provide a summary of the results of this work and suggest future improvements.

I. THEORETICAL BACKGROUND

In this section, the formulas used in the numerical algorithm for the 3-D model are derived. It is based on a parabolic equation (PE) approximation to the full acoustic wave equation in cylindrical coordinates. Similar derivations have been performed previously by various authors.^{8,9}

We begin with the free-field 3-D Helmholtz wave equation in cylindrical coordinates,

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial p}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 p}{\partial \varphi^2} + \frac{\partial^2 p}{\partial z^2} + k_0^2 n^2(r, z, \varphi) p = 0, \quad (1)$$

^{a)}Electronic mail: kevin@physics.nps.navy.mil

where $k_0 = 2\pi f/c_0$ is the reference acoustic wave number at frequency f , $n(r, z, \varphi) = c_0/c(r, z, \varphi)$ is the acoustic index of refraction in the medium with sound speed $c(r, z, \varphi)$, and c_0 is the typical ‘‘reference’’ sound speed. Depth z is defined positive downward from the free surface. The complex acoustic pressure P is assumed to be time harmonic at frequency f such that $p(r, z, \varphi)$ is defined by

$$P(r, z, \varphi, t) = p(r, z, \varphi) e^{-i2\pi ft}. \quad (2)$$

Note that it would then be a simple matter, in principle, to obtain fully 4-D (3-D space and time) acoustic predictions by Fourier synthesizing multiple frequency solutions to Eq. (1). The influence of density contrasts has also been neglected in this presentation, although vertical density variations are included in the model by defining an effective index of refraction,

$$n^2 \Rightarrow n^2 + \frac{1}{2k_0^2} \left[\frac{1}{\rho} \frac{\partial^2 \rho}{\partial z^2} \right]. \quad (3)$$

(Note: this is not the typical form used. However, it has been argued that this approximation is optimal.¹⁰ Further discussions on this will be presented in a forthcoming article by Tappert and Smith.)

Following Tappert,⁸ to obtain the parabolic approximation to Eq. (1), we assume only outgoing solutions to the Helmholtz equation [Eq. (1)], make the substitution

$$p(r, z, \varphi) = \frac{1}{\sqrt{r}} Q_{\text{op}}^{-1/2} u(r, z, \varphi), \quad (4)$$

and introduce the well-known PE ‘‘square-root operator,’’

$$Q_{\text{op}} = (1 + \mu + \epsilon + \nu)^{1/2}, \quad (5)$$

where

$$\mu = \frac{1}{k_0^2} \frac{\partial^2}{\partial z^2}, \quad \epsilon = n^2 - 1, \quad \text{and} \quad \nu = \frac{1}{k_0^2 r^2} \frac{\partial^2}{\partial \varphi^2}, \quad (6)$$

such that the outgoing solution to the Helmholtz equation approximately satisfies an equation of parabolic form,

$$-ik_0^{-1} \frac{\partial u}{\partial r} = Q_{\text{op}} u. \quad (7)$$

In the absence of range dependence (which implies, therefore, no azimuthal dependence), this expression is exact.

Because the Q_{op} operator is only a pseudodifferential operator and cannot be explicitly applied directly, some form of algebraic approximation is required. There exists a vast literature of work on approximations to the corresponding 2-D operator,⁸⁻¹¹ which invokes the uncoupled azimuth (UNCA) approximation, i.e.,

$$Q_{\text{UNCA}} = (1 + \mu + \epsilon)^{1/2}. \quad (8)$$

Three-dimensional environments may still be treated in an $N \times 2$ -D fashion, where the UNCA approximation is utilized to predict the propagation along individual bearings, each of which is defined by a unique index of refraction, $n(r, z, \varphi)$. It is worth noting that some approximations to Eq. (8) are more amenable to specific numerical algorithms than others. In this paper, only solutions based on the highly efficient and

stable split-step Fourier (SSF) technique¹² will be examined. Therefore, it is necessary to approximate Eq. (8) in such a way that the index of refraction term and the depth derivative term are separated.

For similar reasons, it is also desirable to solve the 3-D PE using a 2-D SSF technique. The approximation to Eq. (5) should then also separate the azimuthal derivative term from the others. Since azimuthal coupling may be assumed to be small (especially in light of the success of $N \times 2$ -D models), a simple binomial expansion should be adequate for most environments, thus

$$Q_{\text{op}} \approx (1 + \mu + \epsilon)^{1/2} + \frac{1}{2}\nu = Q_{\text{UNCA}} + \frac{1}{2}\nu. \quad (9)$$

This approximation separates the operator responsible for azimuthal coupling from the typical 2-D UNCA operator defined in Eq. (8). For this work, the ‘‘wide-angle’’ PE (WAPE) approximation¹³ to Eq. (8) is utilized, such that

$$Q_{\text{op}} \approx 1 - (T_{\text{op}} + U_{\text{op}} + V_{\text{op}}), \quad (10)$$

where

$$T_{\text{op}} = -\frac{1}{k_0^2} \frac{\partial^2}{\partial z^2} \left[\left(1 + \frac{1}{k_0^2} \frac{\partial^2}{\partial z^2} \right)^{1/2} + 1 \right]^{-1}, \quad (11)$$

$$U_{\text{op}} = -(n - 1), \quad (12)$$

and

$$V_{\text{op}} = -\frac{1}{2k_0^2 r^2} \frac{\partial^2}{\partial \varphi^2}. \quad (13)$$

In this form, both differential operators have been separated from the index of refraction term, as required for implementation with the SSF technique. It has been shown that the WAPE operators T_{op} and U_{op} do not, in general, decompose into the standard normal mode basis set associated with the Helmholtz wave equation.¹⁴ The recently published c_0 -insensitive version¹⁵ of the WAPE operators is expected to improve on this mismatch significantly, but has not been implemented in these calculations. However, this shortcoming of the WAPE approximation is not expected to affect the results presented here.

To obtain solutions via a forward ‘‘marching’’ algorithm, the customary PE field function ψ is introduced by giving it the usual envelope definition,

$$u(r, z, \varphi) = \psi(r, z, \varphi) e^{ik_0 r}. \quad (14)$$

Substitution into Eq. (7) then yields

$$\frac{\partial \psi}{\partial r} = -ik_0 \psi + ik_0 Q_{\text{op}} \psi \approx -ik_0 (T_{\text{op}} + U_{\text{op}} + V_{\text{op}}) \psi, \quad (15)$$

which can be solved by an approximate marching algorithm according to

$$\psi(r + \Delta r) \approx e^{-ik_0 \Delta r (T_{\text{op}} + U_{\text{op}} + V_{\text{op}})} \psi(r). \quad (16)$$

It can be shown that evaluation of the operators at the midpoint of the range step provides solutions that are second-order accurate in Δr .¹¹

Equation (16) could be implemented using a variety of numerical techniques. At this point, it is worth noting the primary differences between this numerical technique and

other techniques employed to solve similar 3-D propagation problems. Lee *et al.*⁵ have examined applications of implicit-finite difference (IFD), which require wrapping the field around on itself to ensure proper azimuthal boundary conditions. The 2-D SSF algorithm automatically ensures azimuthal continuity, even when limited apertures are applied with proper windowing. Effects of diffraction from the edges cannot be discounted, however, and caution must be observed when examining any but the central most radials. Collins⁶ has examined 3-D finite element (FE) approaches, including combinations of modes and PE solutions. Orris and Collins⁷ also considered a similar FFT technique for the 3-D PE solution to backscattering. However, they employed a line source rather than a point source and did not encounter the issues of zero range singularities described in the next section. It may be possible to include a similar azimuthal aperture limit on such models to improve their efficiency for specific 3-D environmental problems. In general, however, split-step Fourier techniques remain the most stable and efficient routines, and have the benefit of following the evolution of both physical and wave number space.

II. NUMERICAL IMPLEMENTATION OF FINITE AZIMUTHAL APERTURE ALGORITHM

As stated earlier, a 2-D (depth and azimuth) SSF technique will be used due to its efficiency and stability. This is achieved by defining the two-dimensional Fourier transform,

$$\begin{aligned} \psi(r, z, \varphi) &= \iint \hat{\psi}(r, k_z, k_0 s) e^{ik_z z} e^{ik_0 s \varphi} d(k_0 s) dk_z \\ &= F^{2D}[\hat{\psi}(r, k_z, k_0 s)], \end{aligned} \quad (17)$$

with transform variable pairs ($z \leftrightarrow k_z$) and ($\varphi \leftrightarrow k_0 s$). Numerical implementation is then accomplished according to

$$\begin{aligned} \psi(r + \Delta r, z, \varphi) &= e^{-ik_0 \Delta r U_{\text{op}}(r, z, \varphi)} \\ &\quad \times F^{2D}\{e^{-ik_0 \Delta r [\hat{T}_{\text{op}}(k_z) + \hat{V}_{\text{op}}(r, s)]} \\ &\quad \times [F^{-2D}(\psi(r, z, \varphi))]\}, \end{aligned} \quad (18)$$

where the differential operators in their corresponding transform domain become scalar operators, now defined explicitly as

$$\hat{T}_{\text{op}}(k_z) = 1 - \left[1 - \left(\frac{k_z}{k_0} \right)^2 \right]^{1/2} \quad (19)$$

and

$$\hat{V}_{\text{op}}(r, s) = \frac{s^2}{2r^2}. \quad (20)$$

Note that $k_0 s$ is similar to a bearing deviation wave number (but not exactly!). For propagation along a bearing with no azimuthal coupling, $k_0 s = 0$. Furthermore, one can easily recover the corresponding $N \times 2$ -D approximation by setting $s = 0$ in Eq. (20). This provides a nice confirmation of the algorithm. It is important, however, to recognize that the actual bearing deviation wave number¹⁶ is $k_0 s / r$, as suggested by the azimuthal wave number domain operator V_{op} .

This important distinction will have a significant influence in what follows.

Note also that a similar wide-angle approximation could be made to the square-root operator, Eq. (5), which combines the depth and azimuthal derivative terms into a single, wide-angle operator. Such an operator would still become a scalar multiplier in the double-transformed wave number space. However, it would also be an $M \times N$ scalar matrix, rather than two scalar vectors of length M and N . The increased computational memory necessary to handle such an operator does not seem worth what is likely to be only a minor improvement in the solution, and will therefore not be considered further.

Equation (18) together with Eqs. (12), (19), and (20) provide the fundamental formulas needed to numerically predict acoustic propagation in 3-D. There also remains the issue of numerically “gridding” the space in order to sample both the ocean environment and the acoustic field adequately. In terms of the latter, we can simply examine the relationships between the gridding in the physical and transformed domains. Specifically, if we choose to sample the space by M points in depth and N points in azimuth, then depth and bearing resolution are defined by

$$\Delta z = \frac{z_T}{M} \quad (21)$$

and

$$\Delta \varphi = \frac{\varphi_T}{N} \quad (22)$$

where z_T and φ_T are the total depth and azimuthal extents, respectively. By employing discrete Fourier transforms in Eq. (18), the corresponding resolutions in the wave number domains are

$$\Delta k_z = \frac{2\pi}{z_T} = \frac{2\pi}{M \Delta z} \quad (23)$$

and

$$k_0 \Delta s = \frac{2\pi}{\varphi_T} = \frac{2\pi}{N \Delta \varphi}. \quad (24)$$

According to the Nyquist criterion, the maximum wave numbers sampled in each direction are

$$|k_z|_{\text{max}} = \frac{M}{2} \Delta k_z = \frac{\pi}{\Delta z} \quad (25)$$

and

$$|k_0 s|_{\text{max}} = \frac{N}{2} k_0 \Delta s = \frac{\pi}{\Delta \varphi}. \quad (26)$$

Let us pause for a moment and reflect on the subtle, but important, characteristics of what may seem to be obvious. Typically, an acoustic propagation model will define some maximum depth extent, z_T , sampled with a fixed number of mesh points M . This produces a fixed grid size Δz . According to Eq. (25), this fixes $|k_z|_{\text{max}}$ and, therefore, the resolution of vertical structure of propagation that can be treated numerically. Similarly, if a fixed number of mesh points N is

defined over an azimuthal extent of φ_T (typically defined as 2π), then this fixes $|k_{0s}|_{\max}$. However, in contrast to the fixed vertical sampling, the resolution of bearing deviation angle that can be computed will drop off like r^{-1} ! This would be analogous to increasing the depth mesh size linearly in the range; as the range increases, the ability to resolve the vertical structure becomes degraded.

As is often the case, there are additional tricks that must be employed to produce accurate results in an efficient manner. Up to this point, no mention has been made of the boundary conditions that must be applied to the pressure field. In the 2-D problem, there are two boundary conditions: the field must go to zero at the surface and at large depths. In terms of the PE field function, these are defined as

$$\psi(z=0)=0 \quad (27)$$

and

$$\psi(z \rightarrow \infty)=0. \quad (28)$$

The first boundary condition is easily satisfied using the SSF technique by introducing an “imaginary” ocean of identical structure for negative depths and requiring

$$\psi(-z)=-\psi(z). \quad (29)$$

The second boundary condition can be treated effectively by introducing an absorbing layer, or “sponge,” at deeper depths in the calculation. The depths at which this sponge is applied are rather arbitrary, but should be deep enough such that one would not expect acoustic energy to return.

In the 3-D problem, we simply need to add a continuity condition in azimuth, i.e.,

$$\psi(\varphi+2\pi)=\psi(\varphi). \quad (30)$$

If the total azimuthal aperture is 2π , this boundary condition is easily satisfied by enforcing Eq. (30) at the initial range. The symmetry of the SSF algorithm then ensures that this boundary condition is met throughout the calculation. However, as has been argued, the use of the full 2π aperture limits the range of accuracy of the model.

Suppose for a moment that the solution has already been computed out to some range. Furthermore, assume that something has absorbed all (or most) of the energy over half of the total azimuthal aperture. A discrete Fourier transform in bearing space is now twice as large as necessary to obtain the same information since half of the signal appears as zero padding. The effective azimuthal aperture is now only π . In the transformed domain, the sampling is still defined by $k_0 \Delta s = 1$ with a maximum value of $|k_{0s}|_{\max} = N/2$. However, this sampling rate is twice as small as the actual information content. Therefore, every other point in the transformed domain may be dropped, and a subsequent transformation back to bearing space would provide the same information over the limited aperture π (of only half the size, $N/2$). It is important to note that the boundary condition, Eq. (30), is still being satisfied. We are simply choosing to not carry around the extra zeros (which could always be added back by zero padding at any point).

Of course, the same effective result can be achieved by filtering that part of the aperture we wish to exclude. By

zeroing out that part of the aperture, we remove the need to carry this part of the signal (since it contains no information), and we can reduce the transform size by the appropriate amount. Because of the way Fourier transforms react to such filtering (i.e., sidelobes of the window transform), a smooth filter should be designed to reduce the influence of sidelobe leakage and numerical reflections from the boundaries. Naturally, the influence of azimuthal coupling becomes questionable, at best, near the edges where the filtering occurs. Near the center of the unfiltered portion, however, this should not be an issue for most realistic environments.

The technique just described provides a means of reducing the computational load as range increases, but that was not the only significant point of this paper. On the contrary, a considerable problem is the reduced cross-range resolution as range increases. From Eq. (24), it was shown that the maximum “wave number” parameter $|k_{0s}|_{\max}$ remains constant for fixed azimuthal sampling $\Delta\varphi$. This provides the measure of the resolution of the bearing deviation angle. However, the actual maximum azimuthal wave number drops off like r^{-1} . To increase the resolution of the bearing deviation angle that can be computed, a larger $|k_{0s}|_{\max}$ is needed. This can also be easily accomplished in the azimuthal wave number domain. Consider again the case where half the total aperture has been filtered and we have performed a discrete Fourier transform to the azimuthal wave number domain. And again, every other point is dropped, thereby effectively doubling $k_0 \Delta s$. But now this data is zero padded to twice its length, which brings the total size *back* to N . The subsequent transformation back to the bearing space occurs with the same number of points, but only half the original aperture (which was achieved before by dropping every other point in the transform domain) *and* at half the azimuthal sampling. We have simultaneously doubled the cross-range resolution of the environment and the resolution of the bearing deviation angle, and (effectively) maintained the necessary boundary condition with no increase in computational load as the range increases.

Obviously, there is something lost in the process (apart from numerical noise). As a result of increasing the resolution of the bearing deviation angle, the maximum allowable bearing deviation angle that can be computed has also been decreased. However, in the case of azimuthal coupling in real underwater acoustics problems, it is considered more important to maintain adequate sampling of the bearing deviation angle at long range (which will probably always be small) rather than allow for large out-of-plane angles.

There still remains a subtle, but very important, issue that must be addressed. The previous argument for using such a limited azimuthal aperture may suggest that we simply limit the calculation from the beginning, e.g., only use $\pi/2$ total aperture with smooth filtering away from the central radial for *all* ranges. (In fact, this author had incorrectly used this approach in earlier versions of the code. Oddly enough, some of the results were not too bad!) The error of this method can be described by again using the analogy with a vertical array aperture. For a fixed number of elements, their separation, and hence total aperture, grows linearly with range. Similarly, as one considers the near-field

situation, the element separation and total aperture gets very small. Small element separation is not a problem, as it simply increases the sampling of the field. However, if the total aperture gets too small, the “beams” become so “fat” that all directional information is lost, i.e., the entire array acts like a point receiver.

For the corresponding azimuthal aperture, the situation is a little different. Specifically, the total “length” of the aperture (even when covering 2π of azimuth) must shrink to zero at range zero. For the first few range steps, the total azimuthal length may only be on the order of a wavelength. Therefore, it is essential to carry *all* of this information during the calculation with the Fourier transforms. Not until the azimuthal length is many wavelengths, or, equivalently, until the azimuthal sampling is on the order of a wavelength, are we able to justify applying a filter and reducing the total azimuthal aperture.

Therefore, combining all of the above, a limited azimuthal aperture algorithm may be designed as follows. First, the full 2π azimuthal coverage must be defined. This can be accomplished by simply defining the 2-D starting field as in previous $N \times 2$ -D models over N radials. [Note that this automatically satisfies the boundary condition defined by Eq. (30).] The solution is marched out in range according to the 2-D SSF algorithm, Eq. (18). At each range step, the cross-range sampling size, $r \Delta\varphi$, is compared to the acoustic wavelength. If $r \Delta\varphi > \lambda/2$, a smooth filter is applied to the outer quarters of the azimuth, i.e., from $\pi/2$ to π and $-\pi/2$ to $-\pi$ relative to the central radial. (Note that the filter chosen should not reduce the field to zero immediately to avoid numerical reflections. For this application, a smooth filter with a minimum of about 75% was chosen. Multiple applications of this filter then effectively reduce the outer radials to zero after several range steps without generating significant side lobe leakage. This filter was chosen somewhat arbitrarily based on experience and could probably be improved.) When the range increases to the point where $r \Delta\varphi > \lambda$, the aforementioned technique for dropping that half of the aperture that is now effectively zero and resampling the field at half the original azimuthal spacing is applied. Some additional filtering of the outer resampled radials is also necessary at this point to avoid aliasing that may occur as the solution marches forward from this point. However, since the first criteria ($r \Delta\varphi > \lambda/2$) is immediately satisfied after resampling, the first smooth filter is again applied until the cross-range sampling meets the second criteria ($r \Delta\varphi > \lambda$) when another resampling of the field can be accomplished. At any given range of the calculation, roughly half (the “inner” half) of the radial solutions may then be considered accurate.

Note that the specific cross-range criteria stated are not based on anything more than the requirement that the total azimuthal aperture “length” is many wavelengths (assuming N is not too small). Other criteria could be used if it was decided that wavelength sampling cross-range was not needed. Furthermore, one should be cautious not to believe that such a technique could be applied indefinitely. Experience has shown that two or three resamplings (reducing the total aperture by a factor of 4 or 8) is about the current limit,

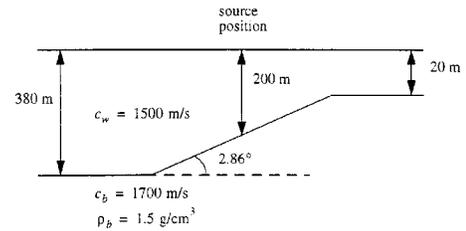


FIG. 1. Schematic diagram of 3-D wedge environment cross section.

although this may be increased by careful application of the resampling for specific problems. As in many numerical applications, the careful user must simply work towards both convergent and physically realistic solutions.

The technique outlined above has been implemented into a 3-D version of the formerly $N \times 2$ -D version of the Monterey–Miami Parabolic Equation (MMPE) model.¹⁷ Results have been computed for various synthetic and real environmental datasets. In the next section, a single, synthetic environment is used to test the model. Results are compared with other work, which provide a good benchmark dataset.

III. NUMERICAL RESULTS

In this section, the limited aperture 3-D PE model just defined will be employed to compute the propagation in the 3-D wedge environment defined in Ref. 18. This environment is a 3-D extension of the 2-D ASA benchmark wedge.¹⁹ In Ref. 18, a combination of techniques was used to solve a similar 3-D parabolic equation. Specifically, an implicit finite difference (IFD) scheme²⁰ was used to solve the range/depth problem while a split-step Fourier technique was used to solve the range/azimuth problem. In that work, it was shown that an adequate number of radials must be used to compute the propagation or the proper azimuthal coupling is not achieved. Furthermore, if too few radials are used, the solution appears to agree very well with the $N \times 2$ -D results.

Following the example of Ref. 18, a 3-D wedge is defined with a cross-sectional profile, as depicted in Fig. 1. The source is at 100-m depth at a point along the slope with a 200-m deep water column. The slope is fixed at 2.86° and the depth varies from a maximum of 380 m to a minimum of 20 m. The water column is isospeed with sound speed $c_w = 1500$ m/s, and the bottom is a homogeneous half-space with sound speed $c_b = 1700$ m/s, density $\rho_b = 1.5$ g/cm³, and no attenuation.

Azimuthal coupling in this environment is most easily observed by exciting individual modes at the source location and propagating them outward in range. For a source frequency of $f = 25$ Hz, the following function is used to define the first propagating mode at the source location:

$$\psi(z) = \sin\left(z \sqrt{\frac{\omega^2}{c_w^2} - K_1^2}\right); \quad z \leq H, \quad (31a)$$

$$\psi(z) = e^{-\sqrt{K_1^2 - \omega^2/c_b^2}(z-H)} \sin\left(H \sqrt{\frac{\omega^2}{c_w^2} - K_1^2}\right); \quad z > H, \quad (31b)$$

where $H = 200$ m is the water depth at the source location, $\omega = 2\pi f$ is the angular frequency of the source, and K_1

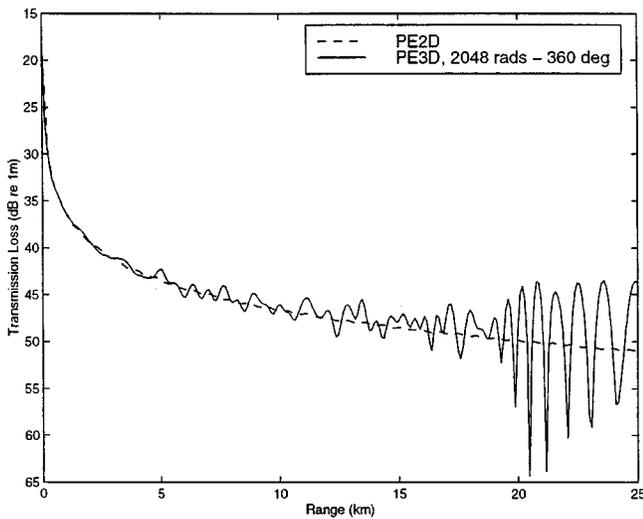


FIG. 2. Mode 1 transmission loss computed at 36-m depth along a 200-m isobath in the cross-slope direction. The dashed curve is a 2-D calculation, the solid curve is a 3-D calculation using 2048 radials over full 360° azimuthal aperture.

$=0.10378 \text{ m}^{-1}$ is the horizontal wave number for mode 1 at the source location, as computed by a standard normal mode algorithm.¹¹

The strongest influence of azimuthal coupling is anticipated along the 200-m isobath radial from the source (in either direction perpendicular to the slope). Along one such radial, transmission loss calculations were extracted at a depth of 36 m. Figure 2 shows the comparison between the 2-D calculation and a 3-D calculation of the mode 1 transmission loss along one of these radials. This 3-D calculation was performed with 2048 radials covering the full 360° azimuthal aperture. The 2-D TL curve exhibits simple cylindrical spreading, as expected. The 3-D TL curve shows the influence of azimuthal coupling beyond ~20 km. As will be shown later, this interference pattern is due to refracted mode 1 paths originally propagating upslope interfering with downslope propagating mode 1 paths. It should be noted that beyond the range computed, the mode 1 TL increases dramatically due to a mode 1 shadow zone. Furthermore, it is important to note that a range-averaged evaluation of the 3-D results will produce a curve very similar to the 2-D results. Thus, the 2-D, uncoupled azimuth approximation still provides good mean level predictions prior to the mode 1 shadow zone.

A similar plot is shown in Fig. 3, which compares the results of the 3-D calculation with 2048 radials covering the full 360° azimuthal aperture and the 3-D calculation with only 512 radials covering 90° of azimuthal aperture. The azimuthal sampling of the two calculations is the same, therefore, and the results are found to agree extremely well. Thus, the limited azimuthal aperture algorithm appears to correctly predict the effects of azimuthal coupling. Of practical significance here is the fact that, due to the smaller azimuthal transform size needed throughout the calculation, the limited azimuthal aperture results were generated in nearly one fourth of the time of the full 360° azimuthal ap-

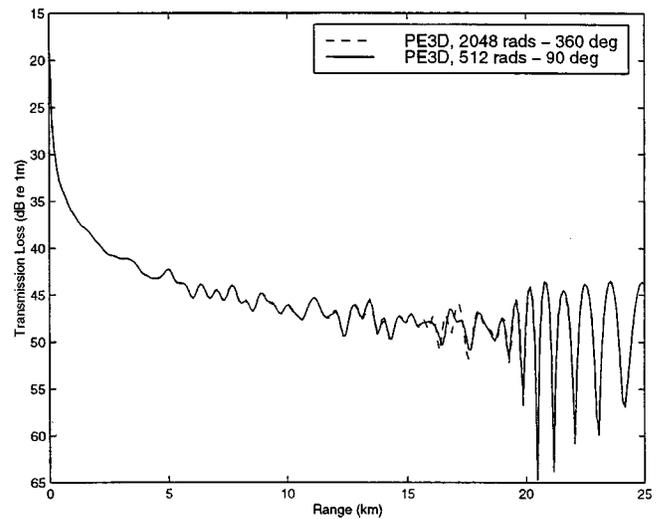


FIG. 3. Mode 1 transmission loss computed at 36-m depth along a 200-m isobath in the cross-slope direction. The dashed curve is a 3-D calculation with 2048 radials over full 360° azimuthal aperture (as in Fig. 2); the solid curve is a 3-D calculation with 512 radials over 90° azimuthal aperture (azimuthal sampling is the same).

erture calculation. The limited azimuthal aperture approach is therefore accurate and efficient.

In Fig. 4, the results of the 3-D calculation with 512 radials covering the full 360° is compared to the results of the limited aperture 3-D calculation with 512 radials covering 90°, a factor of 4 increase in azimuthal sampling. Note that the azimuthal resolution of the full 360° calculation is not adequate to correctly predict the influence of azimuthal coupling. Thus, the limited azimuthal aperture calculation, which has a nearly identical run time, provides much better resolution and correctly predicts the azimuthal coupling.

An initial comparison seemed to indicate a very good match between these results and Fawcett's previous work. The range where the interference pattern began seemed consistent as did the magnitude of the fluctuations. After comparing more recent digital data provided by Fawcett,²¹ some variability exists between the solutions. However, such variability appears to exceed the accuracy of the model with respect to reference sound speed, c_0 , sensitivity. An example of Fawcett's solutions for two different values of c_0 is given in Fig. 5. Similar variability is seen for three different values using the limited azimuthal aperture technique described here, as shown in Fig. 6. While Fawcett's technique seems less sensitive at shorter ranges, presumably due to the IFD implementation in the vertical plane, the location of the onset of the interference pattern appears more sensitive, which is probably due to the use of the WAPE approximation in the 3-D MMPE implementation.

The complete transmission loss data at 36-m depth is displayed in Fig. 7 for the $N \times 2$ -D results, 3-D results with full aperture (2048 radials over 360°) and the limited aperture 3-D results (512 radials over 90°). Note the influence of the filter applied to the outer quarters of this final dataset. The effects of azimuthal coupling are apparent using the limited aperture technique. In this plot, the aforementioned mode 1 interference is evident. Those paths originally propa-

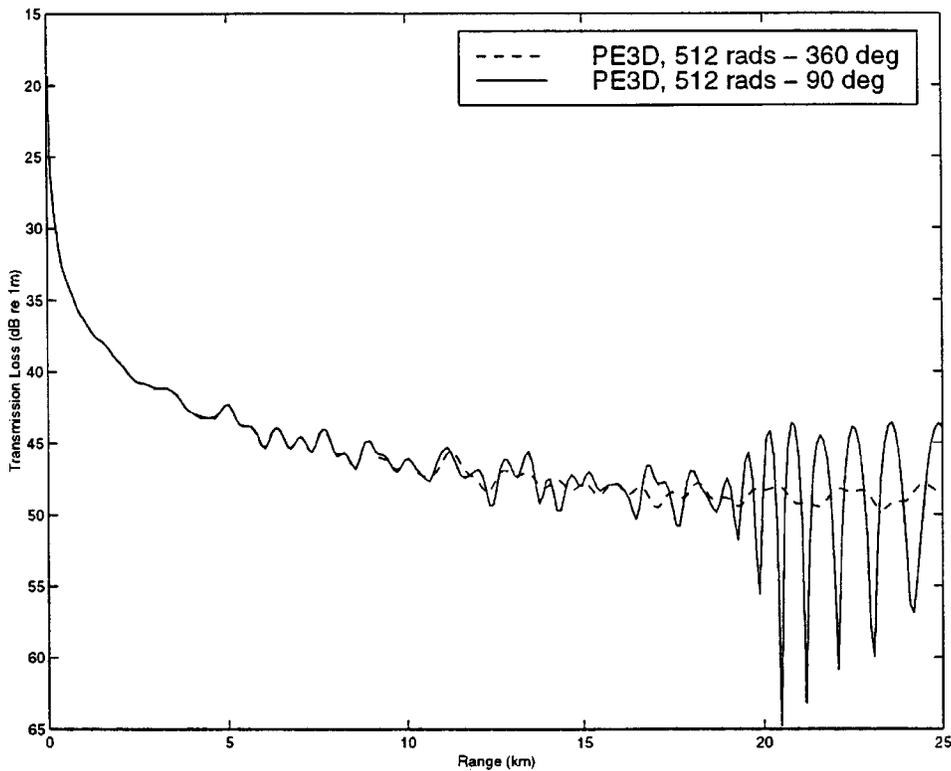


FIG. 4. Mode 1 transmission loss computed at 36-m depth along a 200-m isobath in the cross-slope direction. The dashed curve is a 3-D calculation with 512 radials over full 360° azimuthal aperture (as in Fig. 2); the solid curve is a 3-D calculation with 512 radials over 90° azimuthal aperture (azimuthal sampling increased by factor of 4).

gating upslope are refracted downslope and interfere with other mode 1 paths.

It is the opinion of this author that this test case be used as a benchmark test case for other investigators working on full-wave 3-D propagation models. Obviously from the results presented here, a true benchmark solution is not yet available. Other variations of this environment or source function could also be agreed upon by the modeling commu-

nity in an attempt to validate the latest 3-D codes (e.g., different source frequencies, different bottom slopes, structured water column profiles, etc.).

IV. SUMMARY

In this paper, a novel approach for computing full-wave, 3-D acoustic propagation using a 2-D SSF/PE algorithm with

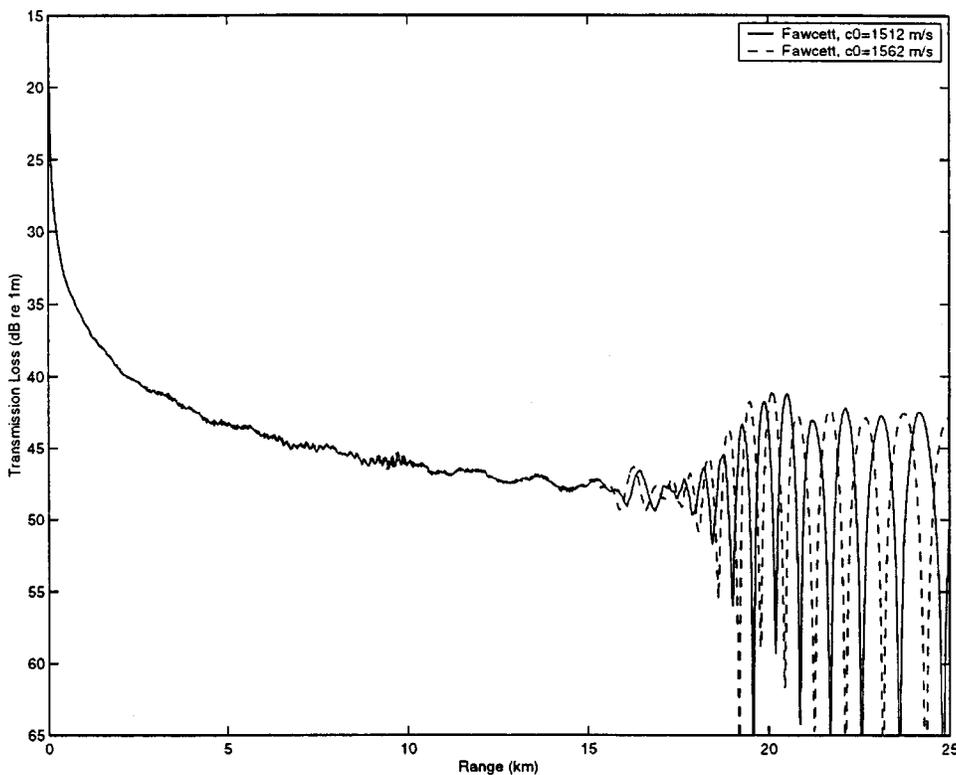


FIG. 5. Mode 1 transmission loss computed at 36-m depth along a 200-m isobath in the cross-slope direction (from Fawcett, Ref. 21). The solid curve is a 3-D calculation with reference sound speed $c_0=1512$ m/s (as in the original paper); the dashed curve is based on same model with reference sound speed $c_0=1562$ m/s.

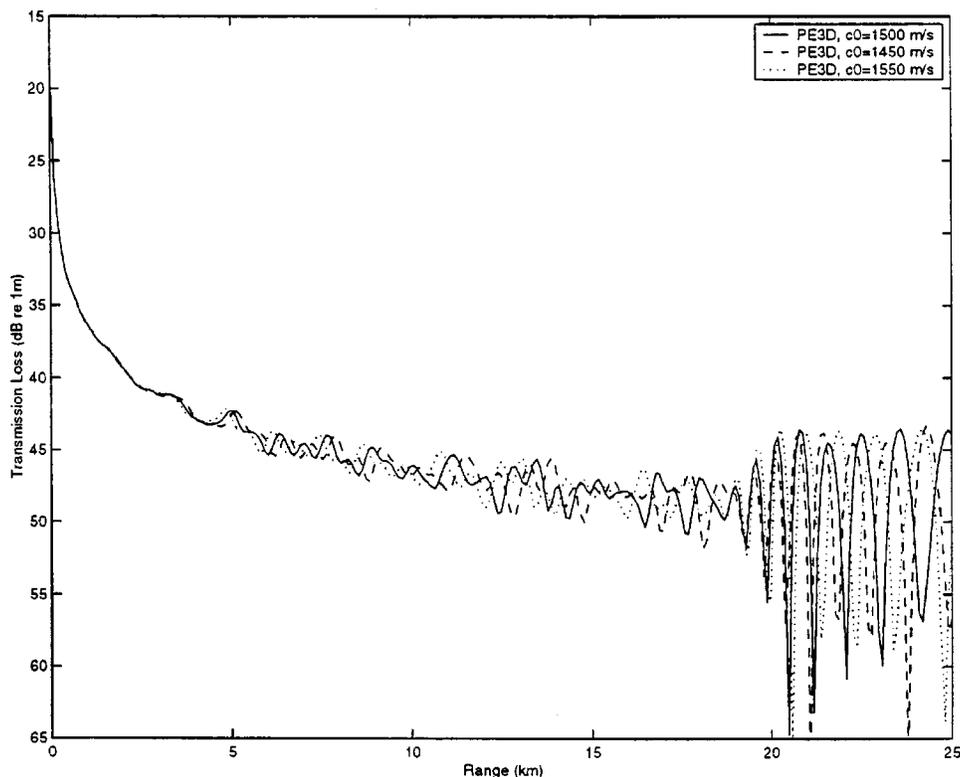


FIG. 6. Mode 1 transmission loss computed at 36-m depth along a 200-m isobath in the cross-slope direction using the limited azimuthal aperture technique. Different solutions found with reference sound speeds $c_0 = 1500, 1450,$ and 1550 m/s are displayed.

limited azimuthal aperture was introduced. The benefit of limiting the size of the aperture is to increase the cross-range resolution without sacrificing efficiency. This is achieved by using smooth filters to remove the outer quarters of the field bearings and then resampling the field. Both of these operations are easily performed in the azimuthal wave number domain. It was noted, however, that such an approach is not valid at very short ranges, and so a scheme must be used to ensure adequate sampling before this technique may be applied. Once this is achieved, it was found that the innermost bearings produce effectively the exact same results as those without filtering. In environments with extreme azimuthal variability, this method would presumably break down at long range, so the limits of its usefulness are unclear. In most ocean acoustics problems, however, it is expected to work well out to significant ranges.

It was also confirmed that adequate azimuthal sampling is a requirement for accurate 3-D propagation effects. Furthermore, it was argued that total aperture sampling is required at short ranges to ensure enough complete wavelengths in the solution. It may be that such a requirement is related to numerical sampling of Huygens' wave fronts or some diffractive phenomenon.

The specific test case of the mode 1 propagation in the 3-D penetrable wedge was examined. The results were found to agree quite well with similar previous work by others. Due to the relative simplicity of this environment and the additional confirmation of the results provided by this paper, it is recommended that this particular propagation problem be used by the acoustics modeling community as a benchmark in future 3-D propagation model development.

Improvements could be made to the current algorithm in the form of better filtering techniques. A wider angle azi-

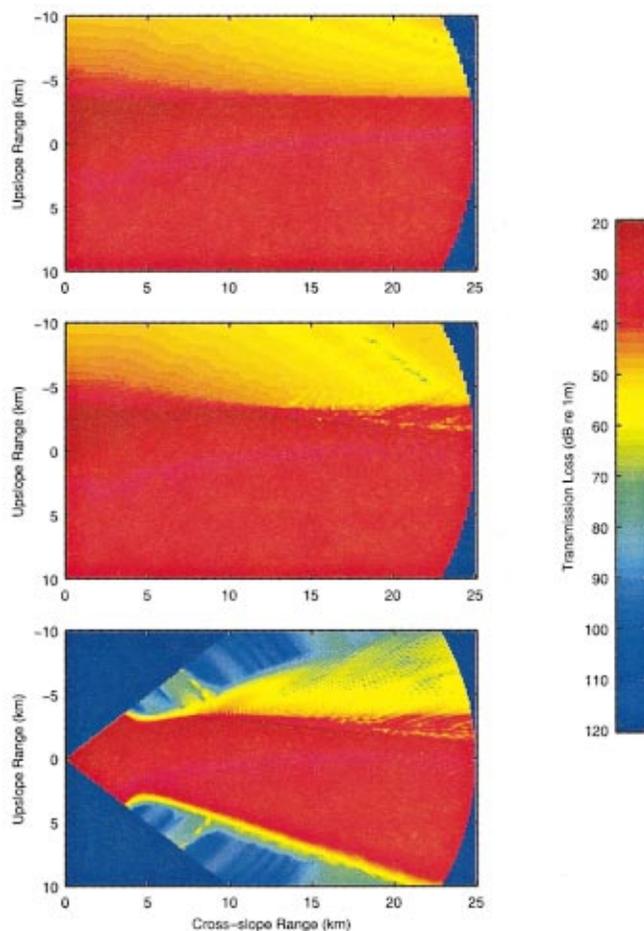


FIG. 7. Mode 1 transmission loss computed at 36-m depth for the wedge environment: (upper) $N \times 2$ -D calculation; (middle) 3-D calculation with 2048 radials over 360° ; and (lower) 3-D calculation with 512 radials over 90° .

muthal operator could also be defined, although it is not expected to improve the results in most ocean environments. The effects of Earth curvature may also be added to test the algorithm on basin scale propagation.

The introduction of a broadband version of the code has already been implemented and the results have been used to contrast the effects of 3-D environments on broadband pulse propagation.^{22,23} It was found that the $N \times 2$ -D solutions show remarkably 3-D characteristics, making the experimental observation of true 3-D azimuthal coupling extremely challenging. As others have noted,²⁴ it appears the best approach for observing azimuthal coupling in a true ocean environment is to measure mode TL at low frequencies (with only a few propagating modes) and search for a mode cutoff range.

As the speed of modern computers increases and the cost of memory decreases, it is natural to expect that fully 4-D (3-D space plus time) calculations covering the full 360° of azimuth will become common. The technique outlined in this paper will then become unnecessary. However, for the foreseeable future, the ability to accurately predict 3-D propagation in shorter time with less memory using the finite azimuthal aperture technique appears to be a useful alternative.

ACKNOWLEDGMENT

This work was supported by ONR Code 3210A. The author would like to acknowledge the help and guidance of Professor Fred Tappert for many useful discussions.

¹K. D. Heaney, W. A. Kuperman, and B. E. McDonald, "Perth-Bermuda sound propagation (1960): Adiabatic mode interpretation," *J. Acoust. Soc. Am.* **90**, 2586–2594 (1991).

²E. K. Westwood, "Broadband modeling of the three-dimensional penetrable wedge," *J. Acoust. Soc. Am.* **92**, 2212–2222 (1992).

³K. B. Smith, "Three-dimensional effects on broadband pulse propagation near shelf breaks," *Proceedings of International Conference on Shallow Water Acoustics*, Beijing, China, 21–25 April 1997 (China Ocean Press, City, 1997), pp. 265–270.

⁴C.-F. Chen, J.-J. Lin, C.-W. Wang, and D. Lee, "The 3-D effect on underwater acoustic propagation in the offshore area of Taiwan's coast," *J. Acoust. Soc. Am.* **103**, 3028 (1998).

⁵D. Lee, G. Botseas, and W. L. Siegmann, "Examination of three-dimensional effects using a propagation model with azimuth-coupling capability (FOR3D)," *J. Acoust. Soc. Am.* **91**, 3192–3202 (1992).

⁶M. D. Collins, "The adiabatic mode parabolic equation," *J. Acoust. Soc. Am.* **94**, 2269–2278 (1993).

⁷G. J. Orris and M. D. Collins, "The spectral parabolic equation and three-dimensional backscattering," *J. Acoust. Soc. Am.* **96**, 1725–1731 (1994).

⁸F. D. Tappert, "The parabolic equation method," in *Wave Propagation and Underwater Acoustics*, edited by J. B. Keller and J. Papadakis (Springer-Verlag, New York, 1977), pp. 224–286.

⁹S. A. Chin-Bing, D. B. King, J. A. Davis, and R. B. Evans, in *PE Workshop II: Proceedings of the 2nd Parabolic Equation Workshop*, Naval Research Laboratory NRL/BE/7181-93-0001, U.S. Government Printing Office, 1993, pp. 49–55.

¹⁰K. B. Smith and F. D. Tappert, "UMPE: The University of Miami Parabolic Equation Model, Version 1.0," Marine Physical Laboratory Technical Memo 432, 1993.

¹¹F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (American Institute of Physics, New York, 1994), Chaps. 5–6, pp. 271–412.

¹²R. H. Hardin and F. D. Tappert, "Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations," *SIAM Rev.* **15**, 423 (1973).

¹³D. J. Thomson and N. R. Chapman, "A wide-angle split-step algorithm for the parabolic equation," *J. Acoust. Soc. Am.* **74**, 1848–1854 (1983).

¹⁴A. R. Smith and K. B. Smith, "Mode functions for the wide angle approximation to the parabolic equation," *J. Acoust. Soc. Am.* **103**, 814–821 (1998).

¹⁵F. D. Tappert, J. L. Spiesberger, and L. Boden, "New full-wave approximation for ocean acoustic travel time predictions," *J. Acoust. Soc. Am.* **97**, 2771–2782 (1995).

¹⁶F. D. Tappert and K. B. Smith, "Analysis of the uncoupled azimuth (UNCA) approximation," to be submitted to *J. Acoust. Soc. Am.*

¹⁷The Monterey-Miami PE Model is an updated version of the University of Miami PE Model (see Ref. 7) and is available via the World Wide Web at the Ocean Acoustics Library (<http://oalib.njit.edu/pe.html>).

¹⁸J. A. Fawcett, "Modeling three-dimensional propagation in an oceanic wedge using parabolic equation methods," *J. Acoust. Soc. Am.* **93**, 2627–2632 (1993).

¹⁹F. B. Jensen and C. M. Ferla, "Numerical solutions of range-dependent benchmark problems in ocean acoustics," *J. Acoust. Soc. Am.* **87**, 1499–1510 (1990).

²⁰D. J. Thomson, "Wide-angle parabolic equation solutions to two range-dependent benchmark problems," *J. Acoust. Soc. Am.* **87**, 1514–1520 (1990).

²¹J. A. Fawcett (private communication).

²²K. B. Smith, "Three-dimensional propagation effects: Modeling, observations, and suggested benchmark cases," *J. Acoust. Soc. Am.* **103**, 2989 (1998).

²³K. B. Smith and F. D. Tappert, "Horizontal refraction and the uncoupled azimuth approximation," *Proceedings of the 6th International Congress on Sound and Vibration*, Copenhagen, Denmark, 5–8 July 1999, pp. 759–766.

²⁴G. B. Deane, "The penetrable wedge as a three-dimensional benchmark," *J. Acoust. Soc. Am.* **103**, 2989 (1998).

Bottom reverberation in shallow water: Coherent properties as a function of bandwidth, waveguide characteristics, and scatterer distributions

Kevin LePage^{a)}

Saclant Undersea Research Centre, Vl. S. Bartolomeo 400, 19138 La Spezia, Italy

(Received 30 October 1998; revised 17 August 1999; accepted 3 September 1999)

Shallow water presents a difficult, reverberation-limited environment for active SONAR operations. It is important to understand the predictable structure of shallow-water reverberation in order to aid the design of processors and detectors which work properly in these environments. In this paper, the temporal characteristics of monostatic reverberation are predicted as a function of source bandwidth, source–receiver depth, and the propagation characteristics of range-independent shallow water. Results show that at early time, reverberation can be highly coherent across a vertical line array, violating the homogeneous noise assumption, while at late time the reverberation becomes increasingly uncorrelated. This is shown to be due to the insonification of independent bottom patches at late time. It is also shown that this decorrelation of the reverberation is dependent both on the propagation characteristics of the particular shallow-water environment, the correlation length scale of the scatterers, and the bandwidth of the source, with high-bandwidth sources causing decorrelated reverberation sooner than low-bandwidth sources. The results also show that there are several identifying characteristics in reverberation time series which may be useful for identifying the types of scatterers which cause reverberation during particular experiments.

[S0001-4966(99)06212-8]

PACS numbers: 43.30.Gv, 43.30.Re, 43.30.Hw [DLB]

INTRODUCTION

Scattering physics is often modeled using the Helmholtz equation for computational reasons. Propagation loss to and from scattering regions is often treated in a like manner. The usual approach for generating reverberation time series therefore requires a numerical Fourier synthesis which can obscure the physics of reverberation phenomenology. Here, we model the reverberation process approximately using normal modes and perform the Fourier synthesis explicitly under a narrow-band approximation. The modes are allowed to interact with the bottom individually, and scattering interactions between the modes at the bottom are retained. Therefore, the approach allows pathological propagation phenomena such as convergence zones to be explicitly modeled. The geometry is monostatic in the horizontal coordinates, although different depths of the source and the receiver may be accommodated. As a generalization, different source and receiver vertical apertures and receiver characteristics may also be accommodated, although they are not explicitly treated in this paper.

As a natural result of the broadband nature of the analysis, the interaction of the modes with each other at the bottom is seen to be governed by the bandwidth of the analysis and the dispersive properties of the waveguide. The result is that modes which interact coherently at bottom patches at early time may decorrelate at late times, in ways which are determined by the mean propagation physics of the waveguide, the bottom correlation length scale, and the reverberation analysis band. In the interest of keeping the approach as

general as possible, the bottom-scattering process is modeled as the interaction between a known modal scattering function and the scatterers which are characterized by the second moments of their distribution in amplitude and space. The bottom-scattering functions are generally related to the mode-shape functions and their derivatives at and in the bottom, which are sensitive to the detailed bottom structure and the incident grazing angles of each mode. They can also be simple parametrizations based on grazing angle. This approach makes it possible to model scattering functions which are based on physics, such as perturbation theory, or “standard” models, such as Lambert’s law, within the framework of a general reverberation model. The interaction between the scattering functions and the scattering potential is assumed to be weak, so that multiple scattering is negligible and a small imaginary part of the model eigenvalues is sufficient to account for propagation loss, and local, so that the field scattered at any particular part of the bottom is due only to the bottom properties and the incident field at that particular point. This latter approximation serves as a restriction on the ability of the model to accommodate large-scale scatterers such as facets or other pathological features, but is useful for modeling clutter, the intended objective.

This work is similar in various respects to other normal-mode reverberation treatments, with the principal difference being the emphasis on the coherent properties of the temporal evolution of the reverberation. The treatment of the problem by Ellis¹ is similar, for instance, most noticeably in the use of the reverberation time over the round-trip modal slowness to determine the patch range. The work presented here may be viewed as an extension to Ellis’s model both because the individual modal dispersion terms, which are important

^{a)}Electronic mail: lepage@saclantc.nato.int

at late time for properly modeling the incident field, have been included, and because the ensemble average of the reverberation intensity has been rigorously obtained directly from the expressions for the complex scattered field. A paper by Tang, which is under review,² properly includes the modal dispersion term and the formal average of the intensity, and thus is more similar to the work presented here. Although the derivations are similar, the work presented here is more complete regarding the exploration of the effects of bandwidth and consequently patch size in relation to correlation length scale, and coherent effects in general. A paper by Tracey and Schmidt³ offers a rigorous frequency-domain derivation of modal volume scattering using the MSP, including the effects of scattering on modal coherence, but the derivations are obtained for the Helmholtz equation, so that numerical integration is required to obtain reverberation time series. Consequently, no exploration of the temporal structure of the reverberation and related temporal coherence issues has been carried out. To summarize, this treatment differs from the other modal scattering treatments in that more emphasis is placed on the importance of the temporal characteristics of modal interaction at and in the seafloor, and consequently a better understanding has been developed of when the modal components of the backscattered field are or are not coherent at various times after the shot. The theory developed here also explores how the ratio of the bottom patch size to the correlation length scale of the scatterers affects the resonant scattering characteristics of the reverberation, and simplified expressions for the time-dependent intermodal coherence are obtained in the limiting cases of very small and very large patch size with respect to a correlation length scale.

I. THEORY

Under the far-field approximation, the total field “incident” from a point source to a scatterer on or in the bottom at range r and frequency ω is given approximately by

$$p_i(\omega, z_s, r) \approx \sqrt{\frac{2\pi}{r}} \sum_{m=1}^N \frac{\phi_m(z_s) \phi_m^-}{\sqrt{k_m}} e^{ik_m r}, \quad (1)$$

where k_m are the modal wave numbers, ϕ_m are the mode shape functions, ϕ_m^- defined below, are the scattering potential for the mode shape functions, and z_s is the source depth.

In Eq. (1), the scattering function ϕ_m^- is intended to represent the excitation of some sediment or sub-bottom scatterer by mode m . At the bottom, we parametrize the scattering amplitude by a local scatterer distribution η , a real variable representing the physical variability of the bottom or sub-bottom which is a function of range and azimuth. From this scatterer distribution, the propagation back to the receiver is given by the Green’s function

$$p_r(\omega, z_r, r) \approx \sqrt{\frac{2\pi}{r}} \sum_{n=1}^N \frac{\phi_n(z_r) \phi_n^+}{\sqrt{k_n}} e^{ik_n r}, \quad (2)$$

where ϕ_n^+ represents the ability of the scatterer to excite the waveguide in the backscattered direction. Taken together, the

backscattered field from scatterers at all ranges and azimuths at frequency ω is

$$p(\omega, z_s, z_r) = 2\pi \int_0^{2\pi} d\theta \int_0^\infty dr \sum_{m=1}^N \sum_{n=1}^N \phi_m(z_s) \phi_m^- \eta \times (r, \theta) \phi_n^+ \phi_n(z_r) \frac{e^{i(k_n+k_m)r}}{\sqrt{k_n k_m}}. \quad (3)$$

In reality, the separability of the scattering functions ϕ_m^- and ϕ_n^+ is not guaranteed. Thus, to add generality we replace the scalar product $\phi_m^- \phi_n^+$ with a general *scattering kernel* T_{mn} , which may be full rank. It is emphasized that under this definition of the scattering kernel, T_{mn} is not a power quantity relating incident to scattered intensities, but a transition matrix relating incident to scattered pressures. Under this definition, the backscattered time series is then given by the inverse Fourier transform

$$p(t, z_s, z_r) = \int_0^{2\pi} d\theta \int_0^\infty dr \eta(r, \theta) \int_{-\infty}^\infty d\omega A(\omega) \times \sum_{m=1}^N \sum_{n=1}^N \phi_m(z_s) T_{mn} \phi_n(z_r) \frac{e^{-i(\omega t - (k_n+k_m)r)}}{\sqrt{k_n k_m}}, \quad (4)$$

where the mode-shape functions $\phi_{n,m}$, the scattering functions $\phi_{n,m}^\pm$, and the modal eigenvalues $k_{n,m}$ are all implicit functions of frequency. Although such an expression can be evaluated numerically, solutions in this form cannot provide significant insight into the controlling characteristics of reverberation. Instead, what is required is the short-time average (STA) of the square of the quantity $p(t)$, since we are interested in intensity; since we are interested in the expected intensity, we seek the ensemble average $\langle p_{\text{STA}}^2 \rangle$. In order to gain insight, we would also like to be able to evaluate the desired quantities analytically. Looking at Eq. (4), it seems that the two largest hurdles to closed-form evaluation are the generally unknown spatial characteristics of the scatterer distribution η , and the frequency integrals over the model Green’s functions p_i and p_r . The first difficulty can be overcome if we assume that the spatial statistics of the scatterers are homogeneous and known. The second difficulty may be overcome for signals of limited bandwidth by making use of the narrow-band approximation.

In the narrow-band approximation, temporal responses of filters are approximated by expanding the filter response about a center frequency. In this application, the filters of interest are the Green’s functions in Eqs. (1) and (2). The mode-shape function and the scattering function T_{mn} are both assumed to remain unchanged in the vicinity of the center frequency; the wave numbers $k_{n,m}$ are assumed to change following the second-order Taylor series expansion about the center frequency

$$k_n = k_n^o + (\omega - \omega_o) \left. \frac{\partial k_n}{\partial \omega} \right|_{\omega=\omega_o} + \frac{(\omega - \omega_o)^2}{2} \left. \frac{\partial^2 k_n}{\partial \omega^2} \right|_{\omega=\omega_o}. \quad (5)$$

Insertion of Eq. (5) into Eq. (4) and integrating over a bandwidth of $\Delta\omega$ yields the approximation

$$p(t, z_s, z_r) = 2A(\omega_o) \Re \left\{ \int_0^{2\pi} d\theta \int_0^\infty dr \eta(r, \theta) \right. \\ \times \sum_{m=1}^N \sum_{n=1}^N \phi_m^o(z_s) T_{mn}^o \phi_n^o(z_r) \frac{e^{-i(\omega_o t - (k_n^o + k_m^o)r)}}{\sqrt{k_n^o k_m^o}} \\ \left. \times \int_{-\Delta\omega}^{\Delta\omega} d\omega e^{-i\omega(t - (S_{nm} + (\omega/2)D_{nm})r)} \right\}, \quad (6)$$

where S_{nm} is the sum of the modal slownesses $\partial k_n / \partial \omega + \partial k_m / \partial \omega$, and D_{nm} is the sum of the modal dispersion terms $\partial^2 k_n / \partial \omega^2 + \partial^2 k_m / \partial \omega^2$ at the frequency ω_o . If we assume for the moment that the modal dispersion is negligible, then the frequency integral in Eq. (6) may be performed with the result

$$p(t, z_s, z_r) = 4\Delta\omega A(\omega_o) \Re \left\{ \int_0^{2\pi} d\theta \int_0^\infty dr \eta(r, \theta) \right. \\ \times \sum_{m=1}^N \sum_{n=1}^N \frac{e^{-i(\omega_o t - (k_n^o + k_m^o)r)}}{\sqrt{k_n^o k_m^o}} \phi_m^o(z_s) \\ \left. \times T_{mn}^o \phi_n^o(z_r) \text{sinc}(\Delta\omega(t - S_{nm}r)) \right\}, \quad (7)$$

where $\text{sinc}(x) \equiv \sin(x)/x$.

Several intuitive observations may be made upon inspection of Eq. (7). The first is that the amplitude of the reverberation time series increases linearly with the bandwidth $\Delta\omega$, consistent with the fact that the energy in the time series is increasing as more frequencies are included. The second is that the contribution to the scattered field at time t comes from a region concentrated about $r = t/S_{nm}$, i.e., from ranges close to the time over the round-trip modal slowness. Furthermore, we see that the response at time t is a local Fourier transform of the scattering process η , which is windowed by a sinc function whose spatial extent is approximately $\Delta r \approx 2\pi/\Delta\omega S_{nm}$ or approximately $c/2\Delta f$, the classic ‘‘patch size’’ from reverberation analysis.

At this point, it is useful to consider the different behavior of the reverberation in the limits of narrow- and broadband frequency excitation. First, there are the effects which exist in general for any particular mode pair and are not related to the coherent nature of the propagation which can introduce structure when Eq. (7) is summing coherently. If the spatial window is very narrow in extent (i.e., broadband excitation), then the individual scatterers dominate the return at any given time and the spatial Fourier transform has inadequate aperture to allow resonant scattering to occur. In addition, we can expect the variance of the backscattered time series to be on the order of the variance of the individual scattering cross sections of the scatterers themselves, since very few scatterers contribute to the total return from each mode pair at any given time. On the other hand, if the spatial window is very broad, as is the case for very narrow-band scattering processes, then the window is much larger than the

correlation length scale of the individual scatterers and resonant scattering can occur. In this case, we can expect much lower amplitudes of backscatter compared to the level of individual scattering cross sections in the window, and lower overall variance of the reverberation intensity, since many correlation lengths of the scatterers are insonified.

The second effect of bandwidth is how it controls the amplitude of the interference terms between the modes. Since the individual modes have different slownesses, we may expect coherent interference terms to disappear between modes where the difference between the range-integrated round trip slowness $\Delta t_{nm} = (S_{nn} - S_{mm})r$ is greater than one over the bandwidth, i.e., $\Delta t_{nm} > 2\pi/\Delta\omega$. Put another way, modes which interrogate different patches of the bottom more than a correlation length apart at a given time are decorrelated and will yield no interference structure associated with coherent propagation to the backscattered field. Thus, for broadband excitation we can expect significant variance in the reverberation intensity, but a fairly smooth mean; conversely, we expect highly structured mean reverberation intensity for narrow-band excitation but lower variance about this structure. We can therefore anticipate that reverberation deviation from smooth parametrized curves comes from a combination of effects induced by propagation (waveguide effects) and effects introduced by the variance of the scatterers themselves.

Even though we can understand how the various terms in Eq. (7) control the character of the reverberation as a function of time, evaluation of the windowed Fourier transform of the spatial distribution of scatterers is difficult to quantify in closed form. More importantly, the modal dispersion terms D_{nm} turn out to be very important to the accurate approximation of the incident field, even at moderate ranges.⁴ For these reasons we adopt a slightly different form of the narrow-band approximation where we assume that the bandwidth is controlled not by absolute frequency limits but by a Gaussian window function

$$A(\omega) = A e^{-(\omega - \omega_o)^2 / \Delta\omega^2},$$

which when inserted into Eq. (4) along with Eq. (5) yields a more useful and accurate form of the narrow-band approximation

$$p(t, z_s, z_r) \\ = 2\sqrt{\pi}A \Re \left\{ e^{-i\omega_o t} \sum_{m=1}^N \sum_{n=1}^N \frac{\phi_m^o(z_s) T_{mn}^o \phi_n^o(z_r)}{\sigma_{nm} \sqrt{k_n^o k_m^o}} \right. \\ \left. \times \int_0^{2\pi} d\theta \int_0^\infty dr \eta(r, \theta) e^{i(k_n^o + k_m^o)r} e^{-(t - S_{nm}r)^2 / 4\sigma_{nm}^2} \right\}, \quad (8)$$

where $\sigma_{nm}^2 \equiv \Delta\omega^{-2} - iD_{nm}r/2$. The term $\exp(-(t - S_{nm}r)^2 / 4\sigma_{nm}^2)$ in Eq. (8) serves the same purpose as the sinc function in Eq. (7), with the dispersion D_{nm} included explicitly. With this expression the windowed Fourier transform

will be easier to evaluate in subsequent manipulations. However, all of the intuition we have gained concerning the effects of bandwidth remains valid, except that the dispersion will have the effect of increasing the patch sizes at late times, both decreasing the variance of the scatterer contribution and increasing the importance of coherent propagation effects.

We now turn our attention to the short-time average (STA)⁵ of Eq. (8). First, we review the complex envelope theory. Equation (8) has the form

$$p(t) = \Re\{e^{-i\omega_o t} A(t)\},$$

where $A(t)$ is the so-called complex envelope which is everything in the bracket in Eq. (8) following the $e^{-i\omega_o t}$. Squaring $p(t)$ and expanding we get

$$\begin{aligned} p^2(t) &= \Re\{A(t)\}^2 (1/2 + \cos(2\omega_o t)/2) \\ &\quad + \Im\{A(t)\}^2 (1/2 - \cos(2\omega_o t)/2) \\ &\quad + 2\Re\{A(t)\}\cos(\omega_o t)\Im\{A(t)\}\sin(\omega_o t), \end{aligned}$$

which upon a short-time average, (where the short time is defined as a time sufficiently long to allow the twice-carrier frequency terms to integrate to zero and to allow the orthogonality of the sin and cos terms to be approached) yields

$$\begin{aligned} p_{\text{STA}}^2(t) &\approx \frac{\Re\{A(t)\}^2}{2} + \frac{\Im\{A(t)\}^2}{2} \\ &= \frac{|A(t)|^2}{2} = \frac{A(t)A^*(t)}{2}. \end{aligned}$$

Under the last identity, the STA of the backscattered intensity, i.e., the STA of the square of Eq. (8), is given by

$$\begin{aligned} p_{\text{STA}}^2(t, z_s, z_1, z_2) &= 2\pi A^2 \sum_{m=1}^N \sum_{n=1}^N \sum_{m'=1}^N \sum_{n'=1}^N \frac{\phi_m(z_s)\phi_{m'}(z_s)T_{mn}T_{m'n'}\phi_n(z_1)\phi_{n'}(z_2)}{\sigma_{nm}\sigma_{n'm'}^*\sqrt{k_n k_m k_n^* k_m^*}} \int_0^{2\pi} d\theta \int_0^\infty dr \eta(r, \theta) \\ &\quad \times e^{i(k_n+k_m)r} e^{-\tau_{nm}^2(r)/4\sigma_{nm}^2} \int_0^{2\pi} d\theta' \int_0^\infty dr' \eta(r', \theta') e^{-i(k_n^*+k_m^*)r'} e^{-\tau_{n'm'}^2(r')/4(\sigma_{n'm'}^*)^2}, \end{aligned} \quad (9)$$

where $\tau_{nm}(r)$ is the reduced time $t - S_{nm}r$, k^* indicates $\Re\{k\} - i\Im\{k\}$, and the superscripts o indicating that the scattering kernel, mode functions, and eigenvalues are evaluated at the center frequency have been dropped for compactness of notation. Note that Eq. (9) gives the STA of the reverberation intensity explicitly in terms of two receiver depths z_1 and z_2 . Thus, it is important to note that this development will yield the quantities necessary to predict the time-dependent vertical coherence

$$\rho^2(t, z_1, z_2) \equiv \frac{p_{\text{STA}}^2(t, z_1, z_2)}{\sqrt{p_{\text{STA}}^2(t, z_1, z_1)p_{\text{STA}}^2(t, z_2, z_2)}}. \quad (10)$$

Since sonar system operators in general have at most only general statistical characterizations of the bottom scattering parameters at their disposal, we will turn our attention to the evaluation of the ensemble average of Eq. (9). The only random variables in Eq. (9) are the scatterer distributions η . The ensemble average of these two distributions in Cartesian coordinates and under assumptions of spatial homogeneity is

$$\begin{aligned} \langle \eta(r, \theta)\eta(r', \theta') \rangle \\ \equiv R_\eta(r \cos \theta - r' \cos \theta', r \sin \theta - r' \sin \theta'). \end{aligned}$$

If we assume that the scatterer autocorrelation function R_η corresponds to an anisotropic Gaussian parametrization with correlation length scales l_x and l_y and a skew angle aligned with $\theta \equiv 0$,

$$R_\eta(x, y) = \frac{\langle \eta^2 \rangle}{2\pi l_x l_y} \exp(-(x^2/2l_x^2 + y^2/2l_y^2)), \quad (11)$$

and perform the change of variables $r'' = r' - r$ and $\theta'' = \theta' - \theta$, we have in the far field

$$\begin{aligned} x &\approx -r'' \cos \theta + r \theta'' \sin \theta, \\ y &\approx -r'' \sin \theta - r \theta'' \cos \theta. \end{aligned} \quad (12)$$

Insertion of Eqs. (12) into Eq. (11) and integration over θ'' yields (see Ref. 6 for details of the derivation)

$$\begin{aligned} \int_{-\theta}^{2\pi-\theta} d\theta'' R_\eta(r, r'', \theta, \theta'') \\ \approx \frac{\langle \eta^2 \rangle}{\pi r l_y \sqrt{2\pi(\sin^2 \theta + l_x^2/l_y^2 \cos^2 \theta)}} \\ \times \exp\left(\frac{r''^2}{l_x^2/2} \frac{(l_x^2/l_y^2 - 1)^2}{\sec^2 \theta + l_x^2/l_y^2 \csc^2 \theta}\right) \\ \times \exp\left(-r''^2 \left(\frac{\cos^2 \theta}{2l_x^2} + \frac{\sin^2 \theta}{2l_y^2}\right)\right), \end{aligned} \quad (13)$$

which delivers the commonly adopted r^{-1} term heretofore missing from the reverberation equations. The result shows that the geometric decay in the reverberation level is due solely to the azimuthal decorrelation of the scattering centers. This term can only be obtained explicitly upon seeking the ensemble average, and is not present in Eq. (9), which on first inspection might seem to indicate that the reverberation intensity does not decay with increasing time. Of course, realizations of the short-time average will show this decay for all times associated with reverberation coming from ranges significantly greater than a correlation length scale, as

the double azimuthal integral will introduce a linearly decreasing number of correlated components for increasing range.

In the limit of isotropic roughness with correlation length l , the ensemble average of the short-time average of the reverberation intensity reduces to

$$\begin{aligned} \langle p_{\text{STA}}^2(t, z_s, z_1, z_2) \rangle &= 2\pi^2 A^2 \sum_{m=1}^N \sum_{n=1}^N \sum_{m'=1}^N \sum_{n'=1}^N \frac{\langle \eta^2 \rangle \phi_m(z_s) \phi_{m'}(z_s) T_{mn} T_{m'n'} \phi_n(z_1) \phi_{n'}(z_2)}{rl \sqrt{2\pi} \sigma_{nm} \sigma_{n'm'} \sqrt{k_n k_m k_n^* k_m^*}} \\ &\times \int_0^\infty dr e^{i(k_n + k_m - (k_{n'} + k_{m'}))r} e^{-(t - S_{nm}r)^2 / 4\sigma_{nm}^2} \int_{-r}^\infty dr'' e^{-r''^2 / 2l^2} e^{-i(k_n^* + k_m^*)r''} e^{-(t - S_{n'm'}(r+r''))^2 / 4(\sigma_{n'm'}^*)^2}. \end{aligned} \quad (14)$$

If we approximate r^{-1} by S_{nm}/t and evaluate σ_{nm}^2 approximately as $\Delta\omega^{-2} - iD_{nm}t/2S_{nm}$ (see Appendix A for the restrictions on these assumptions) then all the remaining range dependence of the integrands resides in linear and quadratic terms in the arguments to the exponentials. If we take the further approximation of taking the lower limits on the range integrals as $-\infty$, the two range integrals may be evaluated exactly (see Ref. 6 for details). Under these approximations, the ensemble average of the short-time average of the reverberation intensity is

$$\begin{aligned} \langle p_{\text{STA}}^2(t, z_s, z_1, z_2) \rangle &= 2\pi^2 A^2 \sum_{m=1}^N \sum_{n=1}^N \sum_{m'=1}^N \sum_{n'=1}^N \frac{\langle \eta^2 \rangle S_{nm} \phi_m(z_s) \phi_{m'}(z_s) T_{mn} T_{m'n'} \phi_n(z_1) \phi_{n'}(z_2)}{tl \sqrt{2\pi} \sigma_{nm} \sigma_{n'm'} \sqrt{k_n k_m k_n^* k_m^*}} \\ &\times \exp \left\{ \left(-t \left(\frac{S_{n'm'}}{2\sigma_{n'm'}^2} R + \frac{S_{nm}}{2\sigma_{nm}^2} \right) - i(k_n + k_m - (k_{n'} + k_{m'}))R \right)^2 \left(\frac{S_{nm}^2}{\sigma_{nm}^2} + \frac{S_{n'm'}^2}{\sigma_{n'm'}^2} R \right)^{-1} \right\} \\ &\times \exp \left\{ - \left(t^2 \left(\frac{1}{4\sigma_{nm}^2} + \frac{R}{4\sigma_{n'm'}^2} \right) + it \frac{k_{n'} + k_{m'}}{S_{n'm'}} R + \frac{(k_{n'} + k_{m'})^2 \sigma_{n'm'}^2}{S_{n'm'}^2} R \right) \right\} \\ &\times \sqrt{\pi \left(\frac{S_{n'm'}^2}{4\sigma_{n'm'}^2} + 1/2l^2 \right)^{-1}} \sqrt{\pi \left(\frac{S_{nm}^2}{4\sigma_{nm}^2} + \frac{S_{n'm'}^2}{4\sigma_{n'm'}^2} R \right)^{-1}}, \end{aligned} \quad (15)$$

where

$$R = 2\sigma_{n'm'}^2 / S_{n'm'}^2 l^2 / (1 + 2\sigma_{n'm'}^2 / S_{n'm'}^2 l^2). \quad (16)$$

In Eqs. (15) and (16), all quantities with primed subscripts are assumed to be conjugated.

Equation (15) gives the ensemble average of the short-time average of the reverberation intensity directly in terms of time, modal slownesses, modal dispersion, and the correlation length scale of the scattering process. The arguments to the exponential functions all sum to a number smaller than

zero. This is intuitive since the exponential functions indicate the contributions to the reverberation between mode sets. For this reason, the product of the exponential functions along with the scaling functions inside the square roots may be viewed as a time-dependent modal correlation function ρ^2 .

In order to gain insight into the expected behavior of the reverberation intensity, it is useful to evaluate the behavior of the modal correlation function in simplified circumstances. In the limit of zero dispersion ($\sigma_{nm}^2 \rightarrow \Delta\omega^{-2}$), ρ^2 , which is the last three lines of Eq. (15), may be written

$$\begin{aligned} \rho_{nmn'm'}^2(t, l, \Delta\omega) &\rightarrow \exp \left\{ \left(\frac{t\Delta\omega}{2} \left(S_{nm} + \frac{S_{n'm'}}{1 + l^2/2L^2} \right) + \frac{i}{\Delta\omega} \left(k_n + k_m - \frac{k_{n'} + k_{m'}}{1 + l^2/2L^2} \right) \right)^2 \left(S_{nm}^2 + \frac{S_{n'm'}^2}{1 + l^2/2L^2} \right)^{-1} \right\} \\ &\times \exp \left\{ - \left(\frac{t^2 \Delta\omega^2}{4} \left(\frac{2 + l^2/2L^2}{1 + l^2/2L^2} \right) + i \frac{t(k_{n'} + k_{m'})}{S_{n'm'}(1 + 2L^2/l^2)} + \frac{(k_{n'} + k_{m'})^2 l^2}{2 + l^2/L^2} \right) \right\} \\ &\times \sqrt{\pi \left(\frac{S_{nm}^2 \Delta\omega^2}{4} + \frac{S_{n'm'}^2 \Delta\omega^2}{4 + 2l^2/L^2} \right)^{-1}} \sqrt{\pi (1/4L^2 + 1/2l^2)^{-1}}, \end{aligned} \quad (17)$$

where L is the patch size $1/\Delta\omega S_{n'm'}$.

It is instructive to evaluate Eq. (17) in the limits of large and small patch size in relation to the correlation length scale. In the large patch-size limit ($L \gg l$) we obtain⁶ for the intermodal correlation function

$$\begin{aligned} \rho_{n'n'm'}^2(t, L \gg l) &\approx e^{-(k_n+k_m-(k_{n'}+k_{m'}))^2 L^2/2} e^{-(k_{n'}+k_{m'})^2 l^2/2} \\ &\times e^{-t^2(S_{nm}^{-2}+2S_{nm}^{-1}S_{n'm'}^{-1}+S_{n'm'}^{-2})/8L^2} \\ &\times e^{it[(k_n+k_m-(k_{n'}+k_{m'}))/S_{nm}]} \end{aligned} \quad (18)$$

The first term of Eq. (18) represents a decorrelation between modes due to very large patch size and differences between the modal wave numbers which cause a cancellation phenomenon. The second term represents the traditional Bragg scattering term, which is the Fourier transform of the correlation function R_η at the sum of the incident and scattered wave numbers. The third term is a spatial decorrelation term which increases with range. This represents the decorrelation between modes due to the fact that different physical parts of the bottom are illuminated by the two round-trip mode sets. The equivalent form of the third term is

$$e^{-t^2 \Delta r^2 / 8L^2},$$

where Δr represents the range separation between the patches associated with each mode set. For broadband processes we will show that this decorrelation is controlled by the correlation length scale l instead. Finally, the argument to the fourth term is the difference between the round-trip phase between the various candidate mode sets in the modal sum. This is identical to the magnitude square of the modal interference terms in the round-trip propagation to the scattering patch and back, and so represents a coherent propagation effect.

Equation (18) represents the traditional view of scattering in the narrow-band limit, where the patch size is sufficient to allow Bragg-type scattering to occur and lower back-scattered levels are achieved. The first term in Eq. (18) also limits the amount of interference structure allowed between mode sets, and in the limit of infinitesimal bandwidth ($L \rightarrow \infty$) decorrelates all the modal cross terms, destroying the propagation structure. This transition is given explicitly in the waveguide and bandwidth parameters.

In the small patch-size limit ($l \gg L$), Eq. (17) reduces to⁶

$$\begin{aligned} \rho_{n'n'm'}^2(t, l \gg L) &\approx e^{-((k_n+k_m)^2+(k_{n'}+k_{m'})^2)L^2} \\ &\times e^{-t^2(S_{nm}^{-2}-2S_{nm}^{-1}S_{n'm'}^{-1}+S_{n'm'}^{-2})/2l^2} \\ &\times e^{it[(k_n+k_m)/S_{nm}-(k_{n'}+k_{m'})/S_{n'm'}]} \end{aligned} \quad (19)$$

where the second term is equivalent to

$$e^{-\Delta r^2/2l^2},$$

and represents a decorrelation between mode sets associated with the ratio between the distance between the mode patches Δr and the correlation length scale l . The third term is the round-trip accumulated phase as in Eq. (18), with the distinction that the phase accumulated on the return trip is

integrated over the range $t/S_{n'm'}$ associated with the primed mode set. This distinction is of little significance.

The first term in Eq. (19) represents decorrelation between mode sets associated with the spatial aperture L , and differs from the analogous term in Eq. (18) in that this decorrelation remains finite for even the diagonal terms where the mode sets are identical. In fact, this term bears more of a resemblance to the Bragg scattering term from the second line of Eq. (18), and may be thought to replace it with a term of approximately the form

$$e^{-(k_{n'}+k_{m'})^2 2/L^2}.$$

This is a key result which shows that in the limit of small patch size, Bragg scattering associated with the correlation function of the scattering centers disappears, and is replaced by a term which looks very much like Bragg scattering from a scattering process with the smaller spatial scale L . This is the manifestation of the extremely strong window effect present at short spatial window scales: the local Fourier transform of the spatial correlation function is severely distorted by the small spatial window, with the result that the amplitude of the backscattered field is determined more by the variance of the individual scatterers themselves than by resonant type terms obtained through the interrogation of multiple scatterers by the larger spatial windows obtained with narrower bandwidths.

II. RESULTS

In this section, we use the theory to predict the short-time average of reverberation intensity for the shallow-water environment observed during the SCARAB'97 experiment in the Capraia basin north of Elba.⁷ The purpose of this section is not to interpret the data from this experiment, but rather to use realistic parameters to show how the reverberation depends on the parameters of center frequency, bandwidth, scattering layer, and correlation length scale.

The SCARAB'97 experiment was conducted during summer, so the measured sound-speed profiles were downward refracting. The bottom properties in the Capraia basin have been the subject of considerable study;^{8,9} here, we adopt a fast, upwardly refracting bottom used by Jensen to adequately characterize the transmission loss.¹⁰ The sound-speed profile in the water column and into the bottom is illustrated in Fig. 1, and the bottom properties used are indicated in Table I.

KRAKEN¹¹ was used to obtain the mode-shape functions ϕ_n and the complex wave numbers k_n for this waveguide at center frequencies of 75, 150, 300, and 600 Hz. The group slownesses $\partial k_n / \partial \omega$ and the modal dispersion $\partial^2 k_n / \partial \omega^2$ were determined for these center frequencies by finite differences, using a 10-Hz step size. The resulting dispersion curves of the waveguide are indicated in Figs. 2 and 3. Inspection of the phase and group speeds of the modes shows that there is typical low-frequency shallow-water behavior, with decreasing group speed for increasing mode number. The modal dispersion illustrated in Fig. 3 indicates that the higher-order modes spread out more in time (and space) at long range than do lower-order modes. Since the sign on the

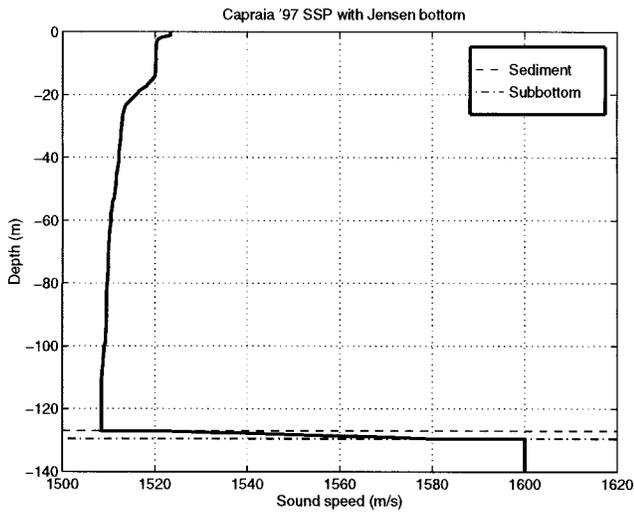


FIG. 1. The ssp of the summertime shallow-water SCARAB '97 environment chosen for this study.

dispersion is negative, higher frequencies have less slowness, and therefore travel faster, than do the lower frequencies. The result is that at a given bottom patch, the highest frequencies for any given mode reach the patch first, and the modes themselves arrive in the order of increasing mode number.

The importance of including the modal dispersion term may be understood in terms of a range at which this term effectively decreases the bandwidth by a factor of 2 (with the effect of increasing the patch size and therefore the number of contributing scatterers at related times by a factor of 2). According to the definition of σ_{nm}^2 , this range is given approximately by the relation $r_{2\Delta\omega} = 4/\Delta\omega^2 \partial^2 k_n / \partial \omega^2$. At 600 Hz for 120 Hz of bandwidth, this range occurs at 30 km for mode 23. As this range is proportional to the inverse square of the bandwidth, the importance of including this term is accentuated for greater proportional bandwidths.

The transmission loss throughout the water column and into the first few meters of bottom was calculated for center frequencies of 150 and 600 Hz. These results are shown in the top frames of Figs. 4 and 5. There are eight homogeneous modes with phase speed less than the sub-bottom speed of 1600 m/s at 150 Hz and 32 homogeneous modes at 600 Hz. The source depth of 10 m was chosen in such a way as to emphasize coherent interaction between the modes. This is particularly evident at 600 Hz, where the downward-refracting nature of the waveguide yields bundles of bottom-interacting rays with a 2- to 2.5-kilometer cycle distance. These strong bottom interactions yield increased reverberation at the two-way travel time to the appropriate ranges.

In the lower two frames of Figs. 4 and 5, the transmis-

TABLE I. Capraia basin bottom properties used as input into KRAKEN for the calculation of modal properties.

Depth	c (m/s)	ρ (g/cm ³)	α (dB/ λ)
0	1520	1.75	0.13
2.5	1580	1.75	0.13
halfspace	1600	1.80	0.15

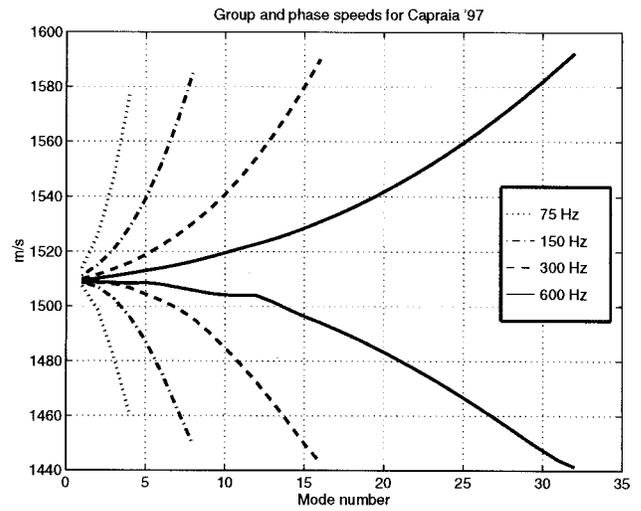


FIG. 2. The phase (upper curves) and group speed (lower curves) obtained for the environment in Fig. 1 using KRAKEN.

sion loss at the water–sediment interface and the short-time average of the reverberation received monostatically for a receiver at 75-m depth are shown for the two frequencies of interest. In both figures, the similarity between the one-way transmission loss (TL) to the scatterers and the backscattered reverberation level (RL) is apparent. The abscissas of the TL plots are scaled in such a way that the range maps directly to an approximate two-way group speed of 750 m/s.

Figure 4 shows that the strong bottom interaction at approximately 7 km at 150 Hz is seen to cause an associated large feature in the RL at $t \approx (7000 \text{ m}) / (750 \text{ m/s})$ or about 9 s. Two other strong bottom interactions at approximately 13 and 14.5 km are also seen to cause strong reverberation features at the appropriate travel time. In Fig. 5 the 600-Hz results show that the theory predicts strong reverberation features at 7.5, 14, and 17.5 s, consistent with regions of strong bottom illumination in the TL at ranges of 6, 11, and 13 km.

These results represent an overview of the types of results which can be obtained with the theory. They were obtained with a perturbation theory scattering-kernel represen-

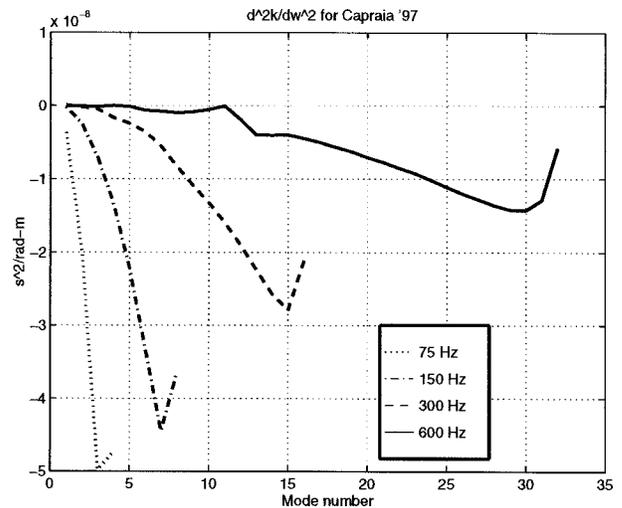


FIG. 3. The modal dispersion term $\partial^2 k_n / \partial \omega^2$ for the SCARAB '97 environment.

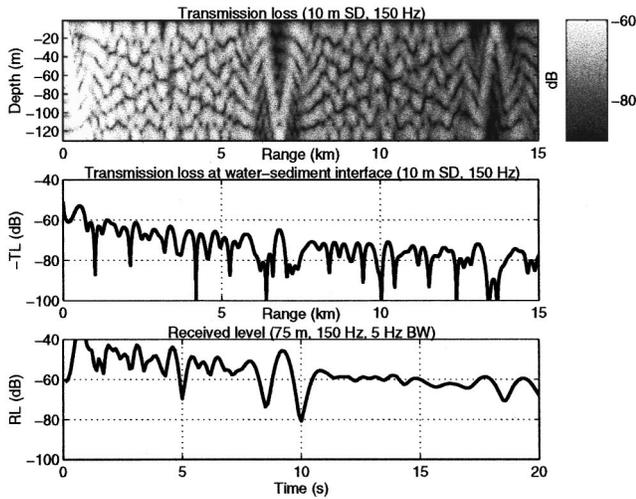


FIG. 4. This comparison between TL and RL at 150 Hz shows that the strong bottom interaction at 7 km causes an associated higher intensity return in the RL at about 9 s.

tation outlined in Appendix B for rough-surface scattering from the bottom. The characteristics of these scattering kernels for water–sediment and sediment–sub-bottom rough-surface scattering is outlined in the following subsection.

A. Contributions from water–sediment and sediment–sub-bottom roughness

In the theory, two separate parameters together determine the importance of scattering into and out of the various modes. These parameters are the scattering kernel T_{mn} and the modal attenuation associated with the complex modal eigenvalues $k_{n,m}$. For the simplified rough-surface scattering theory developed in Appendix B, the scattering kernel is defined as the outer product between ϕ_m^- , the ability of the incident field to excite a scatterer, and ϕ_n^+ , the ability of the scatterer to excite the backscattered field. As discussed in the Appendix, the ability to excite rough-surface scattering is approximately determined by the difference between the sec-

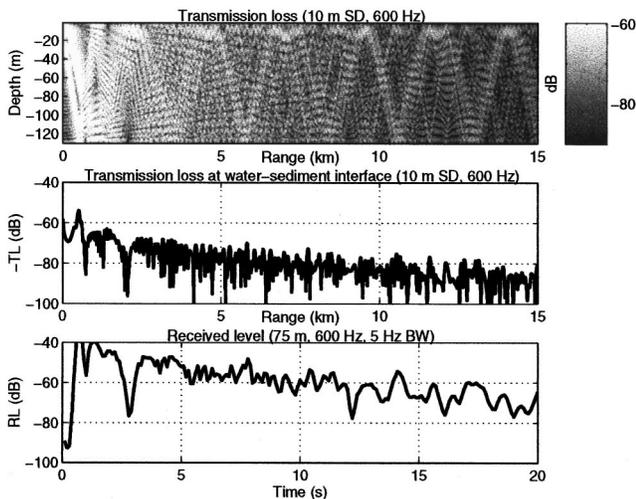


FIG. 5. At 600 Hz the TL structure begins to look strongly raylike. The bottom near 6 and 10.5 km is particularly strongly illuminated by several ray bundles, causing associated features in the RL at 7.5 and 14 s.

ond derivative of the mode-shape function at the interface and the square of the first derivative normalized by the mode-shape function at the scatterer depth

$$\phi_m^- \approx \left(\frac{(\partial \phi_m / \partial z)^2}{\phi_m} - \frac{\partial^2 \phi_m}{\partial z^2} \right) \Bigg|_{z=z_{\text{scat}}}$$

In this case, the ability of the scattered energy to reradiate into modes is taken simply as the point-source expansion into the backward propagating modes

$$\phi_n^+ = \phi_n.$$

Under these assumptions, the scattering kernels T_{mn} for rough-surface scattering are the magnitude square of the outer product of the scattering functions ϕ_m^- and ϕ_n^+ , quantities which can be computed using the known mode-shape functions. It is emphasized that this choice for the scattering function is illustrative and more realistic results may be obtained using volume-scattering kernels or other undetermined kernels, depending on the environment. However, the simplified perturbation theory kernels selected for this discussion contain much of the important structure which is expected to be shown by most scattering kernels, and therefore are sufficient for the purpose of exercising the model.

The results for the simplified perturbation theory rough-surface scattering kernels are illustrated in Fig. 6 for scattering from the water–sediment interface, and in Fig. 7 for scattering from the sediment–sub-bottom interface. These results show that in addition to about 20–40 dB less scattering potential from the sediment–sub-bottom scatterers, due to the lower impedance contrast and the increased difficulty of the scatterers to reradiate energy into the waveguide, there is also a markedly different spectral characteristic to the response. For example, it can be seen that there is much lower scattering out of, and back into, the lower-order modes, especially at 600 Hz. This is caused by the reduced interaction of the lower-order modes with the sediment–sub-bottom interface due to the fast sub-bottom and the lower grazing angle of these modes.

As reverberation time increases, the second factor begins to assert its dominance over the angular characteristics of the backscatter. The modal attenuation, illustrated in dB per km in Fig. 8 for two-way travel in the incoming and the backscattered mode pair, filters out the backscattering contribution from the higher-order modes. Thus, we can define an *effective scattering kernel*

$$T_{mn}^{\text{eff}} \equiv |\phi_m^- e^{-R\Im\{k_m\}} e^{-R\Im\{k_n\}} \phi_n^+|^2,$$

which is an explicit function of scatterer range R and may be calculated from the scattering functions ϕ_m^- and ϕ_n^+ and the complex modal eigenvalues. The results for water–sediment interface scatterers at a scatterer range of 16 km are illustrated in Fig. 9. Comparison between Figs. 6 and 9 shows that the higher incident and scattered mode numbers contribute significantly less to the total scattering from a scatterer at 16 km than they do from a scatterer at short range. Thus, there is expected to be an evolving angular structure to the reverberation, with late-time reverberation having a heavier

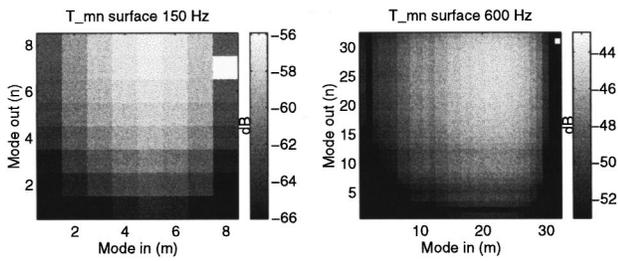


FIG. 6. Scattering kernel T_{mn} for rough-surface scattering from the water–sediment interface. Lower-order modes excite less scattering, and reradiate less effectively. This kernel is obtained under a perturbation approximation. The white square in the upper right-hand corner of these figures is an artifact of the plotting program.

weighting in the lower-order modes than early-time reverberation.

B. Effect of center frequency on RL

The center frequency controls the characteristics of the reverberation by changing the characteristics of the forward propagation and by changing the magnitude and shape of the scattering kernel. The ratio of the wavelength to the correlation length scale also has a very strong influence on the amplitude of the reverberation for the Gaussian scatterer spatial correlation properties assumed in this derivation. For perturbation theory, scattering from homogeneous boundaries is predicted to increase proportional to frequency to the fourth power. However, as frequency increases for a fixed correlation length scale, the Bragg scattering condition begins to reduce the component of energy backscattered to the receiver. The result is that for scatterers distributed according to a Gaussian correlation function, the theory predicts that there is a frequency of maximum backscattering.

The frequency of maximum backscatter is that frequency above which, for sufficiently narrow bandwidths, resonant scattering effects would cause a rapid falloff of the RL. For Gaussian-correlated scatterers, this frequency has a magnitude of approximately $230/l$. At this frequency, the power of the resonant scattering term in Eq. (18) is 6 dB below its maximum value. For example, the falloff should occur for frequencies above approximately 460 Hz for a correlation length scale of 0.5 m. For frequencies below the resonant-scattering threshold, we can expect the RL to rise proportional to the frequency dependence of the scattering kernel. Perturbation-type scattering kernels have a magnitude which increases proportional to f^4 . Thus, for these types of

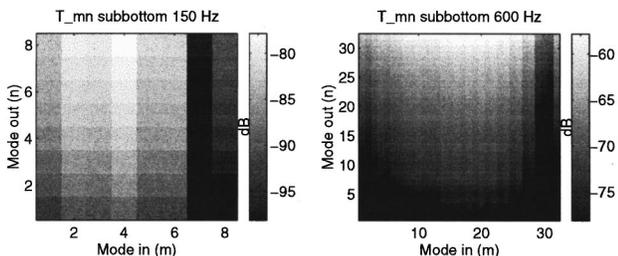


FIG. 7. Scattering kernel T_{mn} for sediment–sub-bottom rough-interface scattering. These kernels have approximately 20–40 dB less power than the water–sediment interface kernels in Fig. 6.

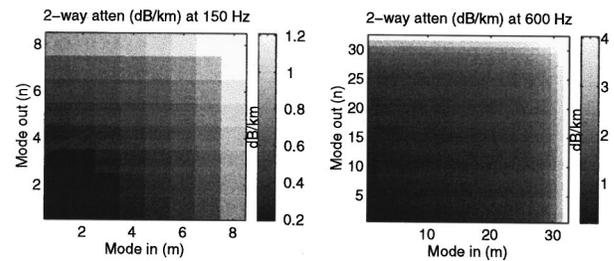


FIG. 8. The two-way loss due to the bottom loss tangent for the SCARAB’97 environment.

kernels, the theory predicts that the RL grows proportional to f^4 below the “resonance” frequency, and falls off faster than any power of frequency above it.

Although the theory cannot explicitly account for correlation function behavior consistent with power-law distributions, it may be easily generalized that the high-frequency asymptote for fractal¹² bottoms with dimension 2 (k^{-3} dependence) will be proportional to f , and that for surfaces with a fractal dimension of 3 (space filling, very rough bottoms with k^{-1} dependence) the RL will grow proportional to f^3 beyond the resonance frequency. In no case should these types of arguments be taken for frequencies above the range of validity of the perturbation approximation itself.

In Fig. 10 the frequency dependence of the RL for scattering from a water–sediment interface roughness with a correlation length scale of 0.5 m is shown for the 20-Hz bandwidth case. The maximum RL is seen to occur at 300 Hz. Our simple relation predicts a maximum RL at $230/0.5$ or approximately 460 Hz. It may also be seen from the results that the RL increases roughly 12 dB between 75 and 150 Hz, as is predicted for perturbation theory. Between 150 and 300 Hz the increase is smaller, probably because 300 Hz is relatively close to the resonant frequency.

C. Effect of bandwidth on RL

In the development of the theory it became evident that as the bandwidth of the interrogating waveform is increased, the amount of coherent structure expected in the RL is decreased. In Fig. 11 the RL expected at 600 Hz from a surface with a 0.5-m correlation length scale is indicated for processing bandwidths of 5, 20, 80, and 160 Hz. The black curves represent scattering from the water–sediment interface with a roughness of rms amplitude 1 m, and the gray curves rep-

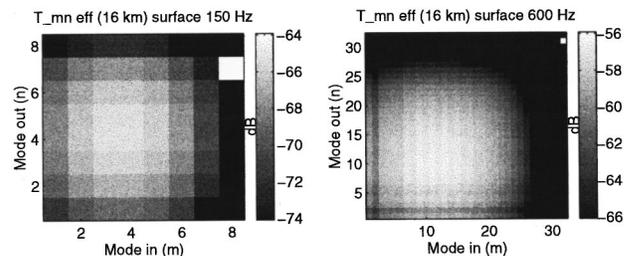


FIG. 9. The effective scattering kernel T_{mn}^{eff} for scattering from water–sediment interface roughness 16 km from the source. The high attenuation for the higher-order modes results in reduced contributions from these mode pairs. The white square in the upper right-hand corner of these figures is an artifact of the plotting program.

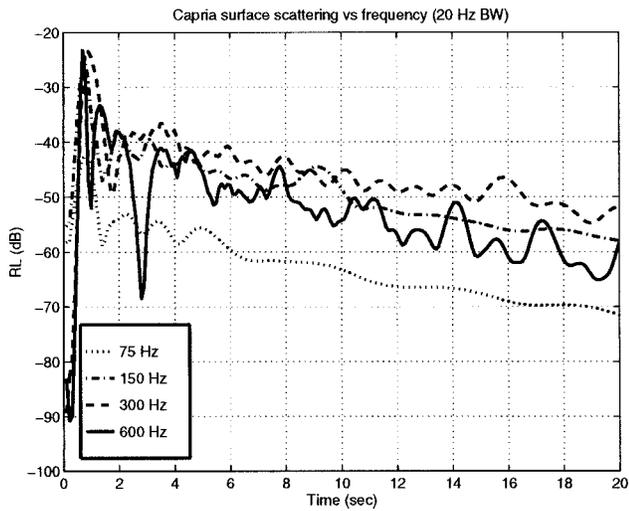


FIG. 10. The RL as a function of frequency for water–sediment interface roughness scattering from a roughness with a 0.5-m correlation length scale assuming 20 Hz of bandwidth. The RL is predicted to have a maximum near 460 Hz due to correlation length scale effects.

resent scattering from the same amplitude roughness at the sediment–sub-bottom interface. The RL caused by the sediment–sub-bottom scatterers is seen to be approximately 30 dB lower than the water–sediment interface result, which is consistent with the fact that the scattering kernel itself is about 20–40 dB lower. The results for the larger bandwidths are seen to be significantly smoother than for the 5-Hz bandwidth, due to the lack of predictable interference terms between the modes, and it can also be seen that there are slight differences between the coherent structure of the RL for scattering from the two different interfaces. Despite the reduction in the modal coherence for increased bandwidths and

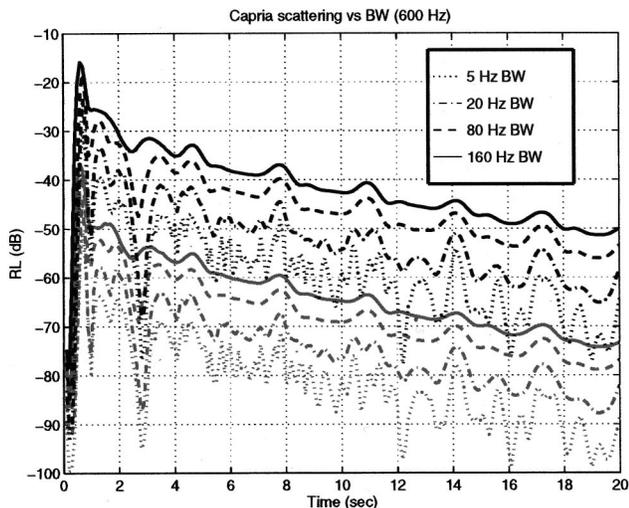


FIG. 11. The RL at 600 Hz for 5, 20, 80, and 160 Hz of bandwidth. Very strong bottom interactions at round-trip travel times of 5, 8, 11, 14, and 17 s are observed on a background of decreasing oscillatory structure at the higher bandwidths. The lack of predictable structure in the RL in no way implies that resolution is lost for higher bandwidths; rather there is a loss of intermodal coherence in the reverberation alone. This decorrelation increases with bandwidth as the patch sizes of the individual modes decrease and the different modes begin to interact with uncorrelated pieces of the scatterer distribution. Water–sediment interface scattering is shown in black; sediment–sub-bottom scattering in gray.

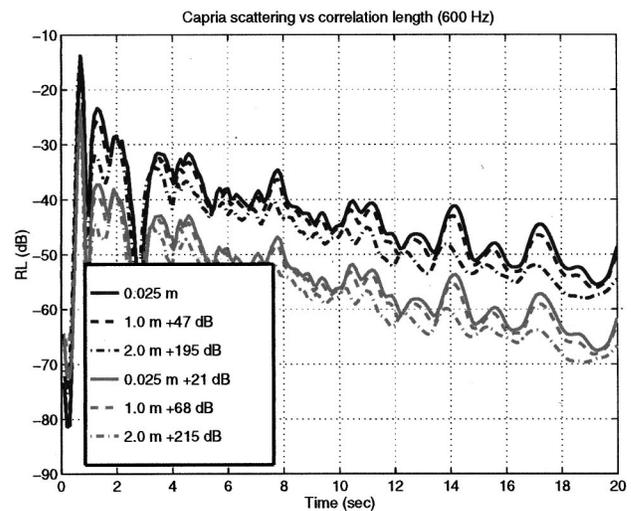


FIG. 12. RL as a function of correlation length scale at 600 Hz. The RL observed to come from the strongest bottom interactions suffers a relative attenuation of 5–8 dB for the longest correlation length. Water–sediment interface scattering is shown in black; sediment–sub-bottom scattering in gray.

late times, it is interesting to note that the RL features associated with raylike interactions with the bottom are clearly visible at even the largest bandwidths considered.

D. Effect of correlation length scale on RL

As discussed above, for Gaussian-correlated scatterers there exists a *resonance frequency* at which the RL is maximized. At fixed frequency and for narrow bandwidths, there is a related *resonance correlation length scale* with a magnitude of $l_r \approx 230/f$. For Gaussian-distributed scatterers with correlation length scales longer than the resonance correlation length scale, the RL is predicted to decrease proportional to e^{-l^2/l_r^2} . In Fig. 12 the effect of correlation length scale is shown for water–sediment and sediment–sub-bottom rough-surface scattering for the 600-Hz center frequency. These results are computed using a bandwidth of 20 Hz. Correlation length scales considered vary between 0.25 and 2 m. To ease comparison between the results, all the curves have been normalized up to the RL level for scattering from the 0.25-m correlation length scale roughness. The black curves represent the RL for water–sediment interface scattering, and the gray curves indicate the RL caused by sediment–sub-bottom interface scattering. The amount the curves have been moved is indicated in the legend. The results indicate that the amount of scattering expected from roughness on either interface is a very strong function of the correlation length scale. This is as predicted, since at this frequency the resonance length scale above which the backscatter drops off very quickly is approximately 0.375 m. The results indicate the response from 1-m roughness is already 47 dB lower than from 0.25-m roughness. At lower frequencies, results which are not shown indicate that the RL plateaus for correlation length scales shorter than the resonant length scale, as predicted.

The results in Fig. 12 also show that there is a very evident signature of decreasing response at the ray cycle

times (8, 11, 14, and 17 s) with increasing correlation length scale, due to the much smaller backscatter contribution from the dominant lower-order modes for the larger resonant wave number $k_n + k_m$ arguments to the Bragg scattering term for these modes. In this example, it is also seen that there is very little difference between the RL caused by the water-sediment and the sediment-sub-bottom scattering mechanisms. The differences seem to lie mostly in the overall decrease in the scattering levels for the sediment-sub-bottom mechanism. Of course, the magnitude of the differences in the RL between the two mechanisms will be dependent on the magnitude of the angular differences between the effective scattering kernels. In general, it should be expected that at later times, the reverberation caused by the sediment-sub-bottom scattering process should die off at a faster rate, as the most efficient propagation is in the lower-order modes which do not effectively excite scattering at this interface.

E. Comparison of closed form and Monte Carlo solutions

To verify that the behavior of the closed-form solution in Eq. (15) for the ensemble average of the short-time average of the reverberation intensity is in agreement with what one would obtain by taking an actual ensemble of independent backscattered intensity observations, a modified version of Eq. (8) for the backscattered pressure was implemented where the azimuthal integral was implicitly performed by replacing the scatterer density in range and azimuth $\eta(r, \theta)$ with an effective scatterer density explicit in range only whose standard deviation falls off proportional to one over the square root of the scatterer range, consistent with the expression for the effective one-dimensional scatterer variance derived in Eq. (12)

$$\hat{\eta}(r) = \eta(r, \theta) \Big|_{\theta = \theta_{\text{ref}}} \frac{(2\pi)^{1/4} \pi}{\sqrt{rl}}. \quad (20)$$

In this case, Eq. (6) may be rewritten as a range integral only involving the known narrow-band properties of the waveguide as defined in Sec. I

$$p(t, z_s, z_r) = A \Re \left\{ e^{i\omega_0 t} \sum_{m=1}^N \sum_{n=1}^N \sigma_{nm} \frac{\phi_m^o(z_s) T_{nm}^o \phi_n^o(z_r)}{\sqrt{k_n^o k_m^o}} \times \int_0^\infty dr \hat{\eta}(r) e^{i(k_n^o + k_m^o)r} e^{-(t - S_{nm}r)^2 / 4\sigma_{nm}^2} \right\}. \quad (21)$$

The transformation in Eq. (20) has the property that the ensemble average of the square of Eq. (21) is identically equal to Eq. (15) for Gaussian-distributed correlation functions. Comparisons between the ensemble average of the square of Eq. (21) averaged over 50 realizations of the effective one-dimensional scatterer realizations in Eq. (20) are illustrated in Fig. 13 for the 600-Hz case with a source depth of 10 m, a receiver depth of 20 m, and a correlation length scale of 0.5 m. Results are shown for 5, 20, 80, and 160 Hz of bandwidth. The results show very good agreement between the Monte Carlo ensemble averages and the closed-

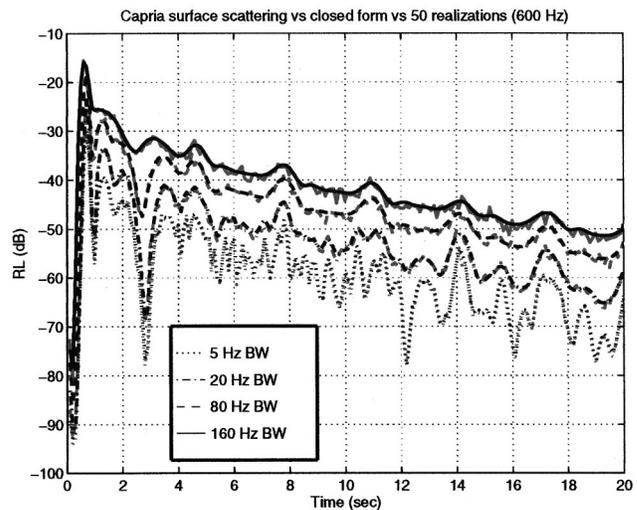


FIG. 13. Comparison between closed-form (black) and Monte Carlo (gray) expressions for the RL at 600 Hz for 5, 20, 80, and 160 Hz of bandwidth. The agreement in all cases is very good, giving confidence that the closed-form expressions are correct.

form expressions, and serve as an important verification of the closed-form expressions.

In general, one may summarize the tradeoff between the closed form expressions and the Monte Carlo technique by stating that the closed-form expressions have the value that (1) they provide insight, and that (2) they are more efficient to calculate at low frequencies. The form of the Monte Carlo solutions provides less insight, but at higher frequencies for moderate bandwidths they are more efficient to calculate. However, as we discuss in the next subsection, the Monte Carlo solutions offer the additional outstanding opportunity to evaluate backscatter from scatterer distributions which are spatially correlated in a non-Gaussian way. The Gaussian assumption places a very unrealistic restriction on the closed-form theory. The ability to treat general scatterer correlation functions is perhaps the most outstanding capability offered by the Monte Carlo solution technique.

F. Effect of correlation function form on RL

Equation (21) has the value that scattering from roughness profiles with various non-Gaussian correlation properties may be evaluated and compared to one another and the closed-form Gaussian results. Almost all real-world scatterer distributions have non-Gaussian spatial correlation characteristics. The most typical spatial correlation properties, which involve surfaces of various fractal dimensions, yield power spectra which are distinctly non-Gaussian and instead obey an inverse power law in the spatial wave number.¹² The effect of these more realistic scatterer distributions is that as the resonant wave number $k_n + k_m$ increases for Bragg scattering, the backscatter falls off much more slowly [$\propto (k_{nm}l)^{-(1-3)}$] than in the case of Gaussian scatterer spectra ($\propto \exp\{-(k_{nm}l)^2\}$). Not only does this imply that the mean levels of the reverberation are higher for surfaces which are correlated in a fractal way, but also that the components backscattered from lower-order modes (which give larger resonant wave numbers) can give a much larger contribution

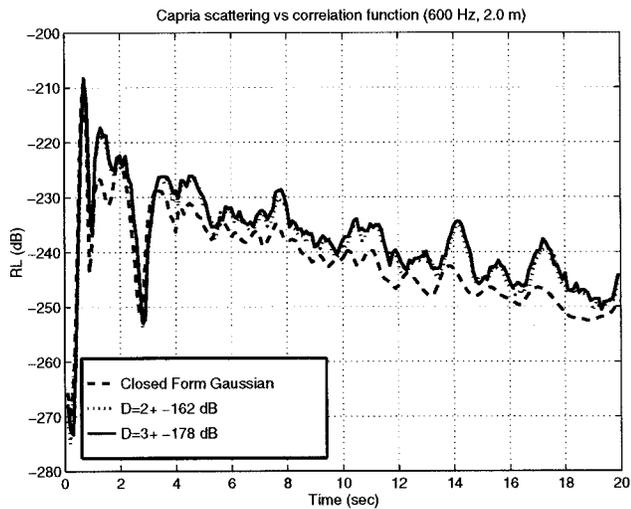


FIG. 14. RL at 600 Hz for Gaussian, and two power-law scatterer distribution power spectra with a correlation length scale of 2.0 m. The very low Bragg scattering component into the lower-order modes in the Gaussian case significantly reduces the backscattering contributions associated with the strong bottom interactions at 8, 11, 14, and 17 s.

to the total backscatter from fractal scatterers than from Gaussian-correlated scatterers, in those cases where the correlation length scale is longer than a wavelength.

The difference between the backscattered intensity from scatterers with Gaussian and non-Gaussian correlation properties at 600 Hz is illustrated in Fig. 14. These results show that for a correlation length scale of 2.0 m, the backscatter from convergence zone-type propagation features on the bottom is much more pronounced for 2-D surfaces with fractal dimension 2 to 3 (solid and dotted, respectively) than it is for the Gaussian surfaces (shown dashed for the closed-form solution,) by a factor of as much as 10 dB. Since the lower-order modes suffer less propagation loss at long range, and they scatter more from the fractal surfaces, the increase becomes more pronounced at later times. Since the rougher surface (fractal dimension 3) has higher backscatter by only about 1–2 dB than the smoother power-law surface, it is also evident that the power-law parameter is not expected to be strongly observable in the reverberation time series.

G. Spatial–temporal coherence of monostatic backscatter

The closed-form theory or the Monte Carlo results may be used to estimate the coherence of the backscatter as a function of vertical separation and time after shot. In the present formulation of the closed-form solution, the coherence between two receivers at two different times has not been obtained, although the modification of the theory to obtain this is straightforward. The coherence between the receiver pair $[z_1, z_2]$ on a vertical line array at time t is given in Eq. (10) as

$$\rho^2(t, z_1, z_2) = \frac{P_{STA}^2(t, z_1, z_2)}{\sqrt{P_{STA}^2(t, z_1, z_1) P_{STA}^2(t, z_2, z_2)}}.$$

This quantity was evaluated at 150 and 600 Hz and plotted as a function of z_1 and z_2 for 20-Hz of bandwidth for early time

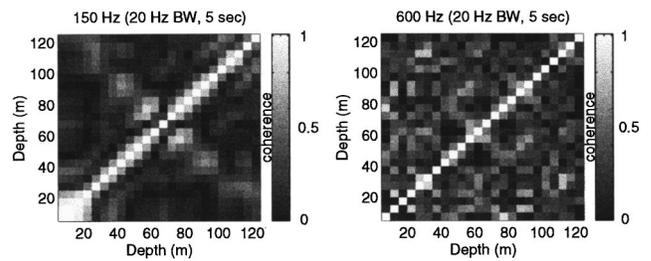


FIG. 15. Vertical coherence at 5 s after the shot for 20 Hz of processing bandwidth.

(5 s) in Fig. 15 and for late time (20 s) in Fig. 16. The results indicate a decrease in spatial coherence for very large offsets at late time. The decrease in coherence for large receiver separation is caused by the modal decorrelation discussed in the Theory section of this paper. Since the spatial coherence may be thought of as a *change of basis* on the modal coherence, a decrease in intermodal coherence directly implies the reduced spatial coherence observed in these results.

In general, it has been observed through exercising the model that there can be a pronounced increase in spatial coherence at late time as bandwidth is decreased. For example, the results for the vertical coherence at 600 Hz for only 5 Hz of processing bandwidth have been observed to be nearly as coherent a full 20 s after the shot as the 20-Hz bandwidth results only 5 s after the shot.⁶

These vertical coherence results are anticipated to be useful as they serve to provide a very important quantity for evaluating the performance of various signal-processing algorithms for active detection in reverberation-limited environments. These results can also be used as constraints in the design of processors which are more robust to waveguide reverberation.

III. CONCLUSION

The short-time average of the monostatic backscattered reverberation intensity in a waveguide has been obtained as a function of (1) scatterer parameters of (i) correlation length scale, (ii) correlation function, and (iii) scatterer interface or depth, (2) the processing bandwidth parameter, and (3) the channel parameters of (i) number of propagating modes, and (ii) their associated wave numbers, shapes, slownesses, and curvatures, and (iii) their ability to excite the scatterers. The results indicate that as processing bandwidth is increased, the predictable structure of the reverberation is decreased, with an accompanying decrease in the vertical coherence, at late

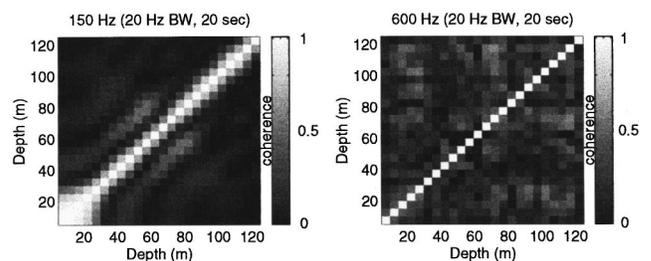


FIG. 16. Coherence at 20 s and 20 Hz of bandwidth. The results are quite diagonal, with decreasing spatial coherence scale as frequency increases.

time. The results also show that scatterer depth, for instance on the water–sediment interface, or at the sediment–sub-bottom interface, can significantly affect the levels, and the temporal characteristics of the received reverberation. Changes in the correlation length scale can also produce observable changes in the reverberation structure for otherwise fixed processing and channel parameters. For Gaussian-correlated scatterers, a frequency of maximum reverberation is predicted, below which reverberation intensity grows proportional to f^4 for perturbation theory scattering kernels, and above which the reverberation intensity falls off faster than any power of frequency.

Results obtained in this paper also indicate that the spatial structure of the forward propagation has a very significant influence on the expected temporal structure of the reverberation. Convergence zones onto the scattering layer and other pathological features in the forward propagation can cause associated large peaks in the reverberation intensity at the appropriate round-trip travel time. The characteristics of the reverberation features turn out to be well-defined functions of the waveguide and scatterer parameters, the center frequency, and the bandwidth. Until now, most scattering theories have ignored the coherent interaction between the various propagation paths. The work presented here shows that these terms need not be neglected, that they can be treated in a robust way, and that when these terms are retained additional insights are gained into the rich physics of waveguide reverberation. In addition, since propagation-induced features are commonly observed in reverberation data sets, retention of these terms is also an advantage for developing a better understanding of these features when they are present.

The closed-form solutions of the short-time average of the backscattered intensity developed in this paper have been benchmarked against self-consistent Monte Carlo predictions, with good results. The Monte Carlo predictions have also been used to predict the differences between the reverberation caused by Gaussian and non-Gaussian correlated scatterers. These results show the observability of scatterer correlation properties in the reverberation signature. The reverberation time series generator used to obtain the Monte Carlo results also delivers a capability to evaluate the higher moments of the reverberation statistics for various scatterer correlation length scales, spatial correlation properties, and amplitude distributions. Although this capability was not explored in this work, it is believed that this capability will prove valuable for developing an understanding of the non-Rayleigh distributions of reverberation intensity which are often observed in shallow-water reverberation data.¹³

In general, the theory enhances understanding of how waves interact with roughness to cause reverberation in waveguides. How the coherence between the backscattered modes is affected by time after shot, correlation length scale, and patch size have all been rigorously identified in the closed-form solution, and various limiting behaviors have been identified and discussed. In particular, the results show that for very narrow bandwidths, the traditional view of reverberation as Bragg scattering from patches of limited spatial extent is correct, but for very large bandwidths, Bragg

scattering does not occur and reverberation levels are much higher and more spatially uncorrelated. The theoretical results also show that coherent interference structure between modes which interrogate different parts of the bottom can be expected when the correlation length scale is longer than the modal separation distance, even if the patch sizes are very small. These results, while heuristically predictable, are given here explicitly for the first time in terms of the various parameters which define a monostatic reverberation experiment.

Finally, the time-dependent coherence between the modes derived in this paper has provided a means for predicting the vertical coherence in the waveguide as a function of relevant parameters. It is believed that this capability of the theory will prove useful for guiding the design of robust algorithms for active localization of targets in reverberation-limited environments.

ACKNOWLEDGMENTS

The author would like to acknowledge the fruitful discussions held with his colleagues at SACLANTCEN and BBN.

APPENDIX A: EVALUATION OF RANGE INTEGRALS

To evaluate the r and r'' integrals in Eq. (14), we must first appeal to two approximations. The first approximation is that the range integrals over r and r'' may both be artificially extended to minus infinity. This approximation is found to be good for times greater than the pulse duration $2\pi/\Delta\omega$. For these times, all the scattering contributions come from ranges greater than zero, and so the scattering integrals may be artificially extended to minus infinity for the convenience of their evaluation.

The second approximation is that for the closed-form evaluation of the scattering integrals, the explicit dependence of σ_{nm}^2 and $\sigma_{n'm'}^2$ on range must be approximated by a dependence on time over round-trip slowness S_{nm} (or $S_{n'm'}$)

$$\sigma_{nm}^2 \approx \Delta\omega^2(1 + iD_{nm}\Delta\omega^2 t/2S_{nm})^{-1}.$$

The same approximation is also required for the r^{-1} term in the denominator of the correlation function which was found in Eq. (13) to result from the azimuthal integration of the scattering correlation function. This term is replaced by the approximation

$$r^{-1} \approx S_{nm}/t.$$

This approximation is good when the range differential over the entire patch is a small fraction of the average range, a condition satisfied for

$$2\pi\omega_o/\Delta\omega k_{nm} \ll t\omega_o/k_{nm},$$

or for times much greater than 1 over the bandwidth of the pulse function in Hertz.

Use of the two above approximations allows the closed-form evaluation of the range integrals through the use of standard techniques. For details of the evaluation, see Ref. 6.

APPENDIX B: SIMPLIFIED EXPRESSIONS FOR SMALL PERTURBATION SCATTERING FROM LOCALLY REACTING BOUNDARY IMPEDANCE

Consider the following boundary condition at a rough interface in a waveguide:

$$Z \equiv \frac{p}{\partial p / \partial n}.$$

Let us call the left-hand side of this equation the boundary operator B which operates on the total pressure p .

In perturbation theory, we expand the boundary operator about a local roughness excursion η , and let it operate on the sum of the unperturbed and scattered fields p and p_s . In the limit of zero scatterer slope, we have

$$Z = \left(B + \eta \frac{\partial B}{\partial z} \right) (p + p_s), \quad (\text{B1})$$

where

$$B = \frac{p}{\partial p / \partial z},$$

and

$$\frac{\partial B}{\partial z} \equiv 1 - \frac{p}{(\partial p / \partial z)^2} \frac{\partial^2 p}{\partial z^2}.$$

In this case, Eq. (B1) may be written

$$Z = \frac{p + p_s}{\partial(p + p_s) / \partial z} + \eta \left(1 - \frac{p + p_s}{(\partial(p + p_s) / \partial z)^2} \frac{\partial^2(p + p_s)}{\partial z^2} \right),$$

or approximately

$$Z \approx \frac{p + p_s}{\partial p / \partial z} - \frac{p + p_s}{(\partial p / \partial z)^2} \frac{\partial p_s}{\partial z} + \eta \left(1 - \left[\frac{p + p_s}{(\partial p / \partial z)^2} + 2 \frac{p + p_s}{(\partial p / \partial z)^3} \frac{\partial p_s}{\partial z} \right] \frac{\partial^2(p + p_s)}{\partial z^2} \right).$$

Retaining terms of order p , we have

$$Z = \frac{p}{\partial p / \partial z},$$

while to first order in η and p_s we have

$$\frac{p_s}{\partial p / \partial z} - \frac{p}{(\partial p / \partial z)^2} \frac{\partial p_s}{\partial z} = -\eta \left(1 - \frac{p}{(\partial p / \partial z)^2} \frac{\partial^2 p}{\partial z^2} \right). \quad (\text{B2})$$

Equation (B2) may be recast to give the vertical derivative of the scattered field in terms of the scattered field itself and the first and second vertical derivatives of the unperturbed pressure p

$$\frac{\partial p_s}{\partial z} = \frac{p_s}{Z} + \eta \left(\frac{(\partial p / \partial z)^2}{p} - \frac{\partial^2 p}{\partial z^2} \right). \quad (\text{B3})$$

Now, we may use Green's third identity to formulate the scattered pressure received at any point in the waveguide in terms of the scattered field and its vertical derivative along the boundary of the waveguide

$$p_s = \frac{1}{2\pi} \int dS \left(p_s \frac{\partial G}{\partial z} - G \frac{\partial p_s}{\partial z} \right) = \frac{1}{2\pi} \int dS \left(p_s \frac{G}{Z} - G \frac{\partial p_s}{\partial z} \right), \quad (\text{B4})$$

which upon the insertion of Eq. (B3) yields

$$p_s = \frac{1}{2\pi} \int dS \ G \eta \left(\frac{\partial^2 p}{\partial z^2} - \frac{(\partial p / \partial z)}{Z} \right). \quad (\text{B5})$$

Equation (B5) may be written in final form by using the following far-field approximations for the unperturbed waveguide:

$$G(z, z_b; r) = \frac{i e^{-i\pi/4}}{\sqrt{8\pi\rho(z_b^+)}} \sum_m \phi_m(z_b) \phi_m(z) \frac{e^{ik_m r}}{\sqrt{k_m r}}, \quad (\text{B6})$$

and

$$p(z_b; r) = e^{-i\pi/4} \sqrt{2/\pi} \sum_n a_n \phi_n(z_b) \frac{e^{ik_n r}}{\sqrt{k_n r}}, \quad (\text{B7})$$

where a_n are the modal amplitude coefficients excited by the source. Inserting Eqs. (B6) and (B7) into Eq. (B5) we obtain the final expression for the scattered field obtained on the receiver as a function of the mode-shape functions and wave numbers of the waveguide and the roughness η

$$p_s(z) = \frac{1}{4\pi^2 \rho(z_b^+)} \sum_n \sum_m a_n \phi_m(z_b) \phi_m(z) \frac{e^{i(k_n + k_m)r}}{\sqrt{k_n k_m}} \times \int_0^{2\pi} d\theta \int_0^\infty dr \ \eta(r, \theta) \left(\frac{\partial \phi_n / \partial z}{Z} - \frac{\partial^2 \phi_n}{\partial z^2} \right) \Bigg|_{z=z_b^+}. \quad (\text{B8})$$

1. Frequency dependence of scattered intensity

For an isospeed waveguide with depth D overlying a rigid bottom, the mode-shape functions have the form

$$\phi_n = \sin((2n-1)\pi z/2D),$$

and the boundary impedance Z is infinite. In this case, the term in the parentheses of Eq. (B8) is

$$-\left(\frac{\partial^2 \phi_n}{\partial z^2} \right) \Bigg|_{z=z_b^+} = -((2n-1)\pi/2D)^2.$$

The effective grazing angle of the mode is

$$\theta_n \approx k_z/k = ((2n-1)\pi/2D)/k = c((2n-1)\pi/2D)/\omega. \quad (\text{B9})$$

Thus, the scattered intensity at the grazing angle associated with the n th mode is proportional to frequency and grazing angle raised to the fourth power

$$I_s(\theta_n) \propto (\omega \theta_n / c)^4.$$

¹D. D. Ellis, "A shallow water normal-mode reverberation model," J. Acoust. Soc. Am. **97**, 2804–2814 (1995).

²D. Tang, "Shallow water reverberation due to sediment volume inhomogeneities," IEEE J. Ocean Eng. (submitted).

³B. H. Tracey and Henrik Schmidt, "A self-consistent theory for seabed volume scattering," J. Acoust. Soc. Am. (to be published).

⁴K. D. LePage, "Time series variability in fluctuating ocean waveguides,"

SACLANT Undersea Research Centre SR-319, La Spezia, Italy, 1999.

⁵I. Dyer, "Statistics of sound propagation in the ocean," *J. Acoust. Soc. Am.* **48**, 337–345 (1970).

⁶K. D. LePage, "Bottom reverberation in shallow water: Coherent properties as a function of bandwidth, waveguide characteristics and scatterer distributions," SACLANT Undersea Research Centre SR-301, La Spezia, Italy, 1998.

⁷C. Holland, "Capraia SCARAB," SACLANT Undersea Research Centre Cruise Report, La Spezia, Italy, July, 1997.

⁸D. F. Gingras and P. Gerstoft, "Inversion for geometric and geoacoustic parameters in shallow water: Experimental results," *J. Acoust. Soc. Am.* **97**, 3589–3598 (1995).

⁹C. Holland and J. Osler, "High resolution geoacoustic inversion in shallow water using a towed broadband source and a bottom moored re-

ceiver," *J. Acoust. Soc. Am.* (to be published).

¹⁰F. B. Jensen, "Comparison of transmission loss data for different shallow water areas with theoretical results provided by a three fluid normal-mode propagation model," in *Sound Propagation in Shallow Water*, edited by O. F. Hastrup and O. V. Olesen, CP-14, SACLANT Undersea Research Centre, La Spezia, Italy, pp. 79–92, 1974.

¹¹M. B. Porter, "The KRAKEN Normal Mode Program," SACLANT Undersea Research Centre SM-245, La Spezia, Italy, 1991.

¹²J. A. Goff and T. H. Jordan, "Stochastic modeling of seafloor morphology: inversion of sea beam data for second order statistics," *J. Geophys. Res.* **93**, 13589–13608 (1988).

¹³D. A. Abraham, "Modeling non-Rayleigh reverberation," SACLANT Undersea Research Centre SR-266, La Spezia, Italy, 1997.

Broadband source localization in shallow water in the presence of internal waves

Kwang Yoo and T. C. Yang

Naval Research Laboratory, Washington, D.C. 20375

(Received 28 July 1998; revised 9 July 1999; accepted 18 August 1999)

Source localization is a challenging problem in a time-varying random shallow-water environment in the presence of internal waves. The short correlation time of the acoustic field (a couple of minutes) makes it difficult to adjust the replica field in response to the environmental changes. Also, mode couplings induced by the internal waves are random and not predictable. In this paper, the effect of mode coupling on source localization in the presence of Garrett–Munk and solitary internal waves is investigated. It is found that mode coupling affects practically all modes in shallow water. Random-mode coupling reduces the matched-field correlation. To improve source localization, a broadband matched-beam processing algorithm is proposed to suppress the noncoherent coupled-mode contributions. It is shown that this technique significantly improves the localization performance over conventional matched-field processing. [S0001-4966(99)03112-4]

PACS numbers: 43.30.Bp [SAC-B]

INTRODUCTION

Source localization in a time-varying ocean waveguide with a complex sound-speed profile, such as many found in coastal waters, is a difficult and challenging problem. Matched-field source localization can be degraded by environmental mismatch between the actual ocean acoustic environment and the environment used to compute the replica fields, as the ocean environment is constantly varying with time and not readily amenable to current remote-sensing methods. Internal waves are a dominant source of ocean variations in shallow waters. They impact significantly signal propagation at low (hundreds of Hz) frequencies. Currently, there does not exist a real-time prediction capability of the random variations caused by the internal waves.

The effect of internal waves on matched-field processing (MFP) in deep ocean has been studied by Jackson and Ewart for signals of a few hundred Hz at ranges of a few tens of kilometers.¹ They assumed a time-independent replica field and time-varying data fields in the presence of Garrett–Munk internal waves. A statistical approach was used to evaluate the average performance degradation of matched-field ambiguity surface near the source location. The calculations were carried out assuming a spatially stationary correlation function for the refractive index in a weak scattering environment.

Krolik developed a minimum variance (MV) matched-field processor designed to be robust in a random range-dependent waveguide.² Book and Nolte used a Bayesian processor for source localization which incorporates *a priori* internal wave variability. The resulting optimal uncertain field processor (OUFP) was applied to narrow-band and broadband signals at very low frequencies (~ 25 Hz) propagating to a long (1000 km) range in deep ocean.^{3,4} Probability of correct range localization and rms error of range estimates were evaluated for both the OUFP and MV processors. The OUFP outperforms the MV processor in general but lags behind the matched case. For both processors, performance degrades significantly as the signal-to-noise ratio decreases.

The study was done assuming an adiabatic normal-mode model.

Matched-mode processing (MMP)⁵ can be used for source localization in an internal waves field by excluding certain modes (mode filtering) which are significantly affected by sound-speed variations.⁶ Matched-mode localization has been successfully applied to broadband data in deep ocean in the presence of internal waves^{7,8} and to acoustic data propagating through an oceanic front.⁹

Internal waves in shallow waters are different from the internal waves in deep ocean. The most prominent difference is the nonlinear solitary internal waves. Solitary internal waves are ubiquitous in the waters of the continental shelves and in the seas near shallow sills. They are usually generated by ocean tides in waters exhibiting strong stratification. The solitary internal waves consist of a series of solitons; each appears like a propagating pulse-like depressions of the thermocline. These waves generally involve large vertical water displacements (~ 10 m) and induce significant perturbations of the acoustic waveguide. In addition to the solitary internal waves, there also exist the random diffused (linear) internal waves of the type described by Garrett and Munk for deep ocean.¹⁰ The random internal waves will be referred to in this paper as the background internal waves. For the properties of internal waves in shallow water, the readers are referred to a recent publication on the initial results from the SWARM 95 experiment.¹¹

Due to the shallow-water depth in coastal waters, the solitary and background internal waves affect signal propagation more drastically in shallow water than in deep water. The characteristics of the acoustic signals in shallow water are different from those in deep water in several aspects:

- (i) Mode coupling.^{12–18} The most dramatic effect induced by internal waves is perhaps the mode coupling between the propagating normal modes. Mode cou-

pling has been studied for narrow-band signals and is responsible for specular transmission loss variation with respect to frequency,^{12,13} and for scintillation index increasing with range.^{14,15} Mode coupling has also been studied for broadband signals propagating through solitary internal waves.^{12–14,16–18}

- (ii) Clustered model arrivals. In deep water, low-frequency signals propagate to a long range. The low-order modes can often be identified and separated in the time series due to different mode travel time. Propagation in shallow water is limited in range (tens of kilometers), and many modes arrive with very little time difference (see illustrations below). Consequently, mode separation often requires a vertical array. Individual modes may not be resolved (mode leakage problem) depending on the array aperture and mode penetration into the bottom.
- (iii) Auto-correlation. Auto-correlation is a good measure of the matched-field performance in a time-varying field. Using one (e.g., the first) of the data fields as the replica field, matched-field correlation becomes the auto-correlation for a vertical array. We note that internal waves are known to have a correlation time on the order of minutes to tens of minutes—the SWARM 95 data indicated that the auto-correlation of the low-order modes^{17,18} and the auto-correlation time of the vertical fields^{19,20} drop to less than 0.6 in 2 min. The auto-correlation time is an indication of how long this replica field is good.

The consequence of the above signal properties is that for a high percentage of data, MFP fails to localize the source correctly in shallow water. Based on the simulation study below, the percentage of incorrect source localization is $\geq 70\%$. This result is to be contrasted with the deep-water results (at similar frequencies) where the degradation of MFP takes the form of moderate fragmentation and wandering of the peaks in the range-depth ambiguity surface.¹

To improve source localization in shallow water in the presence of internal waves, the concept of model filtering can be used as applied previously to the deep-water case. However, there is a significant difference in the use of the mode filter. In deep water, only a couple of (low-order) modes are significantly affected by the internal waves. Source localization can be improved by excluding these modes.^{6–9} In shallow water, the majority of modes is affected by the internal waves due to the “strong” mode coupling among all the modes. To improve source localization, we find that the coupled normal modes need to be excluded. However, the exclusion of the coupled modes cannot be done with a simple mode filter which suppresses a group of modes from the rest of the modes.

The filtering of the coupled normal modes is carried out in this paper using broadband matched-beam processing (MBP) on a vertical array.²¹ Matched-beam processing has been previously applied to vertical arrays in different applications.^{21–24} It was found that beam filtering is easier to implement than mode filtering. It was noted that beam data

are often the only data readily available from existing systems.²⁵

Note that each mode arrives with specific angles at the vertical array; thus, beam filtering is in principle equivalent to mode filtering. Note further that coupled and uncoupled modes travel with different velocities; thus, the coupled modes can be discriminated from the uncoupled modes using a beam-time analysis.^{21,22} It is found that this beam-time analysis is rather effective in suppressing the coupled-mode contribution using broadband matched-beam processing. An alternative is to extract the individual modes from vertical array data using a mode decomposition algorithm⁶ and then separate the coupled modes from the uncoupled modes based on the travel-time differences. We find that the beam domain processor is much simpler than the mode decomposition method. However, a beam filter is not a substitute for the mode decomposition operator when individual mode amplitudes are required.

It is demonstrated below that broadband source localization in the presence of shallow-water internal waves is significantly improved using MBP with a beam filter.

This paper is organized as follows. The acoustic environment used in the simulations is described in the next section. This section describes the simulation of sound-speed variations generated by the Garrett–Munk internal waves and solitary internal waves as a function of time and range (along the propagation path), with the mathematical details left to the Appendix. Section II studies the mode-coupling effect for a broadband signal propagating through this internal wave field. The signal has a center frequency at 400 Hz with a bandwidth of 100 Hz, as was done in the SWARM 95 experiment. In Sec. III, a beam-time analysis of the broadband signals is given and matched-beam processing is reviewed. The source localization results are presented in Sec. IV for MFP and MBP. Section V contains a summary and some conclusions.

I. THE ACOUSTIC ENVIRONMENT

For the simulation of broadband signal propagations, we assume a range-independent bathymetry with a water column 100 m deep, and a range-independent deterministic sound-speed profile representing the background or mean sound-speed profile. This sound-speed profile will be perturbed when internal waves are introduced. The modeling of the diffused (background) internal waves and the solitary internal waves follows closely the work of Tielburger *et al.*,¹⁵ except that the amplitudes of the internal waves have been adjusted to agree more closely with the SWARM 95 measurements. We have also extended the acoustic propagation modeling to broadband pulses.

The deterministic waveguide is used as a reference to illustrate the normal model properties of the signal in the absence of internal waves. The sound-speed profile is shown in Fig. 1(a). This sound speed profile is similar to that used in Ref. 15. The sound speeds in the upper water column are modified to agree better with the actual sound-speed profile at the source location in the SWARM 95 experiment. The corresponding (Brunt–Vaisala) buoyance frequency is shown in Fig. 1(b) as a function of depth. The bottom is

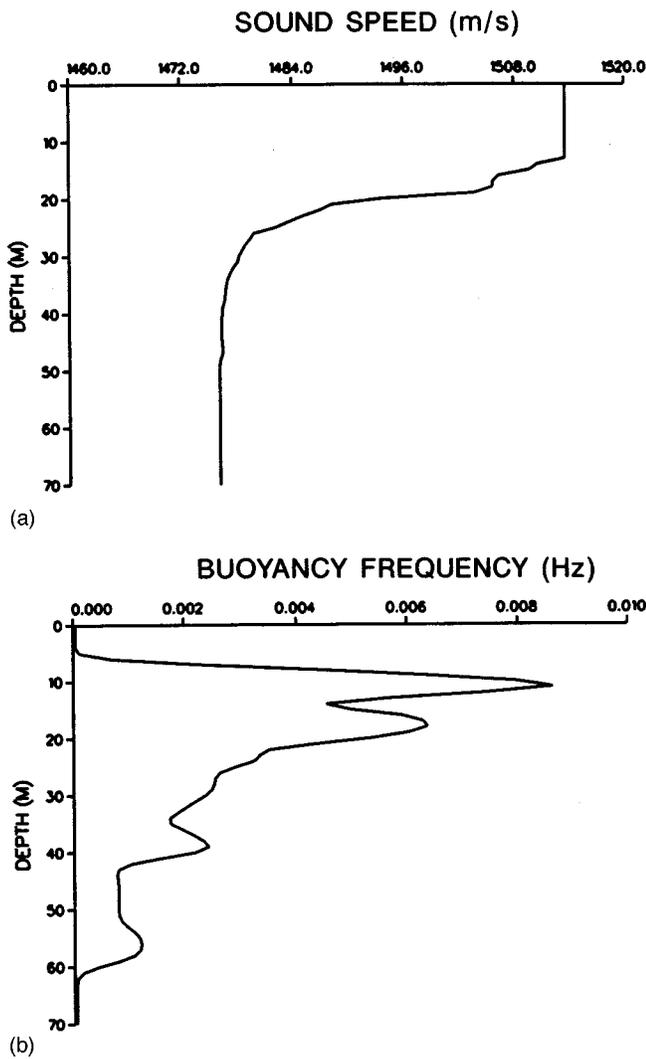


FIG. 1. The sound-speed (a) and buoyancy frequency (b) profile used for the simulation study.

assumed to be a simple hard bottom with a constant sound speed of 1600 m/s. The bottom has a density of 1.5 g/cm^3 and an attenuation coefficient of 0.02 dB/(km Hz) . This environment will be used to calculate the replica field in Sec. IV.

Next, random sound-speed fluctuations due to internal waves are introduced. We use the Garrett–Munk model for the background internal waves and a displacement model for the solitary internal waves. We are aware that the Garrett–Munk model is developed for internal waves in deep water. In shallow water, the forcing (the generation of internal waves) and the oceanography may be very different from that in deep water. Consequently, there is no *a priori* reason to apply the Garrett–Munk model to the shallow-water data. Specifically, the vertical shear, bottom topography, internal tides, surface forcing, mesoscale activity, and coastal edge waves could all affect the internal waves in shallow water. These processes are complicated and not quite understood. For the same reasons, the textbook description of the solitary internal waves,²⁶ using a hyperbolic secant function, may not be correct for solitons in shallow water; other types of solitons have been discussed in the literature.²⁷ In the absence of

a realistic shallow-water internal wave model supported by at-sea data,²⁸ we shall use the “canonical” model as given. This should be adequate for the purpose of simulating the effect of internal waves on sound propagation.

We could use the simulated internal wave fields as given in Ref. 15. But, we shall attempt an “improved” model obtained by constraining the parameters of the model to the oceanographic data taken during the SWARM 95 experiment. We shall use conductivity–temperature–depth (CTD) data at a fixed station to constrain our model.

Why don’t we use the SWARM 95 oceanographic data (as collected) for our simulations? Note that the data were snapshots of the acoustic environments at a particular time and location. Since the background internal wave fields are random in space and fluctuating in time, a snapshot does not represent the total picture and can be misleading in some cases. Also, note that our objective is to investigate internal wave-induced mode-coupling effects on source localization. A range-dependent bathymetry, like the SWARM 95 bathymetry, introduces its own mode coupling and will confuse the issue. For that reason, we choose a range-independent bathymetry and a range-independent *background* sound-speed profile. The sound-speed profile is range dependent when internal waves are present.

In choosing what data to constrain our model, we had several choices, such as the displacement data of the isopycnal layer (acoustic backscattering data), temperature variations at fixed depths (thermistor chain data), velocity profiles as a function of depth (acoustic Doppler current profiler, ADCP data), and CTD data from a fixed or a towed platform; the latter takes samples at a different range and depth each time. We note that the SWARM 95 environment has a range-dependent oceanography; the isopycnal layer depth, the mean (time-averaged) sound-speed profile, etc are location dependent. With sparse data samples, the task of separating internal wave-induced sound-speed variations from that of range-dependent oceanography can be ambiguous. We choose the CTD data at a fixed station because the data are densely sampled in time and depth. For a fixed station, the CTD data have a large number of samples for a statistical measure of the sound-speed variance.

The sound-speed variance is deduced from the CTD data as a function of depth. This variance is used to estimate the amplitudes of the background and solitary internal waves. This method of studying the internal waves is new.²⁸ Note that sound-speed variance is a measure of the sound-speed perturbation and is well determined experimentally. The advantages of using sound-speed variance were discussed in Ref. 28. Figure 2(a) shows the standard deviation (square root) of the sound-speed variance as a function of depth, measured during the SWARM 95 experiment. The variance was measured at close depth intervals (0.5–1 m) taken from data covering a 2-h period.

The details of the internal waves models are given in the Appendix. For the background internal waves, it should be noted that the mode depth functions are uniquely determined by the buoyancy profile shown in Fig. 2(b). The internal wave frequency spectrum is as given by Garret–Munk. The internal wave energy density is unknown and will be fitted

Sound Speed Standard Deviation (m/s)

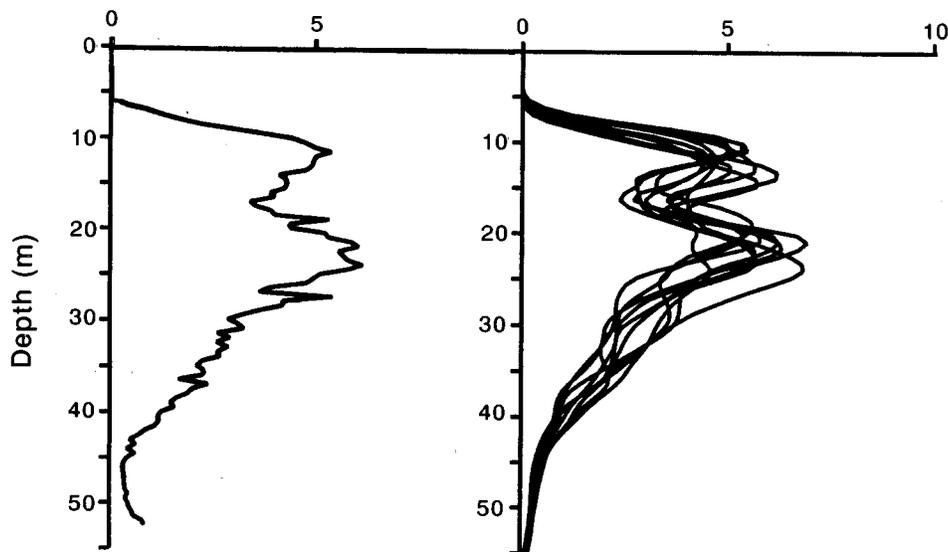


FIG. 2. Standard deviation of sound-speed variations as a function of depth determined from the SWARM 95 CTD data (left figure) and from the ten realizations of the internal wave field (right figure).

with data. For the solitary internal waves, we assume a packet of six solitons with the soliton amplitude, width, velocity, and spacing estimated previously from the temperature sensor data.

We note that sound-speed variance measured from data often involves both solitons and background internal waves. To compare with data, we will simulate sound-speed profiles in the presence of both background and solitary internal waves, and estimate the sound-speed variance from the simulated data. The sound-speed variance for the simulation data is estimated the same way as the real sound-speed data are processed. For the simulations, ten realizations of the internal waves fields were generated as a function of range and depth. (Examples of realization of the sound-speed perturbations are shown below when the impacts of the sound-speed variations on acoustic propagation are studied.) For each realization, the rms value of the sound-speed variation over range is determined for each depth. The standard deviations (square root of the rms variations) for the ten realizations are plotted in Fig. 2(b). The average of the standard deviations is obtained and compared with the SWARM 95 data. To fit the data, we ended up with an internal wave model characterized by its energy density $E_0 = 92 \text{ J/m}^2$ for the background internal waves, and an amplitude $\Lambda_1 = 13.5 \text{ m}$ for the solitary internal waves.

II. BROADBAND MODAL ANALYSIS

Propagation of a broadband pulse is modeled in this section using a full-field propagation model, with and without the internal waves. This section analyzes the mode-coupling effect on pulse propagation induced by internal waves.

The transmitting signal has the same bandwidth and characteristics as that used in the SWARM 95 experiment. Specifically, we use a pulse based on the matched filter output of the 400-Hz M -sequence signal used in the experiment. The pulse consists basically of seven cycles of 400-Hz sinu-

soids with a tapered amplitude. It has a spectrum width of $\sim 100 \text{ Hz}$; the width is 80 Hz at the -8-dB point. We put the source at 50 m depth.

For the receivers, we use a vertical array placed 5 km from the source. The vertical array has 61 phones, spaced at 1.5 m extending from 5 to 95 m in the water column. The signals received on the vertical array are calculated by propagating the broadband signal using a frequency-domain propagation code; 128 frequencies were used with a frequency bin of 1.56 Hz yielding a time series of 0.64 s . Broadband mode decomposition is applied to the received signals. Mode decomposition is based on the mode depth functions as determined from the background (mean) sound-speed profile; this normal mode field represents, in a statistical sense, a mean field in a random medium.

Mode amplitude time series are first calculated from the acoustic field at the receiver array in the absence of internal waves. The received pulses are decomposed into normal mode components as a function of time. In the presence of internal waves, acoustic signals are generated at the receiver array for each realization of the internal wave field. The same mode decomposition procedure is applied to the received signals for each realization. Comparing the arrivals of individual modes with and without internal waves, one identifies the coupled modes due to internal waves. The arrival structure of the coupled and uncoupled normal modes as a function of the arrival time and arrival angle will be explored in the next section using a beam-time analysis to improve source localization. The details of mode analysis are presented below.

A. No internal waves

Figure 3 shows the signal waveforms received on the vertical array at selected depths for the background sound-speed profile. The time series are obtained using the KRAKEN normal mode code.²⁹ One can identify six pulse-like arrivals in the time series at mid-depth receivers, which will be

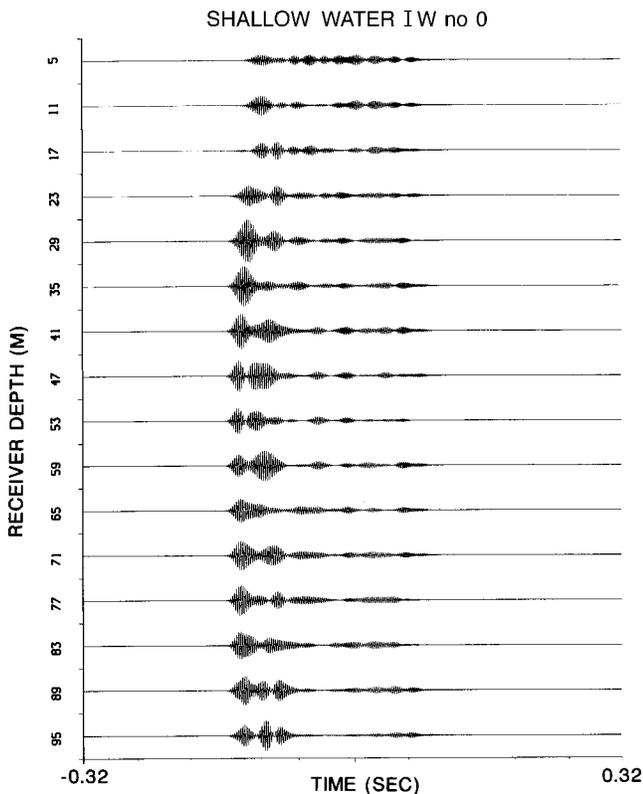


FIG. 3. Broadband time series received on a vertical array at 5 km from the source; no internal waves are present. Receiver depth is shown on the left. Amplitude shown on a relative scale.

named $M1, M2, \dots, M6$. At first glance, one is tempted to associate $M1$ to $M6$ arrivals with the 1st to the 6th normal modes. This is in fact not correct. Correct analysis requires broadband mode decomposition.

To identify the individual normal-mode arrivals for a broadband signal, the received signals on the vertical array (Fig. 3) are first Fourier transformed to obtain the pressure field $p(z_j, f)$ at frequency f for a hydrophone at depth z_j . For each frequency, the mode amplitude $a_i(f)$ for the i th mode is obtained by mode decomposition

$$a_i(f) = \sum_j U(z_j, f) p(z_j, f),$$

where $U(z_j, f)$ is the mode-depth function. The mode-depth functions are calculated using the KRAKEN normal mode code. The mode amplitude time series is obtained by inverse Fourier transform

$$a_i(t) = \int a_i(f) e^{i2\pi ft} df.$$

A fully populated receiver array is used to avoid any mode leakage problem in mode decomposition.

Figure 4 plots the mode amplitude time series for the first ten modes deduced from the time series in Fig. 3. All 61 phone data were used to minimize the mode leakage problem. [There is a small time offset between Figs. 3 and 4 due to the fast Fourier transform (FFT) routine.] We find that the $M1$ arrival can be identified with the 1st and 2nd modes and the $M2$ arrivals can be identified with the 4th and 5th arriv-

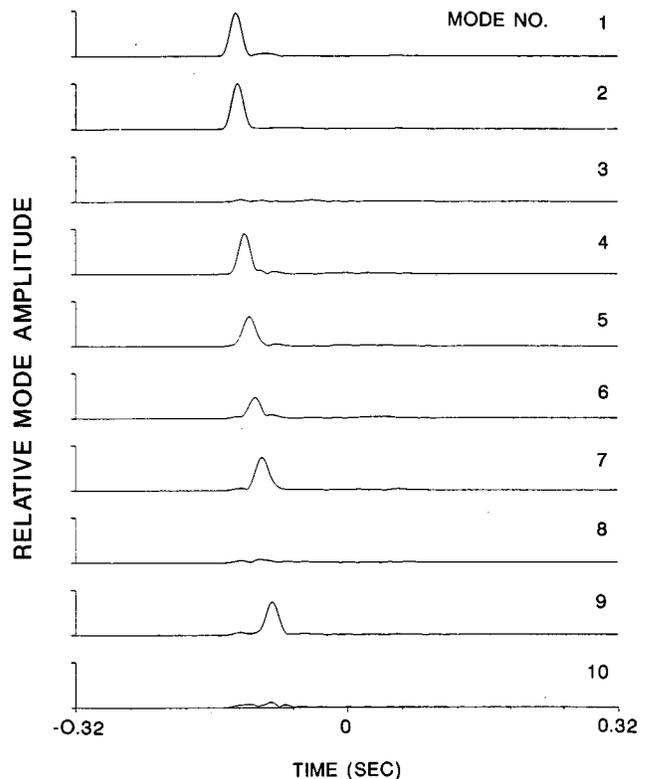


FIG. 4. Mode amplitude time series for the first ten modes deduced from the time series data in Fig. 3; no internal waves. Amplitude shown on a relative scale.

als, etc. This shows that the modes arrive with very small time separations (~ 10 ms) at a range of 5 km.

With this broadband mode decomposition algorithm, we compare mode time series using acoustic fields at the receiver array generated by several propagation codes: FEPE,³⁰ RAM,³¹ UMPE (an improved version),³² and KRAKEN. We find good agreements (with no noticeable differences between the mode time series results by visual comparison) for the low-order (e.g., ten) modes for the given environment. This indicates that the modeling results are reasonable.

In anticipation of the distortion of the mode amplitude time series in the presence of internal waves, we shall define what is meant by coupled and uncoupled normal modes. In a range-independent and stationary waveguide, the normal mode arrival time is given by

$$t_i = \frac{dk_i}{d\omega} r,$$

where $dk_i/d\omega$ is the inverse of the group velocity. Modes arriving with this time are called uncoupled modes in this paper.

In the presence of a discrete scatterer at range r_1 , a j th mode may be converted to an i th mode by the scatterer. (Mode conversion is described by a mode-coupling matrix; hence, mode conversion is also called mode coupling.) The converted mode will arrive at a time

$$t_{ji} = \frac{dk_j}{d\omega} r_1 + \frac{dk_i}{d\omega} r_2,$$

SWARM IT7 IW gmsol

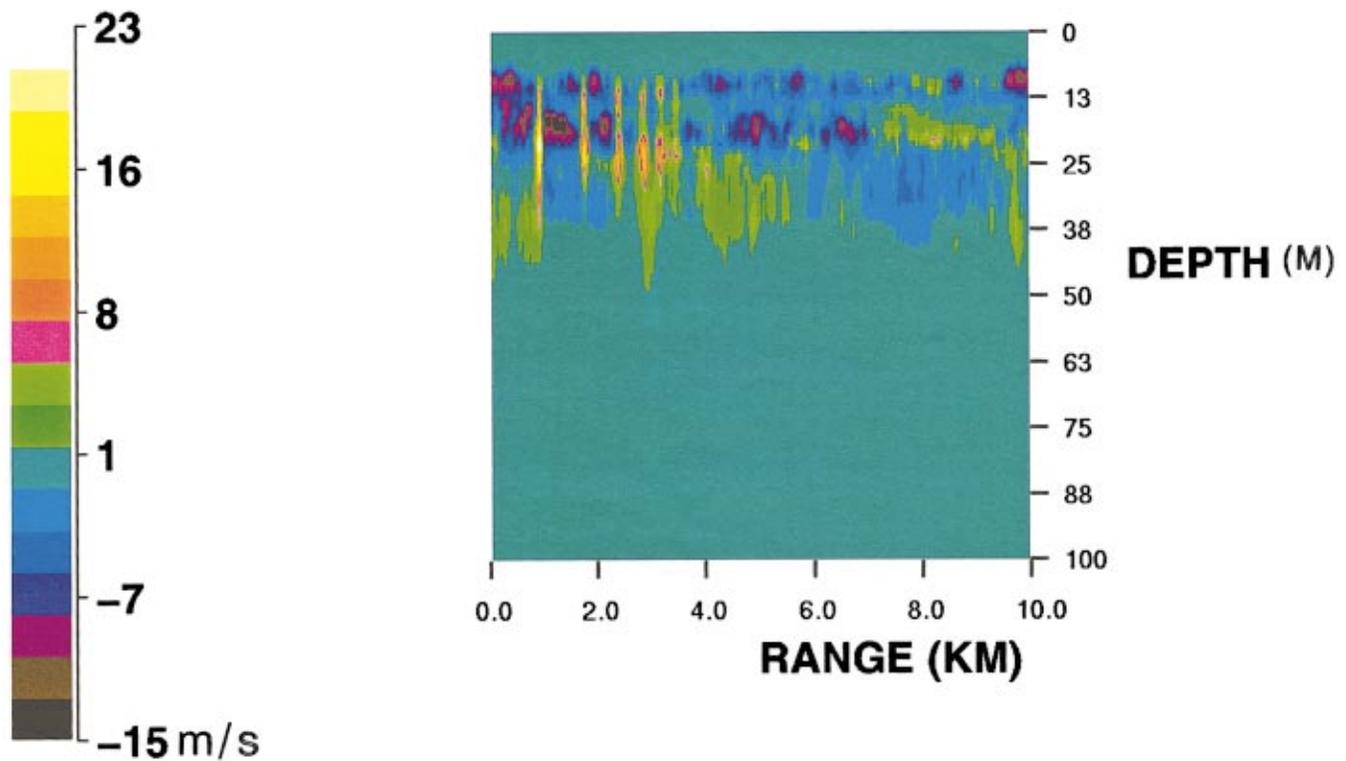


FIG. 5. A realization of sound-speed perturbations due to Garrett–Munk and solitary internal waves.

where $r_2 = r - r_1$. This mode will show up in the i th mode amplitude time series but it will arrive at a different time than the original i th mode. This mode will be referred to as the coupled mode. In the case of internal waves, mode coupling can happen repeatedly at various locations. With respect to time, mode coupling due to the random (background) internal waves is a random process. This makes it difficult to predict the coupled mode arrival time on an event-by-event (ping-by-ping) basis. For the discussions below, peaks in the mode amplitude time series which do not coincide with the original mode arrival time will be categorically called coupled normal modes.

B. With internal waves

In the presence of internal waves, the mode amplitude time series will differ from the previous case of no internal waves. Since internal wave fields are time varying, we shall study the mode amplitudes at a fixed time, corresponding to a realization of the internal wave field. Ten realizations of the internal waves are generated according to the internal wave model described above which include the random Garrett–Munk internal waves and a series of solitons traveling with a fixed speed. A realization is shown in Fig. 5 displaying the corresponding sound-speed profile perturba-

tion by the internal waves. The source is at the origin and the receiver is at a range of 5 km. The solitons are between the source and receivers. For the acoustic propagation, the sound-speed perturbation is frozen for a realization.

Broadband propagation is modeled using the UMPE, RAM/FEPE models for this realization of the sound-speed profile (Fig. 5). The raw time series at the receiver array are plotted in Fig. 6 (using UMPE) for a realization of internal waves field as shown in Fig. 5. Comparing Fig. 6 with Fig. 3, one notes that although the general arrival patterns look similar, there exists substantial differences between the two plots. This difference is better illustrated by the mode amplitude time series. Applying the same mode decomposition algorithm used before to the data in Fig. 6, the resulting mode amplitude time series are plotted in Fig. 7 for the first ten modes. For comparison, we also plot in Fig. 7 the mode time series using pressure fields generated by RAM/FEPE. We find a negligible difference (<1 dB) in the results. Encouraged by this result, UMPE will be used for the rest of the realizations (so that the broadband simulations can be completed in a reasonable time frame). We note that modeling broadband propagation in internal wave fields is computationally intensive. A fine range/depth step commensurate with the sound-speed perturbation is required for each frequency component. (This applies to RAM as well.)

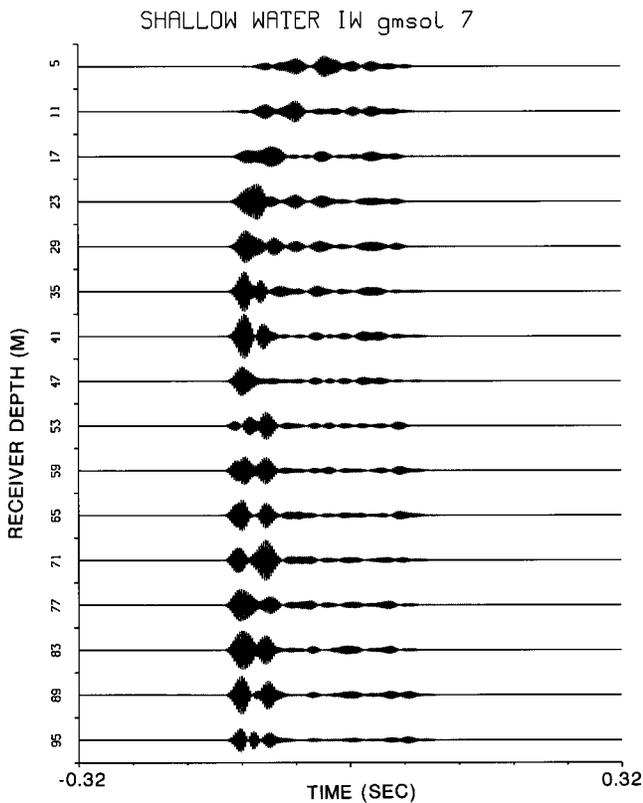


FIG. 6. Broadband time series received on a vertical array at 5 km from the source in the presence of internal waves as shown in Fig. 5. Receiver depth is shown on the left. Amplitude shown on a relative scale.

A comment is in order here. As the title of this paper implies, the purpose of this paper is to demonstrate a methodology for source localization in the presence of internal waves. Since this paper uses only simulated data, the accuracy of the propagation model is not a critical issue (for the purpose of demonstrating the source localization methodology) as the accuracy applies to both the simulated data and simulated replica fields. While UMPE may not be a perfect model for coupled-mode environments, we do find that the mode amplitude time series agree with results using other models. The reason may be due to the fact that the internal waves are confined predominately to the upper water column and hence involve mostly small-angle propagation. We note that the above time-domain comparison of different model predictions may be more “forgiving” than the transmission loss calculations for continuous waves. As such, the above numerical results may not constitute a true test of the model accuracy.

Henceforth, we shall concentrate on the study of mode coupling and its effect on source localization, using the (improved) UMPE model. Comparing Fig. 7 with Fig. 4, we note that

- (i) Mode 2 and 7 amplitudes remain almost unchanged.
- (ii) Mode 1, 4, 5, 6, and 9 amplitudes decrease significantly, indicating that energy has been transferred to other modes.
- (iii) Mode 3, 8, and 10 amplitudes are substantially higher in Fig. 7 than in Fig. 4. The increased amplitudes are due to the internal waves.

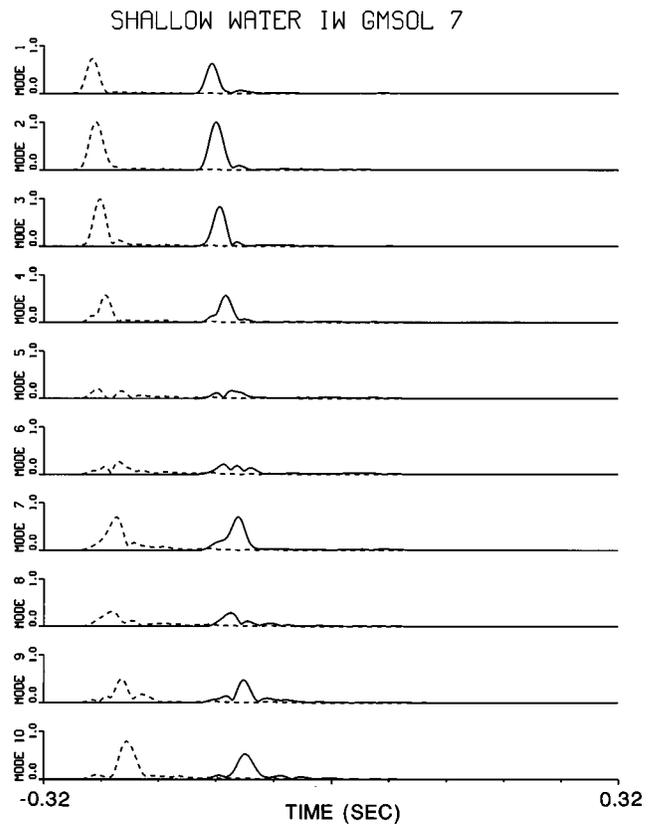


FIG. 7. Mode amplitude time series for the first ten modes in the presence of internal waves based on time series data of Fig. 7. The solid lines represent the UMPE results. The dashed lines represent the RAM/FEPE results, displaced with an offset for comparison purposes. Amplitude shown on a relative scale.

- (iv) For practically all ten modes, one observes multiple peaks in the time series. As discussed above, multiple peaks are associated with the coupled modes; without mode coupling, one peak is associated with each mode. Note that the difference in the arrival time between the peaks is small; hence, coupled modes are difficult to separate from the uncoupled modes based on the travel-time difference alone.

Next, we perform broadband modal analysis for all ten realizations. We also study the mode arrival structure with only the background internal waves, and with only the solitary internal waves; ten realizations for each case. We also vary the amplitudes of the internal waves to study how the mode-coupling effect changes. We find that the coupled-mode amplitudes are different from one realization to the other. Because of the stochastic nature of the background internal waves, mode coupling is random even with solitons present. There is no definitive pattern how modes are coupled for a given random realization. The only general conclusion is that mode coupling affects all the modes.

We note that mode-coupling strength is equal from low-to-high and from high-to-low-order modes. At short ranges, mode energy may be converted from high-to-low-order modes or vice versa, depending on the mode excitation by the source. At long ranges, since high-order modes attenuate faster than the low-order modes, there will eventually be more energy in the low-order modes than in high-order

modes. The net effect of mode coupling at long ranges is to constantly transfer energy from the low-order modes to the high-order modes. As a consequence, the low-order mode amplitudes decrease faster with range in the presence of internal waves than in the absence of internal waves. Conversely, the high-order-mode amplitudes decrease more slowly with range in the presence of internal waves than in the absence of internal waves.

III. SOURCE LOCALIZATION: DIFFERENT APPROACHES

In this section, we investigate the effects of coupled modes on the performance of matched field/mode processing in an internal wave field. Matched-beam processing will be studied next. A beam-time analysis will be used to illustrate a mode-filtering methodology for suppressing the coupled modes.

A. Matched field/mode processing

Matched field/beam processing algorithms are reviewed first. The implication of the SWARM 95 correlation data for matched-field processing will be studied next.

Broadband MFP/MMP processing can be done by coherent or noncoherent summation of the narrow-band range–depth ambiguity surfaces.^{33,34} Coherent frequency domain MFP is related to the time domain MFP by

$$B_{MF}(r, z) = \frac{[\sum_{j=1}^N \int p_j^{\text{rplc}}(t) p_j^{\text{data}}(t) dt]}{[\sum_{j=1}^N \int |p_j^{\text{rplc}}(t)|^2 dt][\sum_{j=1}^N \int |p_j^{\text{data}}(t)|^2 dt]}, \quad (1a)$$

$$= \frac{|\int \sum_{j=1}^N p_j^{\text{rplc}}(f) * p_j^{\text{data}}(f) df|^2}{[\sum_{j=1}^N \int |p_j^{\text{rplc}}(f)|^2 df][\sum_{j=1}^N \int |p_j^{\text{data}}(f)|^2 df]}, \quad (1b)$$

where j is the hydrophone index, and r, z represent, respectively, the range and depth of the replica source; the r, z dependence of the replica field is not explicitly shown. Coherent frequency domain MFP requires time alignment between the replica and data field. If the spectral correlation is squared before it is integrated over frequency in (the numerator of) the above equation, one has the noncoherent MFP.

Equation (1) can be expressed in the mode space as

$$B_{MM}(r, z) = \frac{|\int \sum_{i=1}^M a_i^{\text{rplc}}(f) * a_i^{\text{data}}(f) df|^2}{[\sum_{i=1}^M \int |a_i^{\text{rplc}}(f)|^2 df][\sum_{i=1}^M \int |a_i^{\text{data}}(f)|^2 df]}, \quad (2)$$

where the $a_i(f)$'s are the i th mode amplitudes which are related to the pressure field by

$$p_j(f) = \sum_{i=1}^M a_i^{\text{rplc}}(f) U(z_j, f). \quad (3)$$

For simplicity, we have assumed that the mode-depth functions are orthogonal functions in obtaining Eq. (2) from Eq. (1). Equation (2) can be expressed in the time domain as

$$B_{MM}(r, z) = \frac{|\int \sum_{i=1}^M a_i^{\text{rplc}}(t) a_i^{\text{data}}(t) dt|^2}{[\sum_{i=1}^M \int |a_i^{\text{rplc}}(t)|^2 dt][\sum_{i=1}^M \int |a_i^{\text{data}}(t)|^2 dt]}, \quad (4)$$

where $a_i(t)$'s are the mode amplitude time series investigated in the previous section.

Equipped with the above equations, we can now discuss the effect of internal waves on source localization. We shall start with the matched-mode expression, Eq. (4). We note that range estimation is based predominantly on the phase of the mode amplitude.⁵ In the time domain, range estimation will be based on the (relative) travel time of the (uncoupled) normal modes. Now, if the mode amplitude time series contain coupled modes, it should be no surprise that source range estimation can be wrong.

The effect of the coupled modes on source localization can be estimated from the SWARM 95 data. In the SWARM 95 experiment, the M -sequence at 400 Hz is repeated 23 times, each with a duration of 5.11 s. Mode amplitudes have been deduced from the M -sequence data by Headrick *et al.* in Refs. 17 and 18. Let the replica field in Eq. (4) be the first M -sequence data received, the matched-mode correlation is related to the mode coherence function by

$$\rho_i(t_n) = \frac{|\int a_i^{\text{data}(t_1)}(t') a_i^{\text{data}(t_n)}(t') dt'|}{[\int |a_i^{\text{data}(t_1)}(t')|^2 dt']^{1/2} [\int |a_i^{\text{data}(t_n)}(t')|^2 dt']^{1/2}}. \quad (5)$$

Based on the SWARM 95 data, it was found^{17,18} that the modal coherence drops to <0.6 in ≤ 2 min for mode 1 to 4. One expects that the modal coherence for the higher-order modes will be less. The drop in the modal coherence was apparently due to the mode-coupling effect. The SWARM 95 coherence data suggest that a fixed (time-independent) replica will be good only for a short time (<2 min); it would not be able to localize the source at a later time.

We can also estimate the effect of mode coupling on source localization using MFP. A correlation function is defined for a vertical array following the concept of matched-field correlation, Eq. (1), by using the first data field as the reference field

$$\rho_{MF}(t_n) = \frac{[\sum_{j=1}^N \int p_j^{\text{data}(t_1)}(t') p_j^{\text{data}(t_n)}(t') dt']}{[\sum_{j=1}^N \int |p_j^{\text{data}(t_1)}(t')|^2 dt']^{1/2} [\sum_{j=1}^N \int |p_j^{\text{data}(t_n)}(t')|^2 dt']^{1/2}}. \quad (6)$$

The SWARM 95 data at 400 Hz showed that the MFP coherence drops to ≤ 0.5 in <2 min.¹⁹ This implies that a rep-

lica field is, at best, good for <2 min. In other words, the replica field needs to be updated (rapidly) with time for MFP

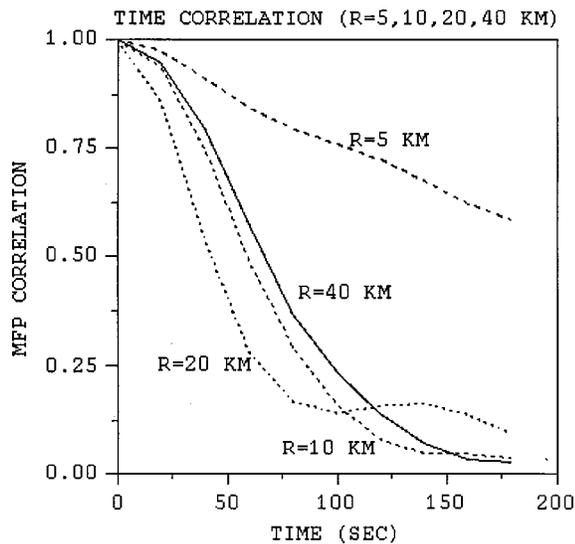


FIG. 8. Matched-field correlation as a function of elapsed time between the data field and the initial field which is used as the replica field.

source localization in the internal wave fields.

Using the internal wave model described above, we simulate the matched-field correlation as a function of time. This simulation starts with one realization of the internal wave field and lets the internal waves evolve with time by the equation of motion (in this case, by the spectral representation). For each elapsed time, we propagate the broadband pulse through the internal wave field, treating it as frozen. The received time series is correlated with the time series at the initial time ($t=0$). Figure 8 shows the MFP correlation as a function of the elapsed time. For the simulation, the internal wave field includes both the Garrett–Munk and solitary internal waves; the solitary internal waves move from the receiver array toward the source. We find that the correlation time at a range of 40 km is less than 2 min, which is consistent with the SWARM 95 finding.¹⁹ For the case of 5 km, the correlation time is several minutes before it drops below 0.5.

The above study indicated that the MFP will have difficulty localizing the source if a fixed replica field is used for all data. As the matched-field correlation decreases, not only does the signal gain degrade significantly but the processor will fail to localize the source correctly. Numerical simulations in the next section support this finding.

As the performance degradation is due to the coupled modes, naturally one is motivated to use only the uncoupled modes to localize the source. The coupled modes can be separated from the uncoupled modes traditionally by two methods: time separation and mode filtering. As the coupled modes arrive at a different time than the uncoupled modes, they can be separated if sufficient travel time difference exists between them. This method works well in the deep ocean at long ranges. For shallow water, as indicated above, the time difference can be shorter than the pulse length, and this method will fail. Using mode filtering, the method normally works when only a small number of modes are affected. An example is low-frequency propagation in the deep ocean internal wave field.^{7–9} For shallow-water internal waves, we

found that the mode coupling affects practically all modes; hence, a simple mode filter will not work in shallow water. Also, the mode filtering works when the mode leakage problem is minimal, which requires a large aperture vertical array. For shallow water, mode decomposition is often hindered by the array size and the fact that some modes penetrate into the ocean bottom.

B. Beam-time analysis and matched-beam processing

In this section, we show that coupled modes can be suppressed using a beam-time analysis. The beam filter is simpler to implement (than a mode filter) and can be applied when mode coupling affects all modes. To suppress the coupled modes, we shall use a replica field which contains only the uncoupled modes. The replica field is the field calculated without the internal waves.

The beam filter is applied in the context of matched-beam processing. Matched-beam processing (MBP) has been studied in previous publications.^{21–25} Briefly, MBP is MFP performed in the beam domain. MBP employs a similar concept as MMP: the data are transformed from the phone domain to the mode domain or beam domain to permit the “matched-field” correlation to be conducted in a smaller (sub)space of the signal. In the presence of internal waves, the subspace is the uncoupled mode space of the signal. The reduction of the dimensionality of the processing space results in robustness against the undesired signals and noise.

1. Beam-time analysis

Beam-time analysis is used to display the arrival angle of the signal with respect to time. Since each mode is expected to arrive at the vertical array at a specific angle, the angle–time relation differentiates the coupled mode from the uncoupled modes more clearly than using the time or angle information alone.

We show below the angle–time plot of the signals with and without the internal waves. The angle–time plot is obtained by conventional beamforming of consecutive (overlapping) short-time segments of the signals on the vertical array. We note that the angle estimated by conventional beamforming is not exactly the mode arrival angle due to the refractive nature of the sound speed as a function of depth. This does not affect the accuracy of matched-beam processing (see below). Despite the distortion, the beam angle is fairly close to the expected arrival angle.

The process starts with a broadband time series (Fig. 3 or 6). For a given channel, we note that the time series in Fig. 3 or 6 lasts about 0.64 s, which is sampled by 1024 points. This time series is divided into segments of 128 elements using a moving window with 75% overlap. Each segment is padded with zeros and Fourier transformed into the frequency domain with a FFT of 256 points. Beamforming is conducted in the frequency domain. Beam outputs are incoherently summed over the frequency components.

A comparison of the beam angle–time display is shown in Fig. 9 with and without internal waves. For the purpose of illustrating the (mode) arrival angle versus time, only the strongest beams with a 3-dB beamwidth are shown for each time segment; the color intensity indicates the peak level

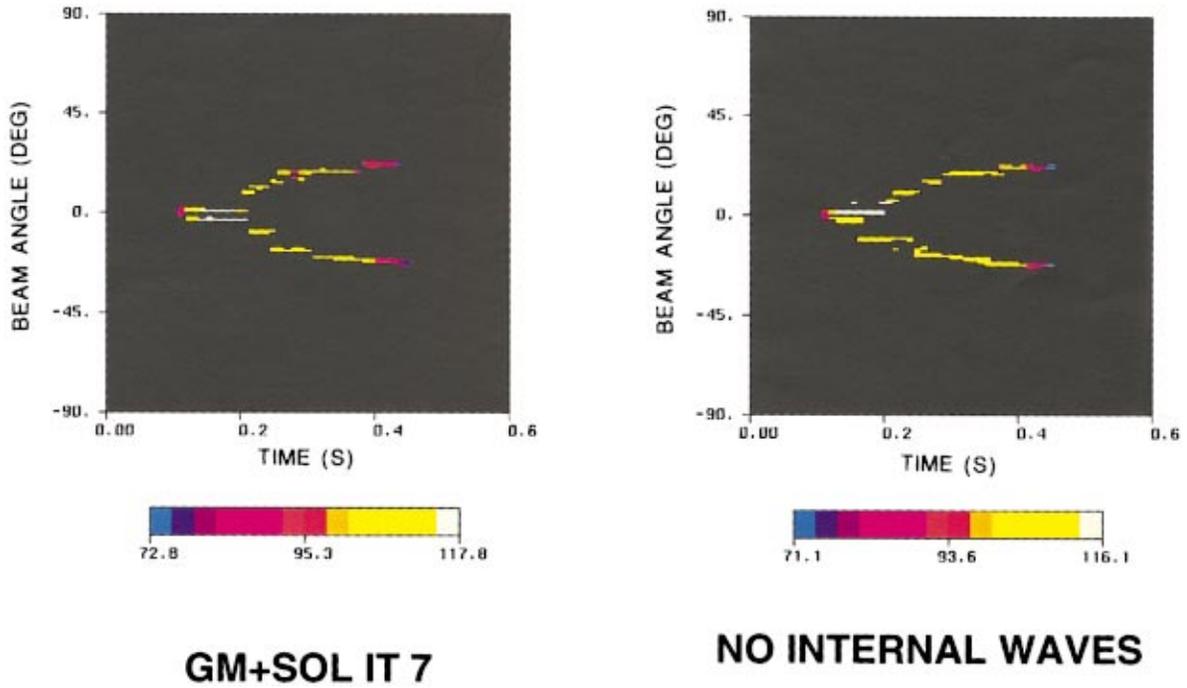


FIG. 9. Beam angle–time traces of the received signal at 400 Hz: (a) no internal waves, (b) with internal waves.

versus time. For simplicity, only one frequency (400 Hz) is used in Fig. 9. The difference in the arrival structure with [Fig. 9(b)] and without [Fig. 9(a)] internal waves is clearly evident. The differences between Fig. 9(a) and (b) are induced by the coupled modes which we have studied in Sec. II. The mode arrival time in Fig. 9(b) can be traced to the mode arrival time in Fig. 7 on a one-to-one basis.

We note that each mode arrives at the vertical array with a symmetric up and down angle. The signal-beam intensity will be symmetric in the up and down angle when the modes are summed incoherently. The signal-beam intensity for a broadband pulse can be unsymmetric due to the mode interference patterns. But, the mode arrival angle and time associated with the uncoupled modes will follow the curve estimated with incoherent (uncoupled) modes. The arrival angle and time of the coupled modes will not follow this curve.

The purpose of the beam angle–time analysis is not to determine precisely the mode arrival angle versus time, which involves the issues of angular resolution, array aperture, and sound-speed variation with depth. The purpose of this analysis is to demonstrate the difference in the arrival patterns and use this difference in MBP. Note that since Fig. 9(a) contains no coupled modes, by overlapping Fig. 9(a) and (b), one identifies the coupled mode. By multiplying the complex beam-time functions, the coupled-mode contributions will be suppressed. This is matched-beam processing. Figure 9(a) is the replica beam which performs the mode-filtering function and source-localization function at the same time.

2. Matched-beam processing

The MBP ambiguity function is computed for each time segment of the time series (denoted by segment number n) as follows:

$$B_{\text{MB},n}(r,z) = \frac{|\int_{\Theta} A^{\text{rplc}_n}(\theta) A^{\text{data}_n}(\theta) d(\sin \theta)|^2}{[\int_{\Theta} |A^{\text{rplc}_n}(\theta)|^2 d(\sin \theta)] [\int_{\Theta} |A^{\text{data}_n}(\theta)|^2 d(\sin \theta)]}, \quad (7)$$

where $A^{\text{data}_n}(\theta)$ and $A^{\text{rplc}_n}(\theta)$ are respectively, the data beams and replica beams associated with the n th segment, and r, z represents the replica source position in range and depth. In the frequency domain, beams are created for each frequency component by conventional beamforming

$$A^{\text{data}_n}(\theta, f) = \sum_j e^{-ikz_j \sin \theta} p_j^{\text{data}_n}(f), \quad (8)$$

$$A^{\text{rplc}_n}(\theta, t) = \sum_j e^{-ikz_j \sin \theta} p_j^{\text{rplc}_n}(t). \quad (9)$$

Equation (7) is calculated for each frequency component. The ambiguity surface is summed (incoherently) over the time segments and then summed over the frequency components. The source spectrum is normalized away using this implementation.

The symbol Θ in Eq. (7) indicates a limit of integration. For example, the integration can be limited to $\theta_1 \leq \theta \leq \theta_2$ where θ_1 and θ_2 are the -3 -dB points of the main peaks (for each segment). Note that although Eqs. (8) and (9) are the conventional beamformer in the frequency domain, no plane wave assumption is made for the signals. Equation (7) is the correlation of the signal in the wave number domain for non-plane wave or plane wave signals alike.

IV. SIMULATION RESULTS

In this section, we simulate the MFP and MBP source-localization results in shallow water in the presence of inter-

nal waves. We start with ten realizations of the internal waves field as mentioned above. The broadband pulse is propagated through each realization of the sound-speed field. The received signals are the data used for MFP and MBP source-localization analysis.

For MFP, it is customary to compute the range–depth ambiguity function in the frequency domain by a Fourier transform of the entire time series data. The MFP ambiguity surfaces are incoherently summed over the frequency components; incoherent frequency averaging is used to suppress sidelobes. For each frequency component, the matched-field ambiguity surface is given by

$$B_{MF}(r, f) = \frac{|\sum_j p_j^{rplc*}(f) p_j^{data}(f)|^2}{[\sum_j |p_j^{rplc}(f)|^2][\sum_j |p_j^{data}(f)|^2]} \quad (10)$$

This procedure will be followed in this section. To save computing time, only eight frequency components are used for

averaging. Preliminary study shows that eight frequencies are adequate for frequency averaging.

MBP is done using the procedure described in the previous section. The ambiguity function is summed over eight frequencies in the same way as in MFP. For MBP, the data are divided into many segments. The data time series is matched filtered with the replica for proper time alignment.

For both processes, the replica field is taken to be the field without the internal waves. In practice, the average sound-speed profile obtained during a period without the solitary internal waves will be used to generate the replica field; the sound-speed variations due to background internal waves will average to zero. (For a range-dependent environment where mode coupling can be induced by bathymetry, the local modes at the receiver, using the mean sound-speed profile, will be used as the reference modes. This subject is beyond the scope of this paper.)

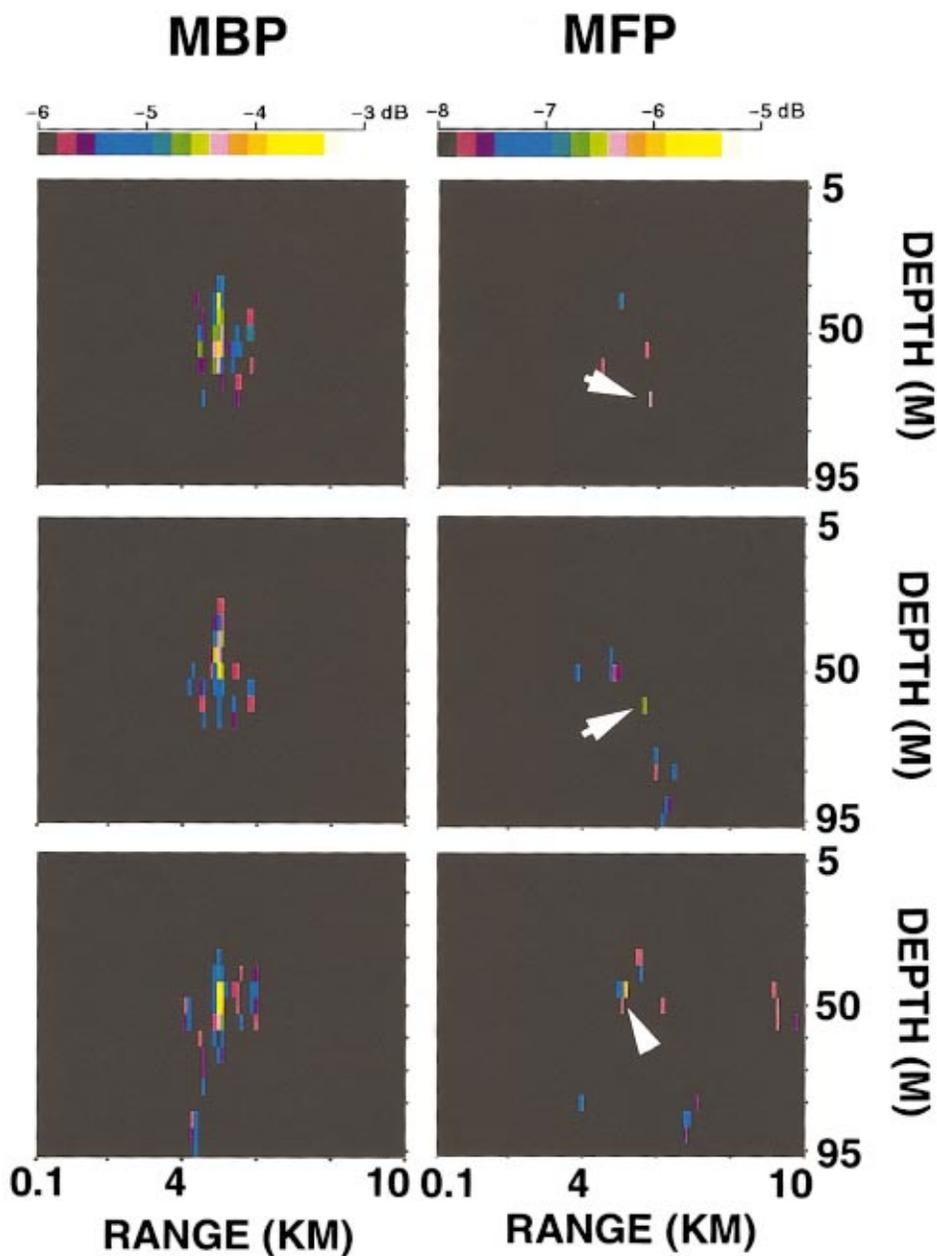


FIG. 10. MFP and MBP range–depth ambiguity surfaces for six realizations of the internal wave field. Arrows are used to help identify the peaks of the MFP ambiguity surfaces.

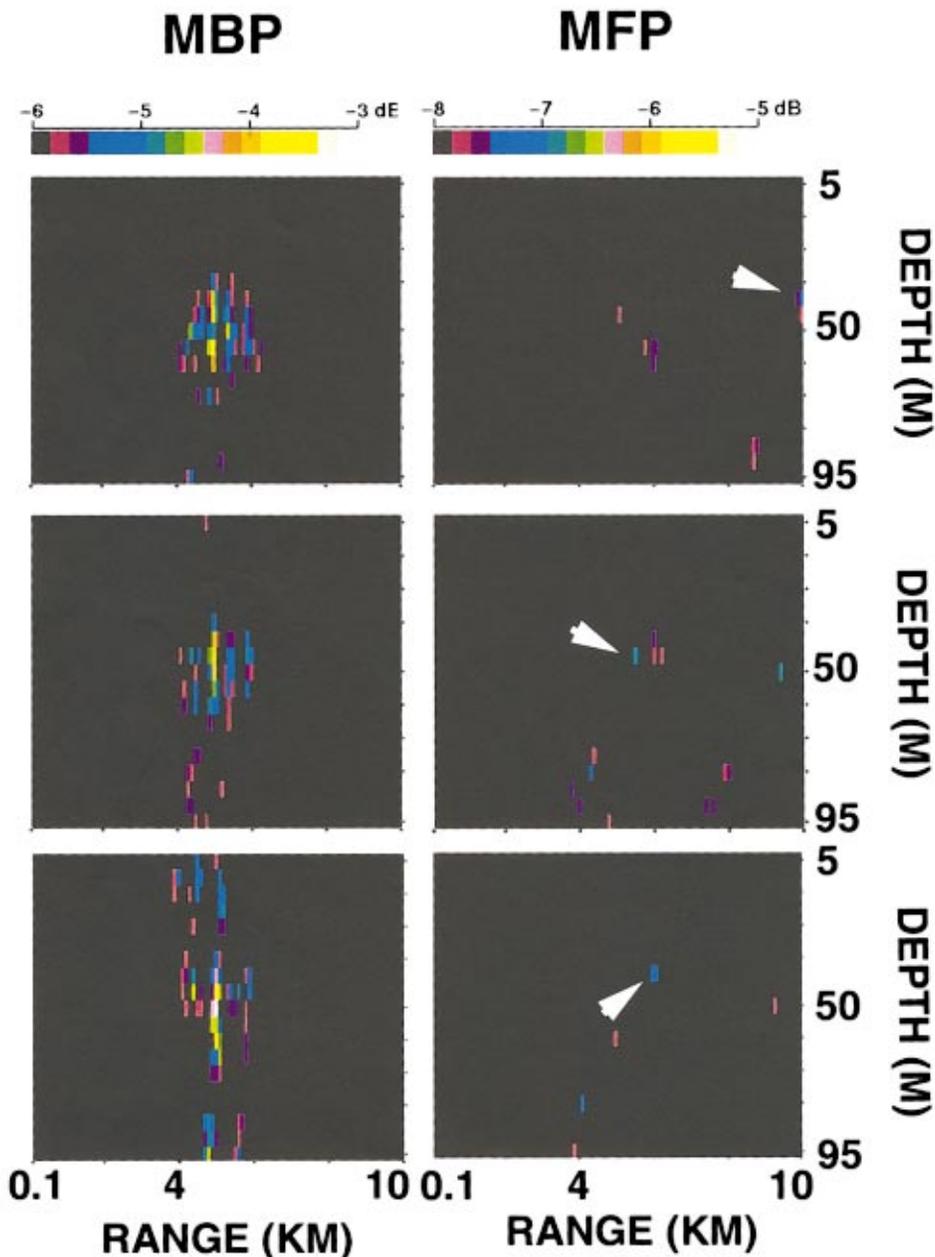


FIG. 10. (Continued.)

A. Simulation results

MFP and MBP ambiguity surfaces are generated for broadband signals propagating through ten realizations of the internal waves field using the procedure described above. Figure 10(a) and (b) shows side-by-side comparison of the MFP and MBP ambiguity surfaces for six realizations. The ambiguity function is calculated for a range spanning from 100 m to 10 km in 100-m steps and a depth covering 5 to 95 m in 5-m steps. As expected, MFP fails to localize the source at the true source location in most cases. The MFP source location appears to be random as the mode-coupling effect due to internal waves seems to be random. On the other hand, MBP consistently localizes the source very close to the true source location. This is evidence that coupled modes have been largely suppressed by this processor.

One notices that the peak level of the MBP surface is consistently ≥ 3 dB higher than the peak level of the MFP

surface for each realization. (The color scales in the figures are adjusted by 3 dB between MFP and MBP.) The peak level of MBP is less than ideal (0 dB) because the incoherent coupled-mode energy is not used. The peak level of MFP is much less than ideal because the distortion (mismatch) of the signal by the internal wave field is not included in the replica field.

Although MBP demonstrated consistent source localization, it has sidelobes close to the main lobe; such sidelobes will generally not be mistaken for a false target. The reason for the sidelobes is the loss of coherent energy (of the uncoupled modes) to the coupled modes.

B. Discussion

To study the effect of internal waves on broadband source localization in shallow water, in addition to the above simulations we have simulated source localizations in the

TABLE I. Summary of matched-field localization results.

	GM+Solitons	GM (92 J/m ²)	GM (154 J/m ²)	Solitons
$\Delta R, \Delta D = 0$	30%	20%	10%	30%
$\Delta R = 0$	30%	50%	20%	50%
Max $\Delta R, \Delta D$	5 km/20 m	0.9 km/30 m	0.8 km/30 m	0.1 km/5 m

presence of Garrett–Munk internal waves only and in the presence of solitary internal waves only. We have also varied the strength of the background internal waves ($E_0 = 92$ and 154 J/m^2)³⁵ to study the effect of different amplitudes on source localization. For each of these cases, we used ten realizations. The results are summarized in Tables I and II for MFP and MBF, respectively. Each table lists the percentage of data with no range/depth error, no range error, and the worst error in range/depth estimation. Comparing Table I with II, we note that MBP performs consistently better than MFP in source localizations.

We note that the effect of internal waves on sound propagation is characterized by two parameters: the rms value of the sound-speed perturbation along the propagation path and the structure or the scale of the sound-speed variation along the propagation path. The rms value of the sound-speed variation is a function of depth for each (realization of) internal wave field. The rms values for the ten realizations were shown earlier in Fig. 2. The horizontal scale of the internal waves field is controlled by the horizontal coherence length, which for linear internal waves field is determined by the (Garrett–Munk) spectrum. We have tried to adjust the rms variation while keeping the horizontal and vertical scale of the internal waves fixed for each realization by renormalizing the calculated rms value to the experimental data. We found no significant changes in the statistics.

V. SUMMARY AND CONCLUSIONS

In summary, this paper addresses the problem of source localization in shallow water in the presence of internal waves. Source localization in shallow water is different from that in deep water in several respects. In shallow water, there are solitary internal waves as well as the background (Garret–Munk) internal waves. In shallow water, significant mode attenuations can result from bottom interactions which cause high-order modes to attenuate faster than the low-order modes. Significant mode couplings are induced by internal waves which affect all modes in shallow water. As a result of the mode coupling and bottom attenuation, acoustic energy is constantly being shifted from low-order modes to higher-order modes after an initial redistribution of mode energy. In deep water, most of the acoustic energy is carried by water-

TABLE II. Summary of matched-beam localization results.

	GM+Solitons	GM (92 J/m ²)	GM (154 J/m ²)	Solitons
$\Delta R, \Delta D = 0$	80%	80%	70%	100%
$\Delta R = 0$	100%	100%	100%	100%
Max $\Delta R, \Delta D$	0 km/5 m	0 km/5 m	0 km/5 m	0 km/0 m

borne modes which have little or equal attenuation. Only a small number of low-order modes are significantly affected by the internal waves.

The effects of internal waves on broadband acoustic signal propagation are illustrated by the presence of coupled modes using a broadband modal analysis. Broadband time series are simulated for a vertical array of receivers by propagating a broadband pulse through the internal wave fields. For the internal wave model, the magnitudes of the background internal waves and solitary internal waves have been set according to the sound-speed variance measured from the SWARM 95 experiment. The presence of the coupled modes in the received signals is indicated by the (different) arrival structure in the mode amplitude time series and is further illustrated using a beam-time analysis.

For performance evaluation of MFP, we calculated the matched-field correlation as a function of time using the initial data field as the replica field. The matched-field correlation decreases with time due to the time-varying nature of the data field. Without internal waves, the correlation is one in theory. With internal waves, the correlation is less than one due to the contribution of the (noncoherent) coupled modes. The simulation result indicates that the matched-field correlation decreases quickly with time as the internal wave fields evolves. The correlation coefficient drops to less than 0.5 in less than 2 min at a range of 40 km, which is consistent with the broadband correlation data centered at 400 Hz from the SWARM 95 experiment. This indicates that a fixed replica field will have difficulty localizing the source in the time-varying internal wave field.

The poor performance of MFP using a fixed replica field is supported by numerical simulations. In our simulations, the acoustic source is incorrectly localized most of time (50%–70%) using MFP. This is in contrast to deep water, where the effect of internal waves on source localization is fragmentation and meandering of the source peak in the ambiguity surface;¹ source is approximately localized.

To improve source localization in the presence of internal waves, a broadband MBP algorithm is used to suppress the coupled normal modes in the data field. The separation of the coupled modes from the uncoupled modes requires a beam angle and time analysis; the coupled modes travel with a different speed than the uncoupled modes for the same arrival angle. Note that the coupled modes generated by the random media are unpredictable and hence useless for source localization. As the coupled-mode energy is not used, signal gain decreases by 1–3 dB. More sidelobes are found near the mainlobe (yielding equivalently a fattened main lobe) due to less coherent modes available for source localization. But, by suppressing the coupled modes, we find that MBP consistently localized the source where the true source was located (with occasional small depth errors).

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research and by a grant of HPC time from the DoD HPC Center, NRL-DC, Origin 2000. We thank S. Finette for sharing his internal wave-generation computer codes with us.

APPENDIX: INTERNAL WAVES MODEL

A. Garrett–Munk internal waves

We use the Garrett–Munk model for the background internal waves. The sound-speed perturbation is related to the displacement of the water mass by

$$\delta C(r, z, t) = \int_z^{z+\eta(r, z, t)} C_0(z) G(z) N^2(z) dz, \quad (\text{A1})$$

where $\eta(r, z, t)$ is the displacement, $C_0(z)$ is the mean sound-speed profile, $N(z)$ is the buoyancy frequency, and $G(z)$ is a function of the derivatives of the salinity and temperature profile as a function of depth. For details, see Ref. 10.

The internal wave displacement can be expressed as a sum of discrete normal modes

$$\eta_D(r, z, t) = \sum_{k_h} \sum_j F(k_h, j) W(k_h, j, z) e^{i[k_h r - \omega(k_h, j)t]}, \quad (\text{A2})$$

where k_h is the horizontal wave number, $\omega(k_h, j)$ is the j th mode frequency, and $W(k_h, j, z)$ is the depth function of the j th mode internal wave which satisfies the following equation:

$$\frac{d^2}{dz^2} W(z) + k_h^2 \left[\frac{N^2(z) - \omega^2}{\omega^2 - f_c^2} \right] W(z) = 0, \quad (\text{A3})$$

where f_c is the Coriolis frequency: $f_c = 2\Omega \sin$ (latitude) with Ω being the earth's angular speed. We chose 40° for the latitude. The depth function $W(z)$ satisfies the boundary conditions $W(0) = W(H) = 0$, where H is the water depth. The boundary conditions determine the frequency dispersion relation of ω as a function of k_h . In Eq. (A2), $F(k_h, j)$ represents the amplitude of the j th mode internal wave which is a zero mean, complex Gaussian random variable with a spectral representation

$$\langle |F(k_h, j)|^2 \rangle = E_0 M [j^2 + j_*^2]^{-p/2} \left(\frac{4}{\pi} \right) k_j k_h^2 (k_h^2 + k_j^2)^{-2}, \quad (\text{A4})$$

where $\langle \dots \rangle$ represents the ensemble average. Equation (A4) is referred to as the Garrett–Munk spectrum. The quantity k_j is given by $k_j = \pi j f_c / \int_0^H N(z) dz$. It corresponds to the spectral peak in the horizontal wave number domain for each mode j . The normalization factor M satisfies $1/M = \sum_{j=1}^{\infty} [j^2 + j_*^2]^{-p/2}$. The parameter E_0 is the average energy density in Joules/m² unit. The characteristic wave number j_* and the spectral power-law exponent p are empirically determined. We use $j_* = 3$ and $p = 4$ as in Ref. 15.

From Eqs. (A1)–(A4), one generates realizations of sound-speed perturbations as a function of range and depth. A realization of the sound-speed perturbation is shown in Fig. 5 where the nonsolitary background sound-speed perturbations are due to the Garrett–Munk internal waves. For the Garrett–Munk internal waves, the sound-speed variance is determined by the internal wave intensity E_0 .

B. Solitary internal waves

The displacement of the solitary internal waves is given by

$$\eta_S(r, z, t) = \sum_{j=1}^{\infty} a_j(r, t) W(k_h, j, z), \quad (\text{B1})$$

where $a_j(r, t)$ is the displacement of the isopycnal layer and $W(k_h, j, z)$ is the depth distribution of the solitary internal waves. The solitary internal waves are assumed to be created outside the acoustic waveguide of interest; Eq. (B1) represents a steady state of the waves propagating in the acoustic channel.

Just like the background internal waves, the solitary internal waves must satisfy the depth-dependent wave equation, Eq. (A3). Thus, in principle, many different eigenmodes can be excited as expressed in Eq. (B1). However, experimental data indicated that the solitons are dominated by the first mode based on the ADCP measurements.¹¹ We shall thus assume only the first mode is involved. Hence,

$$\eta_S(r, z, t) = a_1(r, t) W(k_h, j = 1, z).$$

The isopycnal displacement $a_1(r, t)$ can be expressed in the form

$$a_1(r, t) = \Lambda \operatorname{sech}^2 \left(\frac{r - Vt}{\Delta} \right), \quad (\text{B2})$$

where Λ is the amplitude factor, and Δ is the characteristic width of the soliton. Equation (B2) is known as the solution of the Korteweg–de Vries (KdV) equation.²⁶ Theoretically, one has $\Delta = \sqrt{12\beta/(\alpha\Lambda)}$ where α and β are coefficients of the KdV equations: $\alpha = 0.035$ and $\beta = 100$. V is an amplitude-dependent wave speed and is related to the linear speed ν by $V = \nu + \alpha\Lambda/3$. Equation (B2) describes one soliton.

Experimental data indicated that solitons are grouped into a series. For the simulation study we shall use an empirical representation for the solitons,

$$\eta(r, z, t) = W(k_h, j = 1, z) \sum_{m=1}^6 \Lambda_m \operatorname{sech}^2 \left(\frac{r_m - V_m t}{\Delta_m} \right), \quad (\text{B3})$$

where $\Delta_m = 100$ m and $V_m = 0.6$ m/s (the tidal velocity) are empirically determined. The solitons are separated by a distance from 100 to 400 m, the spacing decreases with the soliton number. The solitons have a decreasing amplitude: we assume that Λ_m decays exponentially¹⁵ with the soliton number m

$$\Lambda_m = \Lambda_1 e^{-0.3(m-1)}, \quad m = 1, \dots, 6.$$

The amplitude Λ_1 as determined from the SWARM 95 data is about 10–15 m.¹¹ We assume a value of 13.5 m.

¹D. R. Jackson and T. E. Ewart, “The effect of internal waves on matched-field processing,” J. Acoust. Soc. Am. **96**, 2945–2955 (1994).

²J. L. Krolik, “Matched field minimum variance beamforming in a random ocean channel,” J. Acoust. Soc. Am. **92**, 1408–1419 (1992).

³P. J. Book and L. W. Nolte, “Narrow-band source localization in the presence of internal waves for 1000-km range and 25-Hz acoustic frequency,” J. Acoust. Soc. Am. **101**, 1336–1346 (1997).

⁴P. J. Book and L. W. Nolte, “Wide-band source localization in the presence of internal waves,” J. Acoust. Soc. Am. **102**, 2597–2606 (1997).

- ⁵T. C. Yang, "A method of range and depth estimation by modal decomposition," *J. Acoust. Soc. Am.* **82**, 1736–1745 (1987).
- ⁶T. C. Yang, "Effectiveness of mode filtering: A comparison of matched-field and matched-mode processing," *J. Acoust. Soc. Am.* **87**, 2072–2084 (1990).
- ⁷K. D. Heaney, "Simultaneous inversion for source location and internal wave strength using long range ocean acoustic signal," in *Scripps Institute of Oceanography* (University of California, San Diego, 1997), pp. 152–160.
- ⁸K. D. Heaney and W. A. Kuperman, "Very long-range localization with a small vertical array," *J. Acoust. Soc. Am.* **104**, 2149–2159 (1998).
- ⁹D. D. Pierce, C.-S. Chiu, C. W. Therrien, and J. H. Miller, "Matched-mode localization using conventional and cumulant-based MUSIC algorithms in a real, complex shallow water environment," *J. Acoust. Soc. Am.* **100**, 2636 (1996).
- ¹⁰See R. Dashen, W. H. Munk, K. M. Watson, and F. Zachariasen, *Sound Transmission Through a Fluctuating Ocean*, edited by S. M. Flatte (Cambridge University Press, Cambridge, MA, 1979).
- ¹¹J. R. Apel, M. Badiay, C.-S. Chu, S. Finette, R. Headrick, J. Kemp, J. F. Lynch, A. Newhall, M. H. Orr, B. H. Pasewark, D. Tielburger, A. Turgut, K. Von de Heydt, and S. Wolf, "An overview of the 1995 SWARM shallow water internal wave acoustic scattering experiment," *IEEE J. Ocean Eng.* **22**, 465–500 (1997).
- ¹²J. X. Zhou, X. Z. Zhang, and P. H. Rogers, "Resonant interaction of sound wave with internal solitons in the coastal zone," *J. Acoust. Soc. Am.* **90**, 2042–2054 (1991).
- ¹³D. B. King, S. A. Chin-Bing, and R. W. McGirr, "Effects of shallow water internal waves on broadband acoustic wave propagation," in *Theoretical and Computational Acoustics*, edited by D. Lee and M. H. Schultz (World Scientific, Singapore, 1994), pp. 793–807.
- ¹⁴D. B. Creamer, "Scintillating shallow-water waveguides," *J. Acoust. Soc. Am.* **99**, 2825–2838 (1996).
- ¹⁵D. Tielburger, S. Finette, and S. Wolf, "Acoustic propagation through an internal wave field in a shallow water waveguide," *J. Acoust. Soc. Am.* **101**, 789–808 (1997).
- ¹⁶J. C. Preisig and T. F. Duda, "Coupled acoustic mode propagation through continental-shelf internal solitary waves," *IEEE J. Ocean Eng.* **22**, 256–269 (1997).
- ¹⁷R. H. Headrick, "Analysis of internal wave induced mode coupling effects on the 1995 SWARM experiment acoustic transmissions," Ph.D. thesis, Massachusetts Institute of Technology/Woods Hole Oceanographic Institution (1997).
- ¹⁸R. H. Headrick, J. F. Lynch, and the SWARM Group, "Acoustic normal mode fluctuation statistics in the 1995 SWARM internal wave scattering experiment," *J. Acoust. Soc. Am.* (submitted).
- ¹⁹B. Pasewark, S. N. Wolf, M. Orr, and J. F. Lynch, "Fluctuations in acoustic propagation seen in the SWARM95 experiment," *J. Acoust. Soc. Am.* **104**, 1765 (1998).
- ²⁰D. Rouseff and A. Turgut, "Coherence of acoustic modes propagating through shallow water internal waves," *J. Acoust. Soc. Am.* **104**, 1765 (1998).
- ²¹T. C. Yang and T. Yates, "Improving the sensitivity of full-field geoacoustic inversion for estimating bottom sound speed profiles," in *Full Field Inversion in Seismic and Ocean Acoustics*, edited by O. Diachok, A. Caiti, P. Gerstoft, and H. Schmidt (Kluwer Academic, Dordrecht, The Netherlands, 1995), pp. 323–328.
- ²²T. C. Yang and T. Yates, "Acoustic inversion of bottom reflectivity and IEEE bottom sound-speed profile," *IEEE J. Ocean Eng.* **21**, 367–376 (1996).
- ²³T. C. Yang and T. Yates, "Matched beam processing: Range tracking with vertical arrays in mismatched environments," *J. Acoust. Soc. Am.* **104**, 2174–2188 (1998).
- ²⁴Kwang Yoo and T. C. Yang, "Improved vertical array performance in shallow water with a directional noise field," *J. Acoust. Soc. Am.* **104**, 3326–3338 (1998).
- ²⁵T. C. Yang and T. Yates, "Matched beam processing: Applications to a horizontal line array in shallow water," *J. Acoust. Soc. Am.* **104**, 1316–1330 (1998).
- ²⁶See, for example, G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).
- ²⁷J. R. Apel, L. A. Ostrovsky, and Y. A. Stepanyants, "Internal solitons in the ocean," The Johns Hopkins University Applied Physics Laboratory, Laurel, MD, Report MERCJRA0695 (1995); J. R. Apel, S. Finette, M. H. Orr, and J. F. Lynch, "The 'dnoidal' model for internal tides and solitons on the continental shelf," *J. Geophys. Res.* (to be published).
- ²⁸Evidence for a modified Garrett-Munk model for shallow water diffused internal waves can be found in T. C. Yang and K. Yoo, "Internal waves spectrum in shallow water: Measurement and comparison with the Garrett-Munk model," *IEEE J. Ocean Eng.* **24**, 333–345 (1999).
- ²⁹M. Porter, "The KRAKEN normal mode program," SACLATENCEN Mem. SM-245 (1991).
- ³⁰M. D. Collins, "A higher-order parabolic equation for wave propagation in an ocean overlaying an elastic bottom," *J. Acoust. Soc. Am.* **86**, 1459–1464 (1989).
- ³¹M. D. Collins, "A split-step Pade solution for parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742 (1993).
- ³²X. Tang and F. D. Tappert, "Effects of internal waves on sound pulse propagation in the straits of Florida," *IEEE J. Ocean Eng.* **22**, 245–251 (1997). See also F. Tappert, J. L. Speisberger, and L. Boden, "New full-wave approximation for ocean acoustic travel time predictions," *J. Acoust. Soc. Am.* **97**, 2771–2782 (1995).
- ³³P. C. Mignerey and S. Finette, "Multichannel deconvolution of an acoustic transient in an oceanic waveguide," *J. Acoust. Soc. Am.* **92**, 351–364 (1992).
- ³⁴T. C. Yang, "Broadband source localization and signature estimation," *J. Acoust. Soc. Am.* **93**, 1797–1806 (1993).
- ³⁵For the Garrett-Munk internal wave model to yield the measured sound-speed covariance, one needs $E_0 = 154 \text{ J/m}^2$. Note that the estimate of E_0 changes depending on the sound-speed and buoyancy profiles used in the model.

Range-dependent matched-field inversion of SWelLEX-96 data using the downhill simplex algorithm

Martin Musil, N. Ross Chapman,^{a)} and Michael J. Wilmot

School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia V8W 3P6, Canada

(Received 10 September 1998; revised 27 July 1999; accepted 1 September 1999)

Matched-field inversion (MFI) techniques have been applied for effective and efficient estimation of geoacoustic parameters of the ocean bottom. This paper presents a new tomographic MFI method for use in range-dependent environments. The MFI correlates modeled data with measured data and uses a search algorithm to determine model parameter values that maximize the correlator. In the present method, a parabolic equation propagation model is used to compute replica fields to account for mode coupling in the environment. The search algorithm is a two-stage hybrid method that combines an initial global component and a final local component. The first stage employs a random search to determine $N+1$ parameter sets with the best correlations, where N is the number of parameters being determined. In the next stage the $N+1$ sets are used as inputs to the local downhill simplex algorithm. The algorithm is shown to perform well for simulated vertical line array data for an environment representative of the SWelLEX-96 experimental site. The inversion technique is then applied to measured data obtained from a radial track in SWelLEX-96. The geometric parameters of the experiment and dominant geoacoustic parameters were successfully recovered for data with relatively low signal-to-noise ratio. © 1999 Acoustical Society of America.

[S0001-4966(99)04012-6]

PACS numbers: 43.30.Pc, 43.60.Pt [DLB]

INTRODUCTION

Inversion methods for estimating ocean bottom properties have generally been applied as range-independent processes. However, it is well known that the spatial variability of the ocean waveguide has a very strong effect on sound propagation in shallow water, and the results of inversion methods based on acoustic field data are consequently limited by the mismatch in modeling the range-dependent bottom environment. Recently, several investigators have designed tomographic inversion methods for range-dependent waveguides. Hermand and Gerstoft¹ reported a tomographic matched field inversion based on adiabatic normal mode theory, and Pignot and Chapman² applied a tomographic inversion technique based on ray theory to data from the Haro Strait experiment³. Both these methods made use of broadband data. The related problem of source localization in range-dependent environments has been investigated experimentally by Booth *et al.*⁴ and D'Spain.⁵

This paper presents a new matched-field tomographic inversion technique for range-dependent environments. The method makes use of a random search algorithm with a multi-range cost function, i.e., the cost function consists of data sets from a number of independent source positions which are evaluated simultaneously. The search algorithm contains a global Monte Carlo component and a local minimizer based on the downhill simplex method. Replica fields are calculated using the parabolic equation (PE) method. Both the source position and the waveguide properties are estimated, and the method provides a measure of the sensitivity of each model parameter.

The inversion method is described in the next section, and is demonstrated on simulated broadband data consisting of multiple CW tones in Sec. III. In Sec. IV, the method is applied to data from the SWelLEX-96 experiment⁶ to estimate both the source position and the geoacoustic properties along a radial track.

I. INVERSION METHOD

A. Sound propagation model

A PE propagation model was chosen in this analysis for its ability to account for mode coupling which takes place in a range-dependent waveguide.^{7,8} Our PE formulation in cylindrical coordinates is as follows:

$$\frac{\partial \psi(r, z)}{\partial r} = ik_0(-1 + \sqrt{1+X})\psi(r, z), \quad (1)$$

where $\psi(r, z)$ is the envelope function [the acoustic pressure $p(r, z) = \psi(r, z)e^{ik_0r/\sqrt{r}}$], k_0 is a reference wave number, and $X = n^2 - 1 + (\rho/k_0)(\partial/\partial z)(1/\rho)(\partial/\partial z)$ (n is the index of refraction and ρ is material density). The PE is solved in small range-independent segments and the field is coupled from one step to the next as follows:

$$\psi(r + \Delta r, z) = e^{ik_0\Delta r(-1 + \sqrt{1+X})}\psi(r, z). \quad (2)$$

The exponential term with the X operator is approximated to high accuracy using a Pade series. The PE uses a complex density in the ocean bottom to approximate shear wave propagation.⁹ The approximation yields good results for low shear wave speeds (~ 300 m/s) and low grazing angles. In the inversion algorithm, the PE is solved on a computational grid which has a finite vertical and horizontal size.

^{a)}Electronic mail: chapman@uvic.ca

B. Matched-field processor

The MFI is generally formulated as a minimization problem which involves finding a global minimum of a multi-dimensional function, called the objective function. In MFI, measured data are compared with simulated replica data computed for trial models to give an objective function value which indicates the goodness of fit. The goal is to determine a set of model parameters which minimizes the objective function. The objective function depends on all the geoacoustic and geometric parameters associated with producing the propagation calculations. The most widely used measure of comparison is based on the normalized Bartlett processor¹⁰

$$E(\mathbf{m}) = 1 - \frac{|\mathbf{r}(\mathbf{m}) * \mathbf{d}|^2}{|\mathbf{r}(\mathbf{m})|^2 |\mathbf{d}|^2}, \quad (3)$$

where $\mathbf{r}(\mathbf{m})$ and \mathbf{d} are the replica and data vectors, respectively, \mathbf{m} is the set of model parameters to be determined, and $*$ indicates complex conjugation. Here $E(\mathbf{m})$ takes on values between zero and one with zero indicating a perfect match.

The difficulty in finding a minimum value of $E(\mathbf{m})$ can be due to several factors such as the size of the search space, the presence of local minima, and parameter correlations. The problem of nonuniqueness makes it difficult to interpret the results because different sets of parameters may yield very similar objective function values. To reduce this ambiguity, it is desirable to use as much information as possible about the environment when defining the objective function. This can be achieved by combining information from data at multiple times (source positions) and/or multiple frequencies.

The single-frequency objective function defined in (3) compares the complex data and replica pressure fields directly. The objective function can be expressed in terms of the covariances (cross-spectra) of the data and replica fields as¹¹

$$E(\mathbf{m}) = 1 - \frac{\sum_{k=1}^{N_s} \sum_{l=1}^{N_s} d^{*kl} r^{kl}}{\sqrt{\sum_{k=1}^{N_s} \sum_{l=1}^{N_s} |d^{kl}|^2 \sum_{k=1}^{N_s} \sum_{l=1}^{N_s} |r^{kl}|^2}}, \quad (4)$$

where N_s is the number of sensors and d^{kl} and r^{kl} are the k th row and l th column entries of the data covariance matrix $D = \mathbf{d}\mathbf{d}^\dagger$ and the replica covariance matrix $R(\mathbf{m}) = \mathbf{r}\mathbf{r}^\dagger$, respectively. $|\bullet|$ indicates absolute value and \dagger indicates conjugate transpose.

In order to take advantage of time averaging and multiple frequency processing, a modified version of the above objective function is used in this study. The frequencies are combined incoherently and scaled with a uniform weighting according to the power at each frequency by dividing each element in the data covariance matrix by the trace, since the trace is a measure of power across the array. Assuming that there are N_t uncorrelated time averaged segments, the objective function is given by

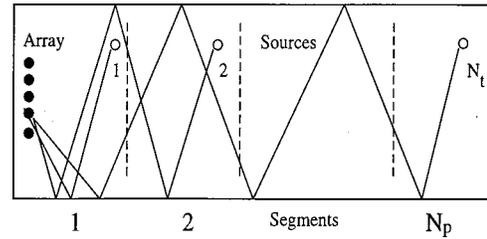


FIG. 1. The waveguide is divided into N_p segments shown by the broken lines. Data are obtained at N_t source positions. The segment profiles nearest the array affect the propagation from all the source positions.

$$E(\mathbf{m}) = 1 - \frac{1}{N_t} \times \sum_{i=1}^{N_t} \frac{\sum_{j=1}^{N_f} \sum_{k=1}^{N_s} \sum_{l=1}^{N_s} d_{ij}^{*kl} r_{ij}^{kl}}{\sqrt{\sum_{j=1}^{N_f} \sum_{k=1}^{N_s} \sum_{l=1}^{N_s} |d_{ij}^{kl}|^2 \sum_{j=1}^{N_f} \sum_{k=1}^{N_s} \sum_{l=1}^{N_s} |r_{ij}^{kl}|^2}}, \quad (5)$$

where i is summed over N_t time periods, j is summed over N_f frequencies, k and l are summed over $N_s \times N_s$ sensor pairs, and d_{ij}^{kl} and r_{ij}^{kl} are the (k, l) elements of the averaged data and replica covariance matrices, respectively, for time period i and frequency j . The minimum and maximum values of $E(\mathbf{m})$ are zero and one, respectively. The value of zero is achieved if and only if the measured and modeled data in each time period yield a value of zero.

For successful range-dependent inversions, the performance is improved by making use of data collected at different source positions, obtained for example, by towing a source slowly behind a ship. Radial source tracks are ideal because only a 2D environment (range and depth) need to be considered, but perfect radials are difficult to obtain in practice. Throughout an experiment the geometry (the relative source–receiver positions) is changing with time while the geoacoustic environment is assumed to remain unchanged. By combining information from different source positions it is possible to resolve environmental range dependence. In our method, the range-dependent waveguide between the array and the farthest source position is partitioned into N_p segments, as shown in Fig. 1. The propagation modeling for the replica fields is carried out for each source position separately. A minimum objective function value is obtained if the trial source positions and the waveguide parameters are close to their true values. The array is assumed to be approximately stationary throughout the experiment (it can tilt and move vertically with time).

How well a particular model parameter is constrained depends on the sensitivity of the data to the parameter and the number of independent data samples which depend on that parameter. The information about each individual source position is independent of that for the others. However, the source position should be accurately estimated since the acoustic field is very sensitive to the experimental geometry. On the other hand, information about the intermediate segments of the environment is contained in multiple snapshots of the data. As indicated in Fig. 1, a moving source transmits energy from different vantage points as it moves through the waveguide, repeatedly probing the environment from differ-

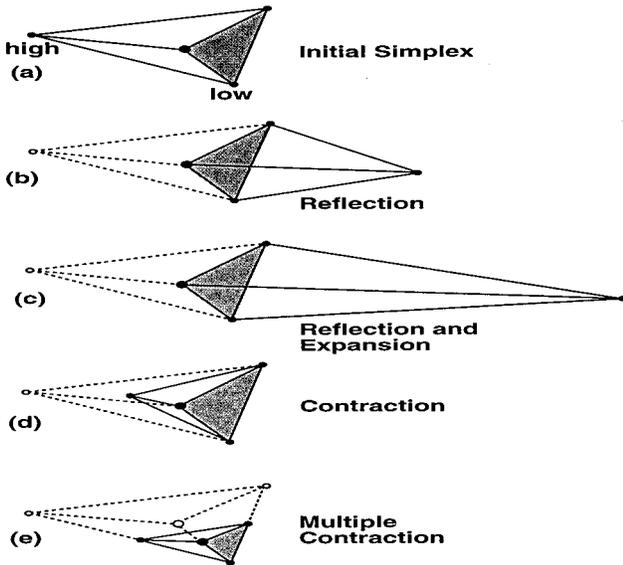


FIG. 2. Possible simplexes for various operations (b)–(e) for initial simplex (a) in 3D. The vertices associated with the highest and lowest objective function values are marked.

ent spatial positions during the collection of N_t data samples, one for each source position. The profiles in the waveguide segment closest to the array affect the propagation from all of the source positions, and hence should be well determined. The profiles in the farthest segment will, in general, be more ambiguous because the information about the properties is carried in a smaller number of the N_t data samples. Overall, the geoacoustic parameters should be relatively well constrained, since the inversion makes use of data simultaneously from a large number of independent samples.

C. Objective function optimization

A number of techniques have been applied to the matched-field inversion problem. Some of the most common ones are simulated annealing (SA)^{12–14} and genetic algorithms (GA).^{15,16} These are based on stochastic techniques as opposed to analytic methods. The greatest hurdle in applying these algorithms is obtaining the appropriate setting of the parameters that determine how the search proceeds. If not properly set, computational time may be wasted or the desired model parameters may not be found. In MFI, calculation of the replicas from the propagation model dominates the computational time, and the required time can easily become prohibitive if the search algorithm is inefficient. Nevertheless, they have been shown to be quite successful in applications with experimental data.

In this study, a technique combining a global search component with a local minimizer is introduced. Previously, a similar method using gradient descent for the local minimizer was developed and tested on synthetic data.¹⁷ Our method makes use of the downhill simplex algorithm (DSA)^{18,19} for the local minimizer. The DSA is a deterministic local minimizer based on a geometrical approach. For a M -dimensional search problem, the DSA begins with $M + 1$ trial models and, in effect, constructs a simplex of vertices,

as shown schematically in Fig. 2(a) for a three-dimensional case, then proceeds to determine new models (vertices) and checks whether an improved model (lower objective function value) is obtained. The operations are shown in Fig. 2(b)–(d). They involve reflections/contractions along the line joining the point with the highest objective function value and the center of the opposite face. The algorithm starts over if it encounters an improvement in one of the three operations. If no improvement is found, the algorithm contracts along all the dimensions towards the point with the lowest objective function value as shown in Fig. 2(e). It has the ability to explore the search space within and around itself to some extent and eventually contracts to a local minimum if allowed to run long enough. Note that the DSA does not require partial derivatives. Hence this method is neither stochastic nor calculus based, but instead is a deterministic geometrical approach. The objective function minimization algorithm consists of an initial search and I restarts:

- (i) initial: N_0 random vectors \mathbf{m} in a bounded search region are generated and the $M + 1$ vectors with the lowest objective function value are input to the DSA. The DSA converges to a local minimum $\mathbf{m}_{\min(0)}$.
- (ii) restart i : N_i random vectors within a fraction f_i ($f_i < 1$) of the parameter search width centered about $\mathbf{m}_{\min(i-1)}$ are generated and the M vectors with the lowest objective function value plus $\mathbf{m}_{\min(i-1)}$ are input to the DSA. The DSA converges to a local minimum $\mathbf{m}_{\min(i)}$. (The f_i decrease to zero as the restart number i increases. In the three-restart example below, $f_1 = 0.15$, $f_2 = 0.10$, $f_3 = 0.05$.)
- (iii) The components of $\mathbf{m}_{\min(i)}$ are the algorithm parameter estimates.

In this scheme the current best model vector is included in the subsequent DSA run and so there is no possibility of losing the best available solution. Since almost all physical problems require constraints on the parameters, it is necessary that all the parameters remain within allowed bounds as the algorithm proceeds. This can be accomplished using the objective function by adding a penalty function to its value whenever one or more trial parameter values occurs outside its lower or upper bounds. The following penalty function was used:

$$Q(\mathbf{m}) = 10^R \sum_{j=1}^M p_j^2$$

$$\text{where } p_j = \begin{cases} \frac{(m_j - m_{Uj})}{(m_{Uj} - m_{Lj})}, & \text{if } m_j > m_{Uj}, \\ \frac{(m_j - m_{Lj})}{(m_{Uj} - m_{Lj})}, & \text{if } m_j < m_{Lj}, \\ 0, & \text{otherwise,} \end{cases} \quad (6)$$

where R is a power coefficient typically between 1 and 6 (here $R = 2$), m_j is the j th parameter estimate, p_j is the penalty for the j th parameter, m_{Lj} is the lower bound for the j th parameter, and m_{Uj} is the upper bound for the j th parameter.

Note that $Q(\mathbf{m})$ is zero if all the parameters are between their bounds. $Q(\mathbf{m})$ is added to $E(\mathbf{m})$ to obtain the objective function,

$$E'(\mathbf{m}) = E(\mathbf{m}) + Q(\mathbf{m}). \quad (7)$$

The penalty function usually causes the parameter to return to the allowed region. At the end of the inversion the recovered parameters are examined to determine whether they are close to lower or upper bounds. A parameter close to a bound indicates that the true solution may have been excluded. This may render the results meaningless depending on the type of parameter and proximity to the constraint. The constraints may have to be readjusted and the inversion performed again.

D. Sensitivity of parameters

In this study, the parameter sensitivity is estimated from the set of N_1 randomly selected models drawn from the search space. A small set of models which gave the lowest objective function value are saved. The sensitivity (or spread) coefficient, Spdc, for each parameter is defined to be the ratio of the observed standard deviation s_i of these best models and the theoretical standard deviation of a random uniform variable of width w_i

$$\text{Spdc}_i = \sqrt{12}s_i/w_i, \quad (8)$$

where w_i is the upper minus the lower bound of the search region. Recall that a uniformly generated random variable over an interval w_i has standard deviation $w_i/\sqrt{12}$. Spdc_{*i*} is near zero if the estimated values of the selected model parameter values are very similar, and approximately unity if they are uniformly distributed over the search region. The Spdc values depend on the parameter bounds so that conclusive physical interpretations cannot be made. However, they indicate relative sensitivities: small values of Spdc indicate high sensitivity and larger ones low sensitivity.

A single noise-free data set from a suite of workshop test cases²⁰ was chosen for this analysis. The data at 105 Hz generated by SAFARI²¹ consists of 100 complex pressure fields from a vertical array with sensors at 1- to 100-m depths. The source is at 1-km range and 20-m depth. A frequency of 105 Hz was selected because it is approximately at the center of the frequency band used for the SWelLEX-96 data set analyzed below.

The true environment is range independent and consists of three fluid layers. The water is a 100-m inhomogeneous layer in which the sound speed decreases linearly from 1480 m/s at the top to 1460 m/s at the bottom. The basement sound speed and density are 1800 m/s and 2.0 g/cm³, and the attenuation increases linearly from 0.23 to 5.0 dB/λ over the computational grid (extending to a depth of 1000 m). The sediment geoacoustic parameters, and the geometric and bathymetric parameters that were varied in the search process are listed in Table I. Range dependence was introduced by allowing the water depth and sediment thickness to vary at both ends of the waveguide; consequently, the inversion could yield a uniformly sloping bottom. There are 13 variable parameters: 3 geometric (source range, source depth,

TABLE I. Model parameters with indicated true values and lower and upper bounds used for the sensitivity study.

Parameter	Lower bound	Upper bound	True value
Source depth (m)	17	30	20
Source range (m)	700	1100	1000
Array tilt (m)	-5	3	0
Water depth at 0 km (m)	60	120	100
Water depth at 5 km (m)	80	140	140
Sediment thickness (m)	20	70	50
Sediment P-wave speed at top (m/s)	1450	1750	1550
Sediment P-wave speed at bottom (m/s)	1600	1850	1700
Sediment P-wave attenuation at top (dB/λ)	0.15	0.4	0.23
Sediment P-wave attenuation at bottom (dB/λ)	0.15	0.4	0.23
Sediment density at top (g/cm ³)	1.4	1.7	1.5
Sediment density at bottom (g/cm ³)	1.4	1.7	1.5

and array tilt), 4 structural (2 water depths and 2 sediment thicknesses), and 6 geoacoustic (density, compressional-wave speed, and attenuation at the sediment top and bottom). The following run parameters were used: $N_0=6000$, $N_1=1000$, $N_2=500$, $N_3=500$, $f_1=0.15$, $f_2=0.1$, and $f_3=0.05$.

In the lower right corner of Fig. 3 is a plot of the evolution of the objective function versus the DSA iteration number (solid line), and the dotted horizontal line at 0.0027 gives the objective function value for the true parameters. The discrepancy from zero is due to the use of different propagation models for the data and replica generation. The solid line consists of DSA evolution curves for the initial evolution and three restarts. Note that each DSA run lowers the objective function value and that very few iterations are wasted. The method required ~670 PE calls per parameter to reach the objective function value of 0.005. In the remaining plots the histograms were computed using the 60 models (out of 6000) which gave the lowest objective function values. The spread coefficients are shown in the left corner of each plot; the Spdc's vary from 0.49 to 1.12 in this particular case. For the array tilt, Spdc=0.49, which makes it the most sensitive parameter. The coefficient with value 1.12 corresponds to the sediment thickness at $r=0$ km, which is among the least sensitive parameters. All other Spdc's fall within the extreme values. Note that it is very likely that different random seeds would result in different Spdc's and so different parameters may have the extreme Spdc values in another random set. The true parameter values are indicated by a solid vertical line. The five diamonds connected by dotted lines represent the evolution of the parameter: the bottom diamond is the best model from the first random search and the remaining four are the final values after each DSA run. These also give an indication whether a parameter has a tendency to converge to the true value and hence indicate sensitivity. The top diamond denotes the final parameter estimate. All the geometric parameters converged closely to the true values. Also, the water depths and sediment thicknesses are very well estimated. The only well-estimated geoacoustic parameter is the sound speed.

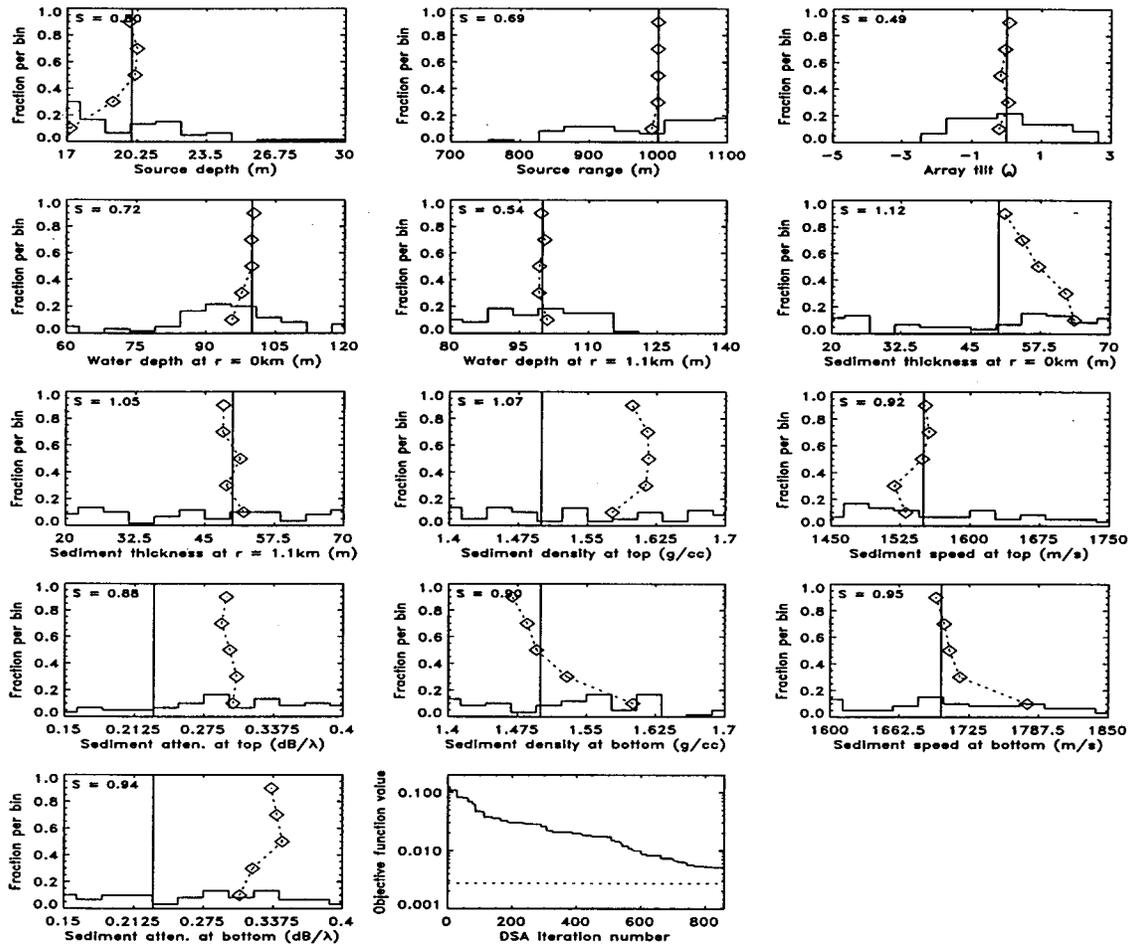


FIG. 3. Histograms and spread coefficients (denoted S) for the best 60 models for the 13 parameters. The solid vertical line on these plots indicates the true parameter value and the diamonds from bottom to top the initial and subsequent restart parameter estimates. The last plot at the lower right shows the objective function values versus DSA iteration number (solid line), and the objective function value for the true parameters (dotted line).

The results in Fig. 3 give an indication of how well various parameters can be estimated. It is to be expected that sensitive parameters such as source depth, source range, water depth, and sediment sound speed will be tightly constrained, and hence different starting models will yield consistent estimates. On the other hand, the less-sensitive parameters such as density and attenuation may not be consistently estimated, particularly for single frequency data. The above method yields parameter sensitivities that are based on all the models explored in the search process, rather than estimates based on only a single point or small neighborhood in the search space, and is therefore a useful complementary process for assessing parameter sensitivities.

II. SIMULATED DATA

Prior to applying the inversion algorithm to experimental data, the algorithm was first tested on a set of synthetic data generated using a range-dependent propagation model COUPLE.²² This code can accurately model range-dependent sound propagation because it accounts for mode coupling and the incoming (backscatter) component of the field. However, the model is limited to purely fluid environ-

ments. Data at 45 and 70 Hz representative of the low-frequency band of the SWellEX-96 data were calculated to carry out parameter estimation.

The simulation scenario is depicted in Fig. 4. There are

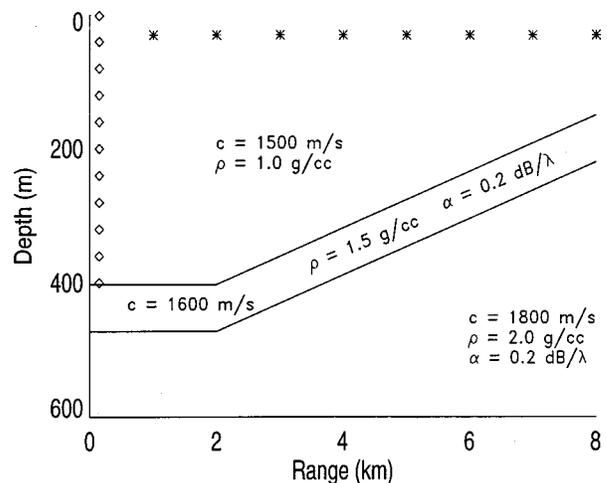


FIG. 4. A schematic of the range-dependent environment with eight source positions (stars) and a 50-element vertical line array (diamonds) used for the simulated data.

TABLE II. Environmental model parameters with indicated true values and lower and upper bounds used for the simulated data.

Parameter	Lower bound	Upper bound	True value
Water depth at left (m)	360	470	402
Water depth at right (m)	100	240	150
Sediment thickness (m)	40	80	70
Sediment P-wave speed at top (m/s)	1500	1750	1600
Sediment P-wave speed at bottom (m/s)	1500	1750	1600
Sediment P-wave attenuation at top (dB/λ)	0.1	0.4	0.2
Sediment P-wave attenuation at bottom (dB/λ)	0.1	0.4	0.2
Sediment density at top (g/cm ³)	1.4	1.7	1.5
Sediment density at bottom (g/cm ³)	1.4	1.7	1.5
Basement P-wave speed at top (m/s)	1650	1900	1800
Basement P-wave speed at bottom (m/s)	1650	1900	1800

eight equally spaced source positions from 1 to 8 km indicated by stars. All eight have a depth of 30 m. At range 0 km there is a 50-element vertical line array. The array sensors are equally spaced from 1 to 400 m, i.e., the array (indicated by diamonds) spans the entire water column. The array more than adequately samples higher-order modes which may be present in the field. The range-dependent water depths are an approximation to a continental shelf environment. A 2-km flat segment at a depth of 402 m originating at the array is followed by a 6-km constant slope segment extending to a maximum range of 8 km at a depth of 150 m. Hence the sloping part of the bottom has an angle of 2.4 degrees. As shown in Fig. 4, the bottom structure consists of a water layer, a 70-m sediment layer over the entire extent of the environment, and a half-space which extends to the bottom of the computational grid. All three layers are homogeneous, i.e., geoacoustic parameters are range independent. The water layer has a sound speed of 1500 m/s and density of 1.0 g/cm³. The sediment has the following physical properties: sound speed of 1600 m/s, density of 1.5 g/cm³, and attenuation of 0.2 dB/λ. The basement has a sound speed of 1800 m/s and density of 2.0 g/cm³. The attenuation at the top is the same as in the sediment and it increases linearly to 5 dB/λ at the bottom of the computational grid. The size of the PE computational grid is 1000 m, as in the previous example. The true values and search limits are tabulated in Table II for the geoacoustic parameters that were varied in the inversion.

There are 16 geometric variable parameters consisting of eight source ranges and eight source depths. The lower and upper bounds and the true values are indicated in Table III. R_T denotes the true range for each of the eight source ranges. It was decided to hold the sensor depth and array tilt fixed at the true values of 1 m (for the top sensor) and 0 degrees, respectively. Range dependence was addressed by partition-

TABLE III. Geometric model parameters with indicated true values and lower and upper bounds used for the simulated data.

Parameter	Lower bound	Upper bound	True value
Source range (m)	$R_T - 300$	$R_T + 100$	1000–8000
Source depth (m)	25	40	30

ing the environment in range into eight segments (nine points). The water depth and sediment thickness were allowed to vary independently at each of the nine points, while only one value was assumed for each geoacoustic parameter at each interface in the bottom layers. This resulted in 26 variable environmental parameters consisting of nine water depths, nine sediment thicknesses, sound speed, density, and attenuation at the sediment top and bottom, and sound speed at the basement top and bottom. Note that the basement bottom implies the bottom of the computational grid. The density and attenuation in the basement are held fixed at the true values. The attenuation gradient in the basement layer minimizes reflections from the bottom of the computational grid.

In this scenario, the farthest source position is in 150 m of water and the field is received in 402 m of water; thus there are fewer modes at the initial range than there are at the final range. The field has to adjust to significant environmental changes during the propagation. The positions where the water depth and sediment thickness were allowed to vary were chosen to be approximately under the source positions. Since mode functions are characteristic of local waveguide depths, the acoustic field contains sufficient information to accurately determine the water depths.

The computational intensity of this problem was the main factor in deciding the extent of the search. Three separate inversions were performed with the following parameters: $N_0 = 2000$, $N_1 = 1000$, $N_2 = 500$, $N_3 = 500$, $f_1 = 0.15$, $f_2 = 0.1$, and $f_3 = 0.05$. Each DSA search was stopped when the difference between the maximum and minimum objective function values was less than 10^{-4} .

The estimated source positions are shown in Fig. 5. The source depths are within 1.5 m of the true depth of 30 m, which confirms the high sensitivity of this parameter. The source ranges are also reasonably well recovered. The best source range estimate coincides exactly with the true value. The largest error in source range is 133 m, most likely due to small mismatch in the water depths. This offset does not appreciably degrade the objective function value. The source range seems to be correlated with the water depth, making it

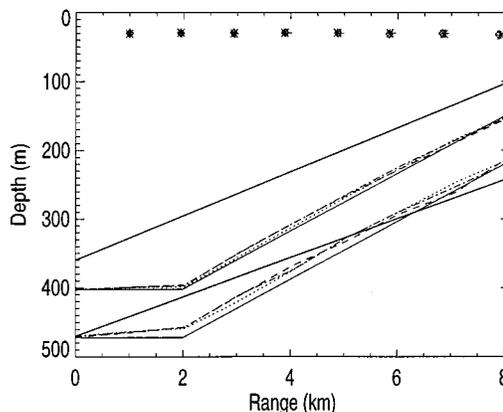


FIG. 5. Three estimates for the source positions (crosses, stars, diamonds) of the simulated data. Three estimates for the bottom structure (dotted, dashed, dash-dotted) for the sea floor and sediment layer. The thin solid lines indicate the true bottom structure. The thick solid parallel lines indicate the lower and upper bounds on the water depths.

TABLE IV. Objective function values and estimated geoacoustic parameters for the simulated data.

Parameter	True value	Inversion 1	Inversion 2	Inversion 3
Objective function value		0.033	0.032	0.030
Sediment P-wave speed at top (m/s)	1600	1603	1607	1599
Sediment P-wave speed at bottom (m/s)	1600	1587	1586	1595
Sediment P-wave attenuation at top (dB/ λ)	0.2	0.25	0.31	0.26
Sediment P-wave attenuation at bottom (dB/ λ)	0.2	0.29	0.27	0.26
Sediment density at top (g/cm ³)	1.5	1.54	1.48	1.53
Sediment density at bottom (g/cm ³)	1.5	1.50	1.55	1.61
Basement P-wave speed at top (m/s)	1800	1762	1802	1753
Basement P-wave speed at bottom (m/s)	1800	1812	1776	1818

difficult to find the true values for these parameters. In the same figure the estimated water depths and sediment thicknesses are plotted as a function range. The water depths are estimated accurately and consistently, which indicates a high sensitivity of this parameter. The sediment thickness values are estimated with less accuracy due to a much lower sensitivity with depth in the bottom.

The objective function values and estimated geoacoustic parameters are shown in Table IV. Note that the objective function values are very near the value of 0.024 achieved when the true parameters are used to calculate the replicas. The sound speed estimates at the upper and lower interfaces verify that the layer is homogeneous with a sound speed of 1600 m/s. The sediment density and basement sound speed

SWellEx-96 SMFP-V Track D to F2
Micro Mariner data, JD 135 10:53 to 11:25 GMT

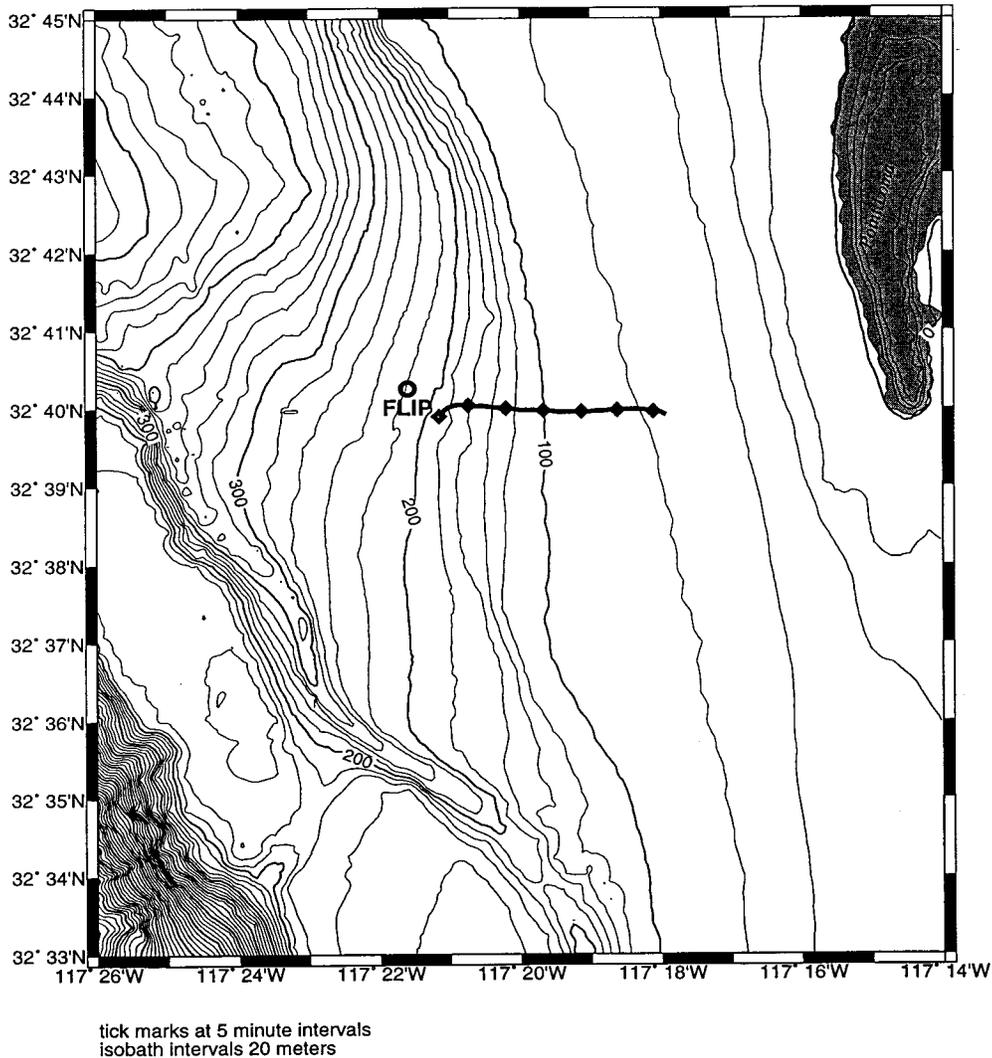


FIG. 6. The radial track from SWellEx-96. Data were analyzed for ship positions roughly between the first and second tick marks closest to FLIP.

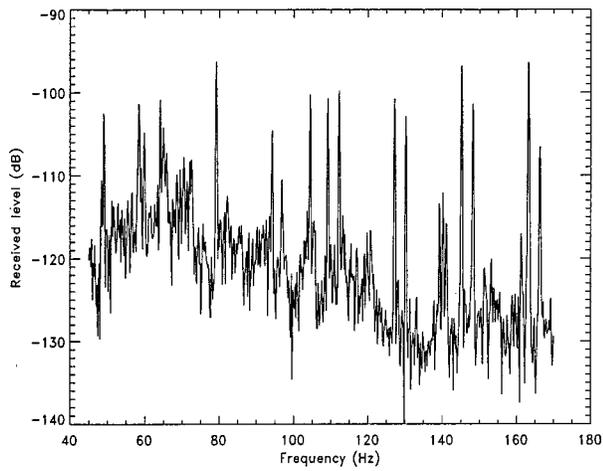


FIG. 7. Example of an estimated power spectrum (an average of three five-second FFT's) for the SWellEX-96 data.

are not as accurately estimated as the sediment sound speed, as expected for less sensitive parameters. For very low sensitive parameters like the sediment attenuation, the estimated value is roughly the mean of the search range.

III. SWellEX-96 DATA

In this section the inversion algorithm is applied to experimental data. The SWellEX-96 data were obtained in shallow waters off the Southern California coast in May 1996.⁶ The experiment provided a large amount of high-quality data for various geometrical and environmental configurations. In this study, vertical array data are analyzed from a radial incoming track (event S107, 14 May²³) over a gently downsloping bottom from about 150 m at the start of the radial to 215 m at the array (Fig. 6). The portion of the data analyzed spans a segment of the track from 1.6 to 1 km from the array position. The source was towed at a speed of about 2.5 m/s and a depth of about 37 m while transmitting CW signals at frequencies between 49 and 400 Hz. Different source levels were available over the entire frequency band to investigate the effects of various signal-to-noise ratio (SNR) values on matched-field processing and matched-field inversion. The band of frequencies used in this study covers frequencies from 49 to 169 Hz. The vertical array consisted of 64 equally spaced hydrophones deployed from the research platform FLIP, making a 118-m aperture. The pressure field was sampled at 1500 Hz at each hydrophone. Source depths were monitored with a pressure sensor, and ranges were determined from differential GPS data on the tow ship and FLIP. Array tilt was measured using an inclinometer on the array cable. The average tilt was 0.5 ± 0.1 degrees, (from the vertical) oriented roughly perpendicular to the ship track during event S107.^{23,24} This magnitude of tilt was observed for several hours before and after the event.

The raw data were Fourier transformed into the frequency domain to obtain complex pressure fields at the hydrophones at the frequencies of interest. Three 5 s FFT's were averaged to obtain covariance matrices at various frequencies and source positions. The data were found to be consistently Doppler shifted to higher values by about 0.2

TABLE V. Geometric and geoacoustic model parameters with indicated lower and upper bounds for the SWellEX-96 data.

Parameter	Lower bound	Upper bound
Source range (m)	$R_{GPS} - 70$	$R_{GPS} + 70$
Source depth (m)	37	40
Array depth (m) (top sensor)	92	98
Water depth at array (m)	190	240
Water depth at 1.6 km (m)	120	170
Sediment thickness (m)	3	50
Sediment P-wave speed at top (m/s)	1500	1700
Sediment P-wave speed at bottom (m/s)	1500	1700
Sediment P-wave attenuation at top (dB/λ)	0.1	0.5
Sediment P-wave attenuation at bottom (dB/λ)	0.1	0.5
Sediment density at top (g/cm ³)	1.4	1.6
Sediment density at bottom (g/cm ³)	1.4	1.6
Basement P-wave speed at top (m/s)	1600	2100
Basement P-wave speed at bottom (m/s)	1600	2100
Basement P-wave attenuation at top (dB/λ)	0.1	0.5
Basement P-wave attenuation at bottom (dB/λ)	5.0	5.0
Basement density at top (g/cm ³)	1.8	2.1
Basement density at bottom (g/cm ³)	1.8	2.1

Hz. Ten covariance matrices (every second one) for each frequency of interest, corresponding to 5 min of data, were selected for the analysis. High SNR data and low SNR data are analyzed separately. A set of five pilot (highest level) low frequencies (LF=49.2, 64.2, 79.2, 94.2, and 112.2 Hz) at the maximum available SNR were used first, followed by a set of three pilot high frequencies (HF=130.2, 148.2, and 166.2 Hz) also at the maximum available SNR. The same procedure was repeated for a band of frequencies transmitted at a lower level (LF=52.2, 67.2, 82.2, 97.2, and 115.2 Hz followed by HF=133.2, 151.2, and 169.2 Hz). Note that the lower level frequencies are offset by 3 Hz from the pilot frequencies.

For successful inversions, the SNR of the data at signal frequencies has to be high. An average power spectrum for one of the VLA hydrophones is shown in Fig. 7. To obtain this estimate of the power spectrum three time segments were averaged and the received power in dB was calculated. The high level tones are noticeably above the ambient noise level. Note that the noise levels are higher in the lower frequency band which may render some signal lines unusable. The pilot signals are about 25 dB above the noise. However, it is quite difficult to see the lines at the low SNR especially in the lower frequency band. From quantitative SNR estimates, the SNR for the pilot frequencies is indeed about 25 dB for the section of the data used. The lower level signals were transmitted at 25 dB below the pilots which would place them at about 0 dB SNR.

The assumed waveguide model was range dependent and consisted of two elastic bottom layers: sediment and basement. Each layer could have a negative or positive gradient in sound speed, attenuation, and density. Ground truth for the geoacoustic parameter values was obtained from the study by Bachman *et al.*²⁵ The bounds for the sediment sound speed span the ranges of values for the known types of ocean bottom materials off Pt. Loma: sands and silt-clays. For fine sands (mean grain size <3.25 phi), the range of

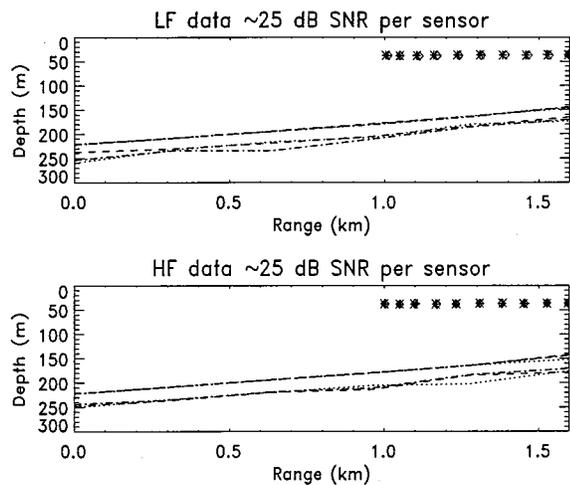


FIG. 8. Estimated source positions, water depths, and sediment thicknesses for the high SNR low-frequency (top) and high-frequency (bottom) data. Three estimates are shown for each parameter for the LF and HF SWellEX-96 data.

values for the upper 30 m of sediment is 1600–1750 m/s, and for silt-clays (mean grain size >5.75 phi), the corresponding range is 1520–1540 m/s. From the ground truth data, the mean grain size is 4.5–5.0 phi so the sediment is likely a mixture of the sand and silt, but more siltlike than sandlike. The sediment thickness is about 30 m.²⁵ The basement is inferred to be mudstone from deep borehole data, with a sound speed of 1860 m/s. Sediment shear speed and attenuation were held fixed at 300 m/s and 2.0 dB/ λ , and similarly, basement shear speed and attenuation were fixed at 500 m/s and 2.0 dB/ λ . The basement compressional attenuation varied linearly from 0.1 to 5.0 dB/ λ to the bottom of the computational grid at 500 m.

The waveguide is divided into five segments (six points) to account for the range-dependent bathymetry. The sound speed profile in the water was measured in the experiment⁶ and is not included in the parameter search. The geometric parameters (source range, source depth, and array depth) are also included in the search. The bounds on these parameters are very tight because accurate *a priori* information is available from the source depth measurements and the GPS data during the track. These tight bounds will constrain the estimation of water depth, since it is known that range and water

depth are correlated.⁵ Based on the inclinometer data, the array is assumed to be straight (no curvature) and vertical. This assumption is not expected to affect inversion performance since the very small tilt (~ 0.5 degrees) is oriented perpendicular to the plane of acoustic propagation. The following unknown parameters were included in the inversion: 30 geometric parameters (10 source ranges, 10 source depths, and 10 array depths), 6 water depths and 6 sediment thicknesses (corresponding to the 5 waveguide segments), and 11 geoacoustic parameters that are assumed to be constant with range (see Table V). The control parameters were set at the following values for all the inversions: $N_0=2000$, $N_1=1000$, $N_2=500$, $N_3=500$, $f_1=0.15$, $f_2=0.1$, and $f_3=0.05$. Each DSA search was stopped when the difference between the maximum and minimum objective function values was less than 10^{-3} . The computational CPU time for each inversion was around 20 h on a 70 MFlop DEC ALPHA server.

A. High SNR data

As mentioned above, the data were available over a wide frequency band. The objective was to demonstrate that similar results can be obtained using data at different frequency bands and hence identify the most favorable model. Three independent inversions with different random seeds were performed for each set of frequencies to demonstrate consistency.

Figure 8 indicates the results for the LF and HF data. The three sets of source range estimates (star, diamond, and cross) are consistent to within 10 m at both LF and HF and are within ± 35 m of the GPS ranges. The source depth estimates are also consistently estimated at approximately 37 m, in excellent agreement with the known source depths of 35–38 m. The estimated depth of the top hydrophone in the array, not shown in the figure, was about 95 m, which is in a good agreement with the known depth of 94 m. The water depth estimates at the array position came consistently to about 220 m, in comparison with the measured depth of 216 m. Other water depths over the track are in good agreement with bathymetric data. Note that the recovered ocean bottom has a nearly uniform slope which is consistent with the known bathymetry over the 1.6-km track.

The low objective function values in Table VI provide

TABLE VI. Objective function values and estimated geoacoustic parameters for the high SNR SWellEX-96 data. LF denotes estimates for the lower frequencies and HF denotes estimates for the higher frequencies.

Parameter	LF 1	LF 2	LF 3	HF 1	HF 2	HF 3
Objective function value	0.058	0.063	0.067	0.102	0.091	0.085
Sediment P-wave speed at top (m/s)	1575	1558	1572	1589	1568	1574
Sediment P-wave speed at bottom (m/s)	1632	1650	1663	1628	1660	1644
Sediment P-wave attenuation at top (dB/ λ)	0.30	0.33	0.40	0.28	0.32	0.22
Sediment P-wave attenuation at bottom (dB/ λ)	0.30	0.32	0.33	0.33	0.28	0.25
Sediment density at top (g/cm ³)	1.57	1.51	1.46	1.52	1.51	1.50
Sediment density at bottom (g/cm ³)	1.56	1.52	1.48	1.47	1.50	1.50
Basement P-wave speed at top (m/s)	1677	1670	1654	1790	1853	1798
Basement P-wave speed at bottom (m/s)	1767	1786	1922	1739	1870	1793
Basement P-wave attenuation at top (dB/ λ)	0.40	0.36	0.38	0.31	0.20	0.35
Basement density at top (g/cm ³)	1.98	2.00	1.90	1.90	2.00	1.98
Basement density at bottom (g/cm ³)	1.94	1.92	1.86	1.98	1.93	2.02

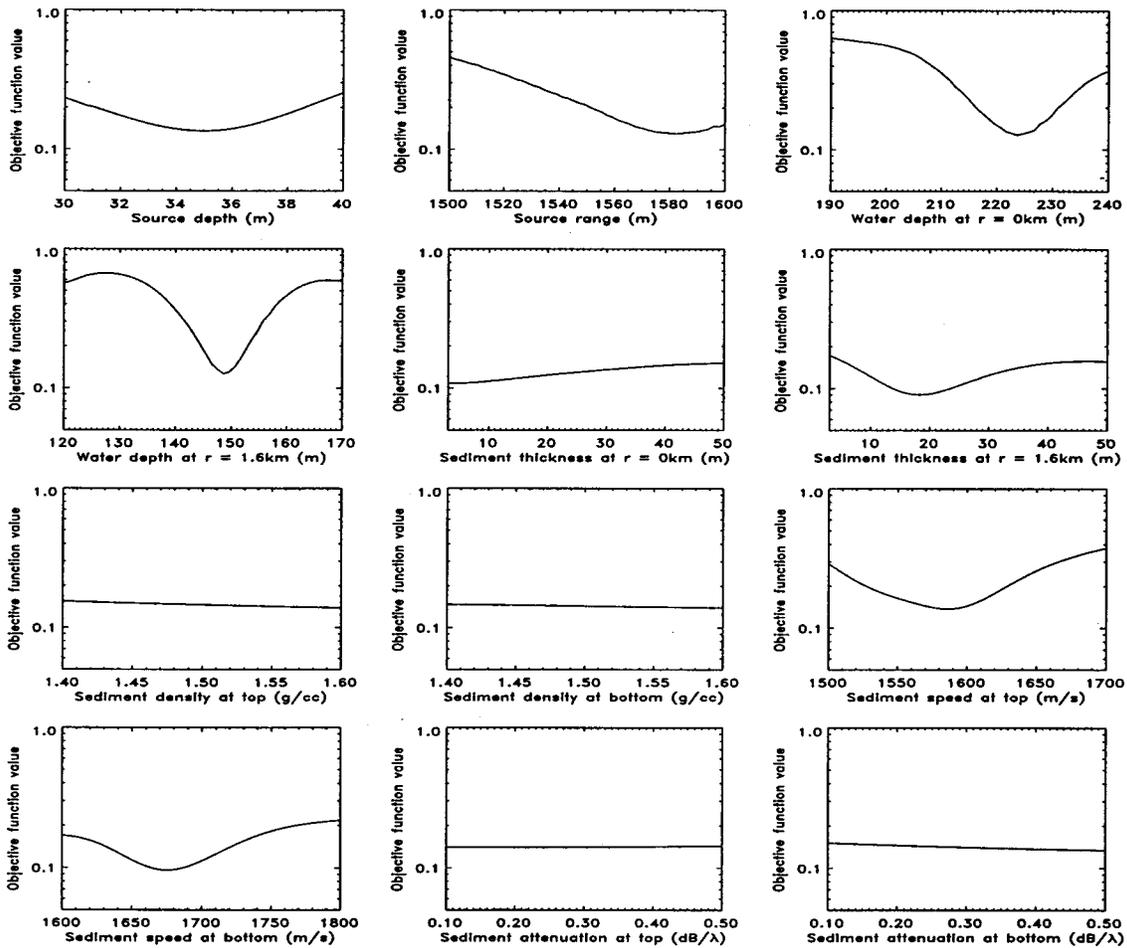


FIG. 9. Plots of the objective function value versus individual model parameters while holding all the other parameters fixed at the best estimated values. The LF high SNR SWelIEX-96 data.

confidence that the data are well modeled by the propagation code. From the sound speed estimates we can conclude that there is a positive gradient in the sediment. The average values at the top and bottom of the sediment are 1573 ± 10 m/s and 1646 ± 14 m/s, respectively. These results indicate a mixed sediment material that is more siltlike, consistent with the ground truth data. The sediment layer varied from 20 to 40 m thick at different positions. Since the sediment thickness is not highly sensitive, it is difficult to conclude whether the variations are real or the result of nonuniqueness. The attenuation estimates cluster around the mean value of the search width, as expected for a very low sensitive parameter. The HF data yield a consistent estimate of the sediment density, very near 1.5 g/cm^3 . The mean of all the estimates for the sound speed at the basement top was 1740 m/s, with a standard deviation of 83 m/s.

Plots of the objective function value versus individual model parameters for values at the algorithm minimum are shown in Fig. 9. One parameter was varied at a time over its search bounds while holding all the other parameters fixed at the best estimated values. The LF high SNR data for 1.6-km range were used and a uniformly sloping ocean bottom was assumed. These plots are useful for examining the structure of the objective function at the best estimated point in the search space. They also give insight into the relative sensi-

tivity of various parameters. Note that the objective function has the largest variation for the two water depths, and the smallest for the density and attenuation. Also note that the plot for the water depth at 1.6 km has a local minimum

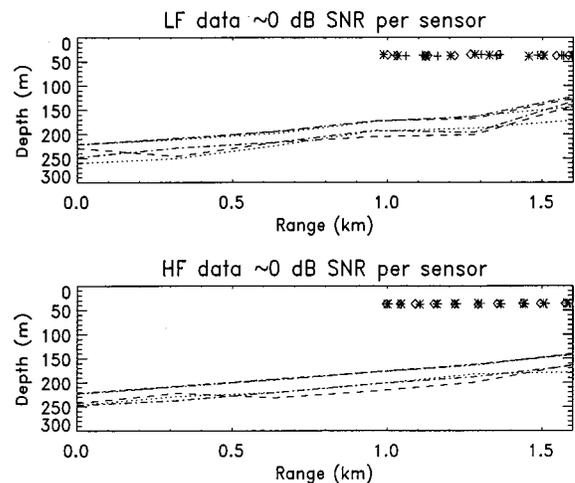


FIG. 10. Estimated source positions, water depths, and sediment thicknesses for the low SNR low-frequency (top) and high-frequency (bottom) data. Three estimates are shown for each parameter for the LF and HF SWelIEX-96 data.

TABLE VII. Objective function values and estimated geoacoustic parameters for the low SNR SWellEX-96 data. LF denotes estimates for the lower frequencies and HF denotes estimates for the higher frequencies.

Parameter	LF 1	LF 2	LF 3	HF 1	HF 2	HF 3
Objective function value	0.58	0.584	0.582	0.233	0.243	0.245
Sediment P-wave speed at top (m/s)	1613	1622	1571	1588	1606	1581
Sediment P-wave speed at bottom (m/s)	1552	1590	1608	1624	1580	1647
Sediment P-wave attenuation at top (dB/ λ)	0.33	0.31	0.27	0.31	0.36	0.30
Sediment P-wave attenuation at bottom (dB/ λ)	0.30	0.35	0.23	0.31	0.35	0.26
Sediment density at top (g/cm ³)	1.53	1.48	1.51	1.50	1.50	1.52
Sediment density at bottom (g/cm ³)	1.50	1.48	1.49	1.49	1.49	1.51
Basement P-wave speed at top (m/s)	1713	1918	1747	1794	1859	1767
Basement P-wave speed at bottom (m/s)	1956	1890	1750	1846	1857	1932
Basement P-wave attenuation at top (dB/ λ)	0.32	0.31	0.27	0.26	0.28	0.40
Basement density at top (g/cm ³)	1.87	1.97	1.83	1.95	1.96	1.89
Basement density at bottom (g/cm ³)	1.98	1.96	2.04	1.94	1.88	1.96

within the search interval. Most likely there are other local minima which are not apparent in these 1D plots.

B. Low SNR data

The next set of tones was transmitted about 25 dB below the pilot frequencies. These levels become comparable with those of surrounding noise sources, which puts them on the threshold of usefulness for our inversion method. With the use of broadband data there should be sufficient information to identify the model parameters. The five lower-frequency tones are combined together and the three higher-frequency tones are used together as was done with the pilot tones. The same model as used above is sought for both sets of tones.

The upper plot in Fig. 10 indicates the results for the LF data. The water depths were again consistent with the previous results but the sediment thicknesses exhibit much more variation. A similar variability exists in the source position estimates, and all the geoacoustic parameter estimates listed in Table VII for the LF data. There was a significant increase in the minimum value of the objective function for these tones which indicates that the data are strongly corrupted by noise in this band.

Surprisingly, the higher-frequency tones yielded much better results. It is our belief that this is a direct consequence of much lower noise levels in this band. Regardless of the absolute signal levels, which are comparable to those for the lower frequency band, the SNR for the higher frequencies is probably several dB higher. Both the environmental and geometric parameter estimates in Table VII and Fig. 10 are mutually consistent and consistent with the previous results for the high SNR data. This comparison indicates that low SNR data at relatively high frequencies can be inverted effectively by our PE-based multi-range, broadband inversion method to provide reasonable estimates of geoacoustic and geometric parameters.

IV. CONCLUSION

In this study, a search algorithm for geoacoustic inversion was presented. The algorithm consists of a global random component which ensures that all the regions of the search space of interest are sampled with equal probability and a local minimizer component which goes rapidly down-

hill to a minimum. The local minimizer downhill simplex algorithm uses a simplex of vertices to move through the search space and converge to a minimum. In combination, these two components have proved to be very effective at minimizing the objective function which is a measure of fit between the data and replica fields, and accurate parameter estimates have been achieved. Furthermore, this method can be applied to range-independent as well as range-dependent waveguides with an acceptable computational efficiency.

The technique was successfully applied to range-dependent synthetic data where the bottom slope was about 2.4 degrees. The most sensitive model parameters were accurately estimated on three occasions. The less sensitive parameters were not as well estimated but the errors were reasonably small. The technique was then used to estimate parameters using the SWellEX-96 data obtained in a range-dependent waveguide with about a 2.5-degree bottom slope. High SNR data yielded very low objective function values and consistent parameter estimates. A simple structure model with two ocean bottom layers was assumed and found to adequately represent the environment. Low SNR data at low frequencies gave consistent estimates for only the bathymetry whereas higher frequency data were just as effective and consistent as their higher SNR counterparts.

ACKNOWLEDGMENTS

The authors are grateful to Newell Booth and Phil Schey, Code D881, Space and Naval Warfare Systems Center, San Diego, CA, for providing data tapes and documentation. We also thank Megan McArthur, Gerald D'Spain, and Dick Bachman for additional information on the bathymetry at the experimental site. SWellEX-96 was sponsored by the U.S. Navy, Office of Naval Research, Code 321US. This work was supported by NSERC and DND, Canada.

¹J. P. Hermand and P. Gerstoft, "Inversion of broad-band multitone acoustic data from the YELLOW SHARK summer experiments," *IEEE J. Ocean Eng.* **21**, 324–346 (1996).

²P. Pignot and N. R. Chapman, "Tomographic inversion of geoacoustic properties in a range dependent shallow water environment," submitted to *J. Acoust. Soc. Am.* (1999).

³N. R. Chapman, L. Jaschke, M. A. Donald, H. Schmidt, and M. Johnson, "Matched field geoacoustic tomography experiments using light bulb sound sources in the Haro Strait sea trial," in *Proc. Oceans 97* (IEEE, Halifax, 1997), pp. 763–768.

- ⁴N. O. Booth, P. A. Baxley, J. A. Rice, P. W. Schey, W. S. Hodgkiss, G. L. D'Spain, and J. J. Murray, "Source localization with broad-band matched-field processing in shallow water," *IEEE J. Ocean Eng.* **21**, 402–412 (1996).
- ⁵G. L. D'Spain, W. S. Hodgkiss, J. J. Murray, P. W. Schey, and N. O. Booth, "Mirages in shallow water matched-field processing," *J. Acoust. Soc. Am.* **105**, 3245–3266 (1999).
- ⁶N. O. Booth, P. W. Schey, and W. S. Hodgkiss, "Detection of low-level broad-band signals using adaptive matched-field processing with horizontal arrays," *J. Acoust. Soc. Am.* **102**, 3170 (1997).
- ⁷M. D. Collins, "A split-step Pade solution for the parabolic equation method," *J. Acoust. Soc. Am.* **93**, 1736–1742 (1993).
- ⁸G. H. Brooke, "3-D PE propagation model for the advanced acoustic signal processor," unpublished contractor's report, 1997.
- ⁹Z. Y. Zhang and C. T. Tindle, "Improved equivalent fluid approximations for a low shear speed ocean bottom," *J. Acoust. Soc. Am.* **98**, 3391–3396 (1995).
- ¹⁰A. Tolstoy, *Matched-field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).
- ¹¹E. K. Westwood, "Broadband matched-field source localization," *J. Acoust. Soc. Am.* **91**, 2777–2789 (1992).
- ¹²S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science* **220**, 671–680 (1983).
- ¹³M. D. Collins, W. A. Kuperman, and H. Schmidt, "Nonlinear inversion for ocean-bottom properties," *J. Acoust. Soc. Am.* **92**, 2770–2783 (1992).
- ¹⁴C. E. Lindsay and N. R. Chapman, "Matched field inversion for geophysical parameters using adaptive simulated annealing," *IEEE J. Ocean Eng.* **18**, 224–231 (1993).
- ¹⁵P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and a posteriori probability distributions," *J. Acoust. Soc. Am.* **95**, 770–782 (1994).
- ¹⁶P. Gerstoft, "Inversion of acoustic data using a combination of genetic algorithms and the Gauss-Newton approach," *J. Acoust. Soc. Am.* **97**, 2181–2190 (1995).
- ¹⁷C. A. Zala and J. M. Ozard, "Estimation of geoacoustic parameters from narrowband data using a search-optimization technique," *J. Comput. Acoust.* **6**, 223–244 (1998).
- ¹⁸J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.* **7**, 308–313 (1965).
- ¹⁹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN* (Cambridge U.P., Cambridge, 1992).
- ²⁰A. Tolstoy, N. R. Chapman, and G. H. Brooke, "Workshop97: Benchmarking for geoacoustic inversion in shallow water," *J. Comput. Acoust.* **6**, 1–28 (1998).
- ²¹H. Schmidt, "SAFARI: Seismo-acoustic fast field algorithm for range-independent environments," Rep. SR-113 SACLANT ASW Research Center, La Spezia, Italy (1988).
- ²²R. B. Evans, "A coupled mode solution for acoustic propagation in a waveguide with stepwise depth variations of a penetrable bottom," *J. Acoust. Soc. Am.* **74**, 188–195 (1983).
- ²³"SWelLEX-96 Preliminary data report," unpublished report, 19 May 1996.
- ²⁴P. Baxley, private communication.
- ²⁵R. T. Bachman, P. W. Schey, N. O. Booth, and F. J. Ryan, "Geoacoustic databases for matched-field processing: Preliminary results in shallow water off San Diego," *J. Acoust. Soc. Am.* **99**, 2077–2085 (1996).

Dispersion and absorption of sound in monatomic gases: An extended kinetic description

Wilson Marques, Jr.^{a)}

Departamento de Física—Universidade Federal do Paraná, Caixa Postal 19081, 81531-990 Curitiba, Brazil

(Received 20 April 1999; accepted for publication 30 August 1999)

The problem concerning the propagation of plane harmonic waves of small amplitudes in monatomic ideal gases is studied using an extended kinetic model that replaces (in the same manner as in the Bhatnagar–Gross–Krook model) the Boltzmann collision operator with a single relaxation-time term. Numerical results for the speed and attenuation of sound are derived and compared with the experimental data of Meyer and Sessler, Greenspan, and Schotter for monatomic gases. At low frequencies, the correct continuum limits for the absorption and dispersion are recovered, while at high frequencies a complete agreement between theory and experiments is observed. © 1999 Acoustical Society of America. [S0001-4966(99)04912-7]

PACS numbers: 43.35.Ae [HEB]

INTRODUCTION

In the low oscillation frequency limit, the problem concerning the propagation of plane harmonic waves arising from an oscillating piston in a simple gas is correctly described by the classical Navier–Stokes and Fourier equations. In this case, sound dispersion is negligible and the absorption per wavelength is proportional to the frequency and can be written as a sum of contributions due to viscosity and thermal conduction. As sound frequency increases, the classical Navier–Stokes–Fourier theory is not able to describe the experimental data and, in particular, leads to an infinite phase velocity as the frequency goes to infinity.

To obtain a theory for high frequencies one has no recourse unless to turn to the Boltzmann equation of kinetic gas theory. A great number of attempts have been made to study the problem concerning sound propagation in kinetic theory. In particular, we mention the works of Wang Chang and Uhlenbeck,¹ Pekeris *et al.*,^{2,3} and Sirovich and Thurber.⁴ In the work of Wang Chang and Uhlenbeck, the sound propagation problem in monatomic gases is studied using the method of eigenfunctions for Maxwellian particles. The solution obtained by Wang Chang and Uhlenbeck proves to be more accurate than the Navier–Stokes–Fourier solution up to the intermediate frequency region. However, in the high-frequency region, it also diverges markedly from the experimental values. Using numerical procedures, Pekeris and his co-workers have applied the eigenfunctions method beyond normal belief to study acoustic oscillations in Maxwellian and hard spheres gases. Nevertheless, in the high-frequency limit, their theoretical results are quite poor. In the work of Sirovich and Thurber, sound propagation in a simple gas is studied using a kinetic model description based on the Gross–Jackson procedure.⁵ The results of Sirovich and Thurber are in close agreement with experiments in the low-frequency limit and in better agreement than the formers in the high-frequency limit.

Our aim in this paper is to study the sound propagation problem in monatomic ideal gases by using an extended kinetic model equation that replaces the Boltzmann collision operator with a single relaxation-time term. The kinetic model equation contains several position and time-dependent coefficients associated with the moments of the distribution function. The identification of these coefficients follows by the requirement that the collisional transfers of the moments be the same as the production terms of the Boltzmann equation for monatomic gases of Maxwellian particles. Using the normal mode method,⁶ we derive specific numerical results for the phase velocity and attenuation as a function of the so-called rarefaction parameter, i.e., the ratio of gas collision frequency to sound frequency. A comparison of the theoretical results with the experimental data of Meyer and Sessler,⁷ Greenspan,⁸ and Schotter⁹ for monatomic gases, shows that the extended kinetic model equation provides a precise transition between low- and high-frequency limits. Moreover, we conclude from this comparison that our kinetic model describes the experimental data in the low rarefaction region providing that the distance between the transmitter and the receiver is larger than the distance traveled by the gas particles during one cycle of the sound source.

The Cartesian notation for tensors with the usual summation convention is used, and angular parentheses around indices denote traceless symmetrization.

I. KINETIC MODEL

We consider a monatomic ideal gas in which the particles of mass m repel each other at all distances by means of a potential energy of the form $\kappa/4r^4$, where κ is the force constant. The state of the gas is characterized by the distribution function $f(\mathbf{r}, \mathbf{c}, t)$ such that $f(\mathbf{r}, \mathbf{c}, t) d\mathbf{r} d\mathbf{c}$ gives at time t the number of molecules in the volume element between \mathbf{r} and $\mathbf{r} + d\mathbf{r}$ and with linear velocities between \mathbf{c} and $\mathbf{c} + d\mathbf{c}$. The distribution function satisfies the Boltzmann equation that in the absence of external forces reads as

^{a)}Electronic mail: marques@fisica.ufpr.br

$$\frac{\partial f}{\partial t} + c_i \frac{\partial f}{\partial r_i} = \mathcal{C}(f), \quad (1)$$

where the Boltzmann collision operator $\mathcal{C}(f)$ is a functional of the distribution function and of the collision cross section. The collision operator obeys the following collisional invariant conditions,

$$\int m\mathcal{C}(f)d\mathbf{c}=0, \quad \int mc_i\mathcal{C}(f)d\mathbf{c}=0, \quad \int mc^2\mathcal{C}(f)d\mathbf{c}=0, \quad (2)$$

which express the conservation of mass, linear momentum, and energy, respectively.

The mathematical complexity of the Boltzmann equation to treat time-dependent problems like sound propagation in dilute monatomic gases can be avoided if the collision operator $\mathcal{C}(f)$ is replaced by a single relaxation-time term of the form

$$\mathcal{C}(f) = -\sigma(f - f_r), \quad (3)$$

where $\sigma = p/\mu$ is the stress relaxation frequency, p is the pressure, μ is the coefficient of shear viscosity, and f_r is a reference distribution function. The reference distribution function depends on the velocity of the particles and contains several position and time-dependent coefficients that are determined by keeping some of the main properties of the Boltzmann description for monatomic gases of Maxwellian particles.

Our aim in this section is to present a kinetic model description for monatomic ideal gases that is compatible with Grad's 35-moment approximation for monatomic gases of Maxwellian particles.¹⁰ In the 35-moment approximation the macroscopic state of a monatomic ideal gas is characterized by the following moments of the distribution function:

$$n = \int f d\mathbf{c}, \quad n v_i = \int c_i f d\mathbf{c}, \quad \frac{3}{2}nkT = \frac{1}{2} \int mC^2 f d\mathbf{c}, \quad (4)$$

$$p_{\langle ij \rangle} = \int mC_{\langle i}C_{j \rangle} f d\mathbf{c}, \quad q_i = \frac{1}{2} \int mC^2 C_i f d\mathbf{c}, \quad (5)$$

$$p_{\langle ijk \rangle} = \int mC_{\langle i}C_j C_k \rangle f d\mathbf{c}.$$

$$\Delta = \int mC^4 (f - f^{(0)}) d\mathbf{c},$$

$$p_{\langle ij \rangle rr} = \int mC_{\langle i}C_j \rangle C^2 f d\mathbf{c}, \quad (6)$$

$$p_{\langle ijkl \rangle} = \int mC_{\langle i}C_j C_k C_l \rangle f d\mathbf{c}.$$

In the above expressions n , v_i , and T represent the number density, the velocity, and the temperature of the gas, while $p_{\langle ij \rangle}$ and q_i represent the pressure deviator (traceless part of the pressure tensor) and the heat flux, respectively. The quantities Δ , $p_{\langle ijk \rangle}$, $p_{\langle ij \rangle rr}$, and $p_{\langle ijkl \rangle}$ are higher-order moments of the distribution function. Furthermore, $C_i = c_i - v_i$ is the peculiar velocity, k is the Boltzmann constant, and

$$f^{(0)} = n \left(\frac{\beta}{\pi} \right)^{3/2} \exp(-\beta C^2), \quad \left(\beta = \frac{m}{2kT} \right) \quad (7)$$

is the local Maxwell distribution function. For the above moment approximation we propose the following expression for the reference distribution function:

$$f_r = f^{(0)} \{ A + A_i C_i + A_{ij} C_i C_j + A_{ijk} C_i C_j C_k + A_{ijkl} C_i C_j C_k C_l \}, \quad (8)$$

where A , A_i , A_{ij} , A_{ijk} , and A_{ijkl} are position and time-dependent coefficients to be determined. A complete specification of these coefficients can be done, if we require that the collisional transfer of the moments is the same one as the production terms of the moments of the Boltzmann equation for monatomic gases of Maxwellian particles. Hence, together with (2) we impose the conditions

$$\int mC_i C_j \mathcal{C}(f) d\mathbf{c} = -\sigma p_{\langle ij \rangle}, \quad (9)$$

$$\int mC_i C^2 \mathcal{C}(f) d\mathbf{c} = -\frac{4}{3} \sigma q_i, \quad (10)$$

$$\int mC_{\langle i} C_j C_k \rangle \mathcal{C}(f) d\mathbf{c} = -\frac{3}{2} \sigma p_{\langle ijk \rangle}, \quad (11)$$

$$\int mC^4 \mathcal{C}(f) d\mathbf{c} = -\frac{2}{3} \sigma \Delta, \quad (12)$$

$$\int mC_{\langle i} C_j \rangle C^2 \mathcal{C}(f) d\mathbf{c} = -\frac{7}{6} \sigma \left(p_{\langle ij \rangle rr} - \frac{p}{\varrho} p_{\langle ij \rangle} \right), \quad (13)$$

$$\int mC_{\langle i} C_j C_k C_l \rangle \mathcal{C}(f) d\mathbf{c} = -\alpha \sigma p_{\langle ijkl \rangle}, \quad (14)$$

where $p = nkT$ is the thermodynamic pressure, $\varrho = mn$ is the mass density, and $\alpha \approx 1.8731$ is a pure number. Insertion of (3) and (8) on the left-hand side of Eqs. (2) and (9)–(14) leads to

$$A = 1 + \frac{1}{6} \frac{\beta^2}{\varrho} \Delta, \quad A_i = -\frac{4}{3} \frac{\beta^2}{\varrho} q_i, \quad A_{rr} = -\frac{2}{3} \frac{\beta^3}{\varrho} \Delta, \quad (15)$$

$$A_{\langle ij \rangle} = \frac{1}{3} \frac{\beta^2}{\varrho} \left\{ \beta p_{\langle ij \rangle rr} - \frac{7}{2} p_{\langle ij \rangle} \right\}, \quad (16)$$

$$A_{rri} = \frac{8}{9} \frac{\beta^3}{\varrho} q_i, \quad A_{\langle ijk \rangle} = -\frac{2}{3} \frac{\beta^3}{\varrho} p_{\langle ijk \rangle},$$

$$A_{rrss} = \frac{2}{9} \frac{\beta^4}{\varrho} \Delta,$$

$$A_{\langle ij \rangle rr} = -\frac{1}{9} \frac{\beta^3}{\varrho} \left\{ \beta p_{\langle ij \rangle rr} - \frac{7}{2} p_{\langle ij \rangle} \right\}, \quad (17)$$

$$A_{\langle ijkl \rangle} = \frac{2}{3} \frac{\beta^4}{\varrho} (1 - \alpha) p_{\langle ijkl \rangle}.$$

In summary, the extended kinetic model presented in this paper for monatomic ideal gases is defined by the equation

$$\frac{\partial f}{\partial t} + c_i \frac{\partial f}{\partial x_i} = -\sigma(f - f_r), \quad (18)$$

where the reference distribution function f_r is given by Eqs. (8) and (15)–(17).

We call attention to the fact that an extension of our kinetic model for more realistic interaction potentials (e.g., Lennard-Jones 6–12 potential) can be easily be done. However, for time-dependent problems like sound propagation⁴ and light scattering¹¹ in monatomic gases, the interaction potential does not play an important role.

Closing this section, we remark that the expression for the reference distribution function changes in the 13-moment and 20-moment approximations. The expressions for the reference distribution function in these approximations are given in Appendix A. The kinetic model that corresponds to the 13-moment approximation is the same as that proposed by Shakhov¹² and gives the correct value for the Prandtl number. Moreover, if the reference function is the local Maxwell distribution, the BGK model¹³ is recovered.

II. DISPERSION RELATION

We want to study now the problem concerning the propagation of plane harmonic waves of small amplitudes in monatomic gases using the extended kinetic model presented in the previous section. In order to describe the propagation of such waves, it is appropriate to assume the following expansion for the distribution function:

$$f(\mathbf{r}, \mathbf{c}, t) = n_0 f_0 \{1 + h(\mathbf{c}) \exp[i(\kappa x - \omega t)]\}, \quad (19)$$

where

$$f_0 = \left(\frac{\beta_0}{\pi}\right)^{3/2} \exp(-\beta_0 c^2). \quad (20)$$

In Eq. (19), ω is the angular frequency of the forced wave, $\kappa = \kappa^r + i\kappa^i$ is the complex wave number, x is the direction of propagation of the wave, and $h(\mathbf{c})$ is the amplitude of the perturbation. By inserting expression (19) into the kinetic model equation (18) and keeping only linear terms, we obtain

$$(c_x - \zeta)h(\mathbf{c}) = -i \frac{\sigma}{\kappa} \mathcal{L}, \quad (21)$$

where $\zeta = (\omega + i\sigma)/\kappa$ and

$$\begin{aligned} \mathcal{L} = & \bar{n} + 2\sqrt{\beta_0} c_i \bar{v}_i + (\beta_0 c^2 - \frac{3}{2}) \bar{T} \\ & + \frac{1}{6} (\beta_0 c^2 - \frac{7}{2}) \beta_0 c_i c_j \bar{p}_{\langle ij \rangle} \\ & + \frac{4}{15} (\beta_0 c^2 - \frac{5}{2}) \sqrt{\beta_0} c_i \bar{q}_i - \frac{1}{3} \beta_0^{3/2} c_i c_j c_k \bar{p}_{\langle ijk \rangle} \\ & + \frac{1}{45} (\beta_0^2 c^4 - 5\beta_0 c^2 + \frac{15}{4}) \bar{\Delta} - \frac{1}{21} (\beta_0 c^2 - \frac{7}{2}) \\ & \times \beta_0 c_i c_j \bar{p}_{\langle ij \rangle rr} + \frac{(1-\alpha)}{3} \beta_0^2 c_i c_j c_k c_l \bar{p}_{\langle ijkl \rangle}. \end{aligned} \quad (22)$$

The dimensionless amplitudes of the perturbations of the macroscopic moments that appear in the above expression are given by

$$\bar{n} = \int f_0 h(\mathbf{c}) d\mathbf{c}, \quad \bar{v}_i = \int \sqrt{\beta_0} c_i f_0 h(\mathbf{c}) d\mathbf{c}, \quad (23)$$

$$\bar{T} = \frac{2}{3} \int \left(\beta_0 c^2 - \frac{3}{2} \right) f_0 h(\mathbf{c}) d\mathbf{c},$$

$$\bar{p}_{\langle ij \rangle} = 2 \int \beta_0 c_{\langle i} c_{j \rangle} f_0 h(\mathbf{c}) d\mathbf{c}, \quad (24)$$

$$\bar{q}_i = \int \left(\beta_0 c^2 - \frac{5}{2} \right) \sqrt{\beta_0} c_i f_0 h(\mathbf{c}) d\mathbf{c},$$

$$\bar{p}_{\langle ijk \rangle} = 2 \int \beta_0^{3/2} c_{\langle i} c_j c_{k \rangle} f_0 h(\mathbf{c}) d\mathbf{c}, \quad (25)$$

$$\bar{\Delta} = 2 \int \left(\beta_0 c^4 - 5\beta_0 c^2 + \frac{15}{4} \right) f_0 h(\mathbf{c}) d\mathbf{c},$$

$$\bar{p}_{\langle ij \rangle rr} = 2 \int \beta_0^2 c_{\langle i} c_j \rangle c^2 f_0 h(\mathbf{c}) d\mathbf{c}, \quad (26)$$

$$\bar{p}_{\langle ijkl \rangle} = 2 \int \beta_0^2 c_{\langle i} c_j c_k c_l \rangle f_0 h(\mathbf{c}) d\mathbf{c}.$$

Since the amplitudes of the perturbations of the macroscopic moments (23)–(26) are independent of the particle velocity \mathbf{c} , the dependence of the right-hand side of Eq. (21) upon \mathbf{c} is explicit, and we can therefore obtain an expression for $h(\mathbf{c})$ by dividing (21) by $(c_x - \zeta)$. However, if ζ is real this factor will vanish for certain values of c_x , and we are not allowed to divide by zero. Hence, we must distinguish two cases, depending upon whether ζ is real or not. If ζ is real, a set of linear inhomogeneous equations for the amplitudes of the perturbations is obtained. The solution of this set of equations gives rise to the so-called continuous (or particle) modes. For these modes we do not have a dispersion relation, since there is no functional relation between ω and κ . If ζ is not real, then the solution of Eq. (21) is

$$h(\mathbf{c}) = -i \frac{\sigma}{\kappa} \frac{\mathcal{L}}{c_x - \zeta}. \quad (27)$$

The amplitudes of the perturbations are determined by inserting Eq. (27) into Eqs. (23)–(26). The result in this case is a set of linear homogeneous equations, which is best expressed in the matrix form

$$(\mathbf{M} - i(\kappa v_0 / \omega) \mathbf{I}) \mathbf{X} = 0, \quad (28)$$

where \mathbf{I} is the identity matrix, $v_0 = \sqrt{2kT_0/m}$ is the equilibrium thermal velocity, and

$$\mathbf{X} = \begin{pmatrix} \bar{n} \\ \bar{v}_x \\ \bar{T} \\ \bar{p}_{\langle xx \rangle} \\ \bar{q}_x \\ \bar{\Delta} \\ \bar{p}_{\langle xxx \rangle} \\ \bar{p}_{\langle xx \rangle rr} \\ \bar{p}_{\langle xxx \rangle} \end{pmatrix}. \quad (29)$$

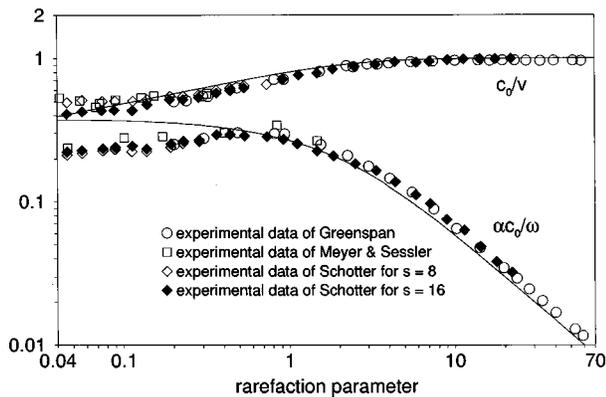


FIG. 1. Phase velocity and attenuation coefficient in the BGK approximation.

The elements of the matrix \mathbf{M} (listed in Appendix B) depend on the dimensionless parameter $z = (1 + ir)(\omega/\kappa v_0)$ and on the plasma dispersion function,¹⁴

$$W(z) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{\exp(-t^2)}{t-z} dt, \quad (30)$$

where $r = \sigma/\omega$ is the rarefaction parameter. We note that the plasma dispersion function $W(z)$ is an analytic function, except on the real axis of the z plane.

The condition for the existence of a plane wave solution is that the dispersion relation vanish, i.e.,

$$\det(\mathbf{M} - i(\kappa v_0/\omega)\mathbf{I}) = 0. \quad (31)$$

Following the same procedure described in details in the literature, the dispersion relation is solved for $\text{Im}(z) > 0$, i.e., up to the cut of the plasma dispersion function $W(z)$ on the real axis in the z plane. This leads to a critical rarefaction parameter r_c (or critical frequency ω_c) below which the collective sound mode ceases to exist. The critical rarefaction parameter (or critical frequency) is determined by the condition $\text{Im}(z) = 0$, i.e., when $\kappa^i/\kappa^r = \sigma/\omega_c$. This criterion for the critical frequency can be put into a more familiar term such as the phase velocity $v = \omega/\kappa^r$ and the attenuation coefficient $\alpha = \kappa^i$. Thus, the criterion for the critical frequency can be stated as $(\alpha v/\omega)_c = r_c$. When $\text{Im}(z) < 0$, the analytic continuation of the plasma dispersion function is used in the dispersion relation. This leads to the analytic continuations of the

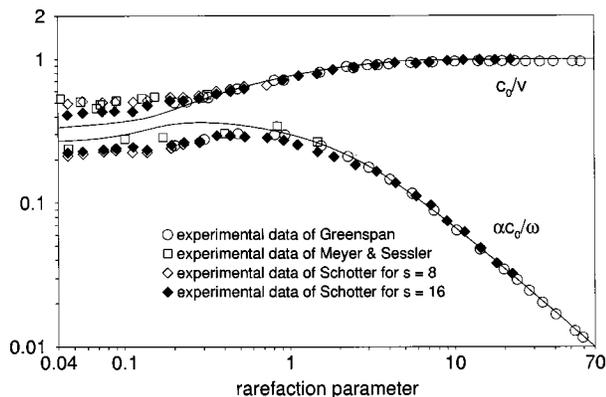


FIG. 2. Phase velocity and attenuation coefficient in the 13-moment approximation.

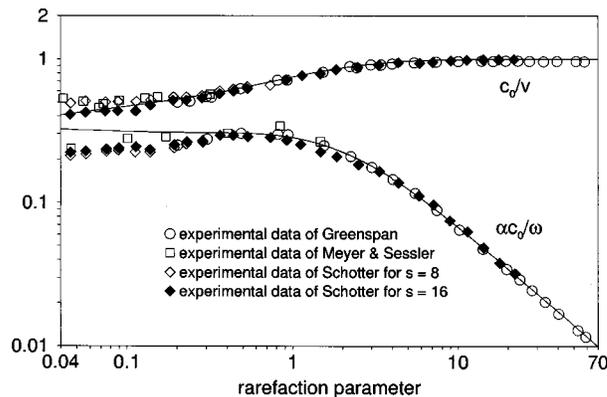


FIG. 3. Phase velocity and attenuation coefficient in the 20-moment approximation.

phase velocity and attenuation coefficient below the critical frequency.

We close this section by examining the roots of the dispersion relation in the low-frequency limit, i.e., when $r \rightarrow \infty$. In this limit, we see that $|z|$ is large and we can use the following asymptotic expansion for the plasma dispersion function:

$$W(z) \sim -\frac{1}{\sqrt{\pi}} \sum_{n=0}^{\infty} \Gamma(n+1/2) z^{-(2n+1)}. \quad (32)$$

Expanding the dimensionless wave number $(\kappa v_0/\omega)$ in powers of $1/r$ and retaining terms up to first order, we get

$$v = c_0 \quad \text{and} \quad \alpha = \frac{7}{6} \frac{\mu \omega^2}{\rho c_0^3}, \quad (33)$$

where $c_0 = \sqrt{5kT_0/3m}$ is the adiabatic sound velocity. The above expressions for the phase velocity and attenuation coefficient agree with the sound propagation results obtained from the Navier–Stokes equations.

III. NUMERICAL RESULTS

In order to calculate the discrete sound mode in the transition region we must solve the dispersion relation numerically. Below the critical frequency (or critical rarefaction parameter), the analytic continuation of the dispersion relation

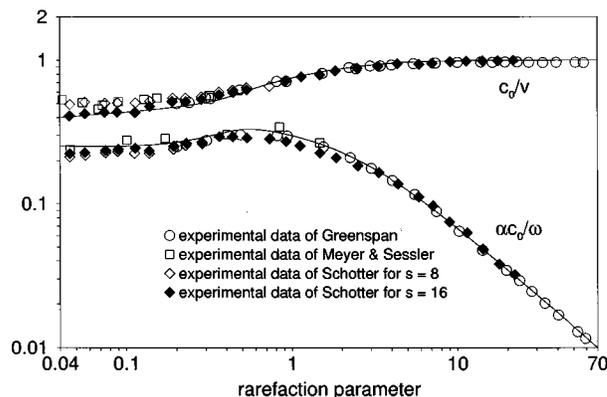


FIG. 4. Phase velocity and attenuation coefficient in the 35-moment approximation.

is used to extend the discrete sound mode solutions to negative $\text{Im}(z)$. The value of the critical rarefaction parameter $r_c = \sigma/\omega_c$ for the 35-moment approximation is 0.538. Given an initial guess, the roots of the dispersion relation are determined by using a multidimensional secant method.¹⁵ As an initial guess, we use the roots of the dispersion relation in the low-frequency limit.

Figures 1, 2, 3, and 4 show on a logarithmic scale the normalized attenuation coefficient $\alpha c_0/\omega$ and the normalized reciprocal velocity c_0/v versus the rarefaction parameter r for the BGK, 13-moment, 20-moment, and 35-moment approximation, respectively. The figures also show the experimental results of Meyer and Sessler⁷ for Ar and the experimental results of Greenspan⁸ and Schotter⁹ for He. The measurements of Meyer and Sessler, and Greenspan were made at room temperature and constant sound frequency for different values of the gas pressure. Greenspan employed a quartz crystal oscillator and receiver at a frequency of 11 MHz, while Meyer and Sessler used an electrostatic condenser transmitter and receiver at frequencies of 100 and 200 kHz. The attenuation of sound was obtained by a determination of the logarithmic decrement in the sound level of the signal as a function of the sound path by varying the distance between transmitter and receiver. The sound speed measurement was obtained by measuring the phase difference between a direct signal from the driving oscillator and the signal received at the receiver as a function of sound path. In the work of Meyer and Sessler, the distance x between transmitter and receiver for which the experiments were made was such that $\omega x = 3.0 \times 10^3$ m/s. Schotter used in his work a high barium titanate sandwich transducers and narrow-band phase-sensitive signal detection to investigate sound propagation in the noble gases He, Ne, Ar, and Kr. For helium, the sound propagation experiments of Schotter have been made for propagation distances such that $\omega x < 15.7 \times 10^3$ m/s at collision frequencies $\sigma > 0.002\omega$.

To discuss the sound propagation results derived from our extended kinetic model equation we introduce the dimensionless parameter,

$$s = \sqrt{\frac{5}{3}} \frac{\omega x}{c_0}, \quad (34)$$

which represents the ratio of the source to receiver distance to the mean distance traveled by a gas particle during one cycle of the sound source. In the low-frequency limit, the absorption and dispersion functions are independent of the propagation distances and binary collisions are much more frequent than wall collisions. In the high-frequency limit, absorption and dispersion become strongly dependent upon the propagation distances and sound propagation is not only governed by binary collisions.

As we can see from Figs. 1, 2, 3, and 4, the numerical results derived from the 13-moment, 20-moment, and 35-moment approximations for the attenuation and dispersion are in complete agreement with the experimental data in the low-frequency limit. In this limit, the attenuation curve derived from the BGK model is slightly displaced from the other curves. This divergence arises from the fact that the

BGK model does not yield the correct value for the Prandtl number.

With respect to the velocity ratio c_0/v in the small rarefaction region ($r < 1$), we see that the 13-moment, 20-moment, and 35-moment approximations are in good agreement with the experimental data for large values of the parameter s . However, the prediction derived from the 35-moment approximation is decidedly closer to the experiments for small values of the rarefaction parameter r . For the attenuation curves, we observe that in the so-called transition region ($1 < r < 7$) the theoretical predictions are in good agreement with the experiments. In the small rarefaction region, we see that the 35-moment approximation gives the best fit with the experimental data. Moreover, we also observe that the dimensionless parameter s (the ratio of the source to receiver distance to the mean distance traveled by a gas particle during one cycle of the sound source) does not affect the sound attenuation in the high-frequency limit.

IV. CONCLUSIONS

In this paper a kinetic model equation that is compatible with Grad's 35-moment approximation have been used to study the problem concerning the propagation of plane sound waves of small amplitudes in monatomic ideal gases. The sound propagation results derived from this model are in remarkably good agreement with the experimental results of Meyer and Sessler, Greenspan, and Schotter, showing that the 35-moment approximation gives a precise transition between low- and high-frequency limits, providing that in the low rarefaction region the distance between transmitter and receiver is larger than the distance traveled by the gas particles during one cycle of the sound source.

Last, we call attention to the fact that the 35-moment kinetic model equation has no free adjustable parameters and that it gives the correct value for the Prandtl number.

ACKNOWLEDGMENTS

The author would like to thank Sergio R. Lopes and Marcio H. F. Bettega for several discussions concerning numerical computation.

APPENDIX A

The values of the coefficients A , A_i , $A_{\langle ij \rangle}$, A_{rr} , $A_{\langle ijk \rangle}$, A_{rri} , $A_{\langle ijkl \rangle}$, $A_{\langle ij \rangle rr}$, and A_{rrss} in 13-moment and 20-moment approximations are the following: (i) 13-moment approximation:

$$A = 1, \quad A_i = -\frac{4}{3} \frac{\beta^2}{\varrho} q_i, \quad A_{rri} = \frac{8}{9} \frac{\beta^3}{\varrho} q_i, \quad (A1)$$

$$A_{rr} = A_{\langle ij \rangle} = A_{\langle ijk \rangle} = A_{rrss} = A_{\langle ij \rangle rr} = A_{\langle ijkl \rangle} = 0; \quad (A2)$$

(ii) 20-moment approximation:

$$A = 1, \quad A_i = -\frac{4}{3} \frac{\beta^2}{\rho} q_i, \quad A_{rri} = \frac{8}{9} \frac{\beta^3}{\rho} q_i, \quad (A3)$$

$$A_{\langle ijk \rangle} = -\frac{2}{3} \frac{\beta^3}{\rho} P_{\langle ijk \rangle},$$

$$A_{rr} = A_{\langle ij \rangle} = A_{rrss} = A_{\langle ij \rangle rr} = A_{\langle ijk \rangle l} = 0. \quad (A4)$$

APPENDIX B

The elements of matrix \mathbf{M} occurring in Eq. (28) are

$$M_{11} = W(z), \quad M_{12} = 2(zW(z) + 1), \quad (B1)$$

$$M_{13} = (z^2 - \frac{1}{2})W(z) + z,$$

$$M_{14} = \frac{1}{6}[(z^4 - 3z^2 + \frac{3}{4})W(z) + z^3 - \frac{5}{2}z], \quad (B2)$$

$$M_{15} = \frac{4}{15}[z(z^2 - \frac{3}{2})W(z) + z^2 - 1].$$

$$M_{16} = \frac{2}{15}M_{14}, \quad M_{17} = -\frac{5}{4}M_{15}, \quad (B3)$$

$$M_{18} = -\frac{2}{7}M_{14}, \quad M_{19} = \frac{75}{35}(1 - \alpha)M_{14},$$

$$M_{21} = \frac{M_{12}}{2}, \quad M_{2i} = zM_{1i} \quad (i = 2, 3, \dots, 9), \quad (B4)$$

$$M_{31} = \frac{2}{3}M_{13}, \quad M_{32} = \frac{4}{3}zM_{13}, \quad (B5)$$

$$M_{33} = \frac{2}{3} \left[\left(z^4 - z^2 + \frac{5}{4} \right) W(z) + z^3 - \frac{z}{2} \right].$$

$$M_{34} = \frac{1}{9}[(z^6 - \frac{7}{2}z^4 + \frac{11}{4}z^2 - \frac{5}{8})W(z) + z^5 - 3z^3 + \frac{7}{4}z], \quad (B6)$$

$$M_{35} = \frac{8}{45}[z(z^4 - 2z^2 + \frac{7}{4})W(z) + z^4 - \frac{3}{2}z^2 + \frac{3}{2}], \quad (B7)$$

$$M_{36} = \frac{2}{135}[(z^6 - \frac{7}{2}z^4 + \frac{17}{4}z^2 - \frac{11}{8})W(z) + z^5 - 3z^3 + \frac{13}{4}z], \quad (B8)$$

$$M_{37} = -\frac{2}{9}[z(z^4 - 2z^2 - \frac{3}{4})W(z) + z^4 - \frac{3}{2}z^2 - 1], \quad (B9)$$

$$M_{38} = -\frac{2}{7}M_{34},$$

$$M_{39} = \frac{8}{35}(1 - \alpha)[(z^6 - \frac{7}{2}z^4 - \frac{3}{4}z^2 + \frac{9}{8})W(z) + z^5 - 3z^3 - \frac{7}{4}z], \quad (B10)$$

$$M_{41} = \frac{4}{3}M_{13}, \quad M_{42} = \frac{8}{3}zM_{13}, \quad (B11)$$

$$M_{43} = \frac{4}{3} \left[\left(z^4 - z^2 - \frac{1}{4} \right) W(z) + z^3 - \frac{z}{2} \right],$$

$$M_{44} = \frac{2}{9}[(z^6 - \frac{7}{2}z^4 + 2z^2 - \frac{1}{4})W(z) + z^5 - 3z^3 + z], \quad (B12)$$

$$M_{45} = \frac{16}{45}z[(z^4 - 2z^2 + \frac{1}{4})W(z) + z^3 - \frac{3}{2}z], \quad (B13)$$

$$M_{46} = \frac{4}{135} \left[\left(z^6 - \frac{7}{2}z^4 + \frac{5}{4}z^2 + \frac{1}{8} \right) W(z) + z^5 - 3z^3 + \frac{z}{4} \right], \quad (B14)$$

$$M_{47} = -\frac{4}{9}[z(z^4 - 2z^2 + \frac{3}{2})W(z) + z^4 - \frac{3}{2}z^2 + \frac{5}{4}], \quad (B15)$$

$$M_{48} = -\frac{2}{7}M_{44},$$

$$M_{49} = \frac{16}{35}(1 - \alpha)[(z^6 - \frac{7}{2}z^4 + \frac{15}{4}z^2 - \frac{9}{8})W(z) + z^5 - 3z^3 + \frac{11}{4}z], \quad (B16)$$

$$M_{51} = \frac{15}{4}M_{15}, \quad M_{52} = 2zM_{51}, \quad M_{53} = \frac{45}{8}M_{35}, \quad (B17)$$

$$M_{54} = \frac{1}{6}z[(z^6 - \frac{9}{2}z^4 + \frac{23}{4}z^2 - \frac{11}{8})W(z) + z^5 - 4z^3 + \frac{17}{4}z], \quad (B18)$$

$$M_{55} = \frac{4}{15}z[z(z^4 - 3z^2 + \frac{13}{4})W(z) + z^4 - \frac{5}{2}z^2 + \frac{5}{2}], \quad (B19)$$

$$M_{56} = \frac{1}{45}z[(z^6 - \frac{9}{2}z^4 + \frac{29}{4}z^2 - \frac{17}{8})W(z) + z^5 - 4z^3 + \frac{23}{4}z], \quad (B20)$$

$$M_{57} = -2z^2M_{14},$$

$$M_{58} = -\frac{2}{7}M_{54}, \quad (B21)$$

$$M_{59} = \frac{12}{35}(1 - \alpha)z[(z^6 - \frac{9}{2}z^4 + \frac{9}{4}z^2 + \frac{3}{8})W(z) + z^5 - 4z^3 + \frac{3}{4}z],$$

$$M_{61} = 12M_{14}, \quad M_{62} = 2zM_{61}, \quad M_{63} = 135M_{36}, \quad (B22)$$

$$M_{65} = 24M_{56}, \quad M_{68} = -\frac{2}{7}M_{64},$$

$$M_{64} = \frac{1}{3}[(z^8 - 6z^6 + \frac{23}{2}z^4 - \frac{11}{2}z^2 - \frac{19}{16})W(z) + z^7 - \frac{11}{2}z^5 + \frac{37}{4}z^3 - \frac{19}{8}z], \quad (B23)$$

$$M_{66} = \frac{2}{45}[(z^8 - 6z^6 + \frac{29}{2}z^4 - \frac{17}{2}z^2 + \frac{89}{16})W(z) + z^7 - \frac{11}{2}z^5 + \frac{49}{4}z^3 - \frac{31}{8}z], \quad (B24)$$

$$M_{67} = -\frac{2}{3}z[(z^6 - \frac{9}{2}z^4 + \frac{9}{4}z^2 + \frac{3}{8})W(z) + z^5 - 4z^3 + \frac{3}{4}z], \quad (B25)$$

$$M_{69} = \frac{24}{35}(1 - \alpha)[(z^8 - 6z^6 + \frac{9}{2}z^4 + \frac{3}{2}z^2 + \frac{9}{16})W(z) + z^7 - \frac{11}{2}z^5 + \frac{9}{4}z^3 + \frac{9}{8}z], \quad (B26)$$

$$M_{71} = 3M_{15}, \quad M_{72} = 2zM_{71}, \quad M_{73} = -\frac{18}{5}M_{37}, \quad (B27)$$

$$M_{74} = \frac{24}{5}z[(z^6 - \frac{9}{2}z^4 + \frac{9}{2}z^2 - \frac{3}{4})W(z) + z^5 - 4z^3 + 3z], \quad (B28)$$

$$M_{75} = \frac{32}{25}z^2M_{14}, \quad (B29)$$

$$M_{76} = \frac{4}{225}z[(z^6 - \frac{9}{2}z^4 + \frac{9}{4}z^2 + \frac{3}{8})W(z) + z^5 - 4z^3 + \frac{3}{4}z].$$

$$M_{77} = -\frac{12}{5}z[z(z^4 - 3z^2 + \frac{9}{2})W(z) + z^4 - \frac{5}{2}z^2 + \frac{15}{4}], \quad (B30)$$

$$M_{78} = -\frac{2}{7}M_{74}.$$

$$M_{79} = \frac{48}{175}(1 - \alpha)z[(z^6 - \frac{9}{2}z^4 + \frac{39}{4}z^2 - \frac{27}{8})W(z) + z^5 - 4z^3 + \frac{33}{4}z], \quad (B31)$$

$$M_{81} = \frac{4}{3} \left[\left(z^4 + \frac{z^2}{2} - 1 \right) W(z) + z^3 + z \right], \quad M_{82} = 2zM_{81}, \quad (B32)$$

$$M_{83} = \frac{4}{3} \left[\left(z^6 - \frac{3}{4}z^2 - \frac{3}{2} \right) W(z) + z^5 + \frac{z^3}{2} \right], \quad (B33)$$

$$M_{84} = 8[(z^8 - \frac{5}{2}z^6 - \frac{3}{2}z^4 + \frac{9}{4}z^2 + \frac{3}{4})W(z) + z^7 - 2z^5 - 2z^3 + \frac{3}{2}z], \quad (B34)$$

$$M_{85} = \frac{16}{45}z \left[\left(z^6 - z^4 - \frac{5}{4}z^2 - \frac{1}{2} \right) W(z) + z^5 - \frac{z^3}{2} - z \right], \quad (B35)$$

$$M_{86} = \frac{4}{135}[(z^8 - \frac{5}{2}z^6 - \frac{3}{4}z^4 - \frac{9}{8}z^2 - \frac{3}{4})W(z) + z^7 - 2z^5 - \frac{5}{4}z^3 - \frac{3}{2}z], \quad (\text{B36})$$

$$M_{87} = -\frac{4}{9}\left[z\left(z^6 - z^4 - \frac{5}{2}z^2 + \frac{9}{2}\right)W(z) + z^6 - \frac{z^4}{2} - \frac{9}{4}z^2 + \frac{35}{8}\right], \quad (\text{B37})$$

$$M_{88} = -\frac{2}{7}M_{84},$$

$$M_{89} = \frac{16}{35}(1 - \alpha)\left[(z^8 - \frac{5}{2}z^6 - \frac{13}{4}z^4 + \frac{81}{8}z^2 - \frac{9}{2})W(z) + z^7 - 2z^5 - \frac{15}{4}z^3 + \frac{17}{2}z\right]. \quad (\text{B38})$$

$$M_{91} = \frac{96}{35}M_{14}, \quad M_{92} = 2zM_{14}, \quad (\text{B39})$$

$$M_{93} = \frac{48}{105}\left[(z^6 - \frac{7}{2}z^4 - \frac{3}{4}z^2 + \frac{9}{8})W(z) + z^5 - 3z^3 - \frac{7}{4}z\right],$$

$$M_{94} = \frac{8}{105}\left[(z^8 - 6z^6 + 9z^4 - 3z^2 - \frac{9}{16})W(z) + z^7 - \frac{11}{2}z^5 + \frac{27}{4}z^3 - \frac{9}{8}z\right], \quad (\text{B40})$$

$$M_{95} = \frac{48}{7}M_{76},$$

$$M_{96} = \frac{16}{1575}\left[(z^8 - 6z^6 + \frac{9}{2}z^4 + \frac{3}{2}z^2 + \frac{9}{16})W(z) + z^7 - \frac{11}{2}z^5 + \frac{9}{4}z^3 + \frac{9}{8}z\right], \quad (\text{B41})$$

$$M_{97} = -\frac{16}{105}z\left[(z^6 - \frac{9}{2}z^4 + \frac{39}{4}z^2 - \frac{27}{8})W(z) + z^5 - 4z^3 + \frac{33}{4}z\right], \quad (\text{B42})$$

$$M_{98} = -\frac{2}{7}M_{94}.$$

$$M_{99} = \frac{192}{1225}(1 - \alpha)\left[(z^8 - 6z^6 + \frac{39}{2}z^4 - \frac{27}{2}z^2 + \frac{27}{8})W(z) + z^7 - \frac{11}{2}z^5 + \frac{69}{4}z^3 - \frac{51}{8}z\right]. \quad (\text{B43})$$

¹C. S. Wang Chang and G. E. Uhlenbeck, in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland, Amsterdam, 1970), Vol. V, pp. 43–75.

²C. L. Pekeris, Z. Alterman, and L. Finkelstein, in *Symposium on the Numerical Treatment of Ordinary Differential Equations, Integral and Integro-Differential Equation of P.I.C.C.* (Birkhäuser-Verlag, Basel, 1960), pp. 388–398.

³C. L. Pekeris, Z. Alterman, L. Finkelstein, and K. Frankowski, *Phys. Fluids* **5**, 1608–1616 (1962).

⁴L. Sirovich and J. K. Thurber, *J. Acoust. Soc. Am.* **37**, 329–339 (1965).

⁵E. P. Gross and E. A. Jackson, *Phys. Fluids* **2**, 432–441 (1959).

⁶G. E. Uhlenbeck and G. W. Ford, in *Lectures in Statistical Mechanics* (American Mathematical Society, Providence, RI, 1963), pp. 89–101.

⁷E. Meyer and G. Sessler, *Z. Phys.* **149**, 15–39 (1957).

⁸M. Greenspan, *J. Acoust. Soc. Am.* **28**, 644–648 (1956).

⁹R. Schotter, *Phys. Fluids* **8**, 1163–1168 (1974).

¹⁰E. Ikenberry and C. Truesdell, *J. Rat. Mech. Anal.* **5**, 1–54 (1956).

¹¹W. Marques, Jr. and G. M. Kremer, *Continuum Mech. Thermodyn.* **10**, 319–329 (1998).

¹²E. M. Shakhov, *Fluid Dyn.* **3**, 95–99 (1968).

¹³P. L. Bhatnagar, E. P. Gross, and M. Krook, *Phys. Rev.* **94**, 511–525 (1954).

¹⁴B. D. Fried and S. D. Conte, *The Plasma Dispersion Function* (Academic, New York, 1961).

¹⁵W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran* (Cambridge University Press, New York, 1992).

Internal circulation in a drop in an acoustic field

Hong Zhao and S. S. Sadhal

Aerospace & Mechanical Engineering Department, University of Southern California, Los Angeles, California 90089-1453

Eugene H. Trinh

NASA Headquarters, 300 E Street S.W., Washington, DC 20546

(Received 10 November 1998; accepted for publication 26 July 1999)

An investigation of the internal flow field for a drop at the antinode of a standing wave has been carried out. The main difference from the solid sphere case is the inclusion of the shear stress and velocity continuity conditions at the liquid–gas interface. To the leading order of calculation, the internal flow field was found to be quite weak. Also, this order being fully time dependent has a zero mean flow. At the next higher order, steady internal flows are predicted and, as in the case of a solid sphere, there is a recirculating layer consisting of closed streamlines near the surface. In the case of a liquid drop, however, the behavior of this recirculating Stokes layer is quite interesting. It is predicted that the layer ceases to have recirculation when $|M| > \frac{5}{2}\sqrt{2}[2 + 5(\hat{\mu}/\mu)]$, where $\hat{\mu}$ is the liquid viscosity, μ is the exterior gas-phase viscosity, and M is the dimensionless frequency parameter for the gas phase, defined by $M = i\omega a^2 \rho/\mu$, with a being the drop radius. Thorough experimental confirmation of this interesting new development needs to be conducted. Although it seems to agree with many experiments with levitated drops where no recirculating layer has been clearly observed, a new set of experiments for specifically testing this interesting development need to be carried out. © 1999 Acoustical Society of America. [S0001-4966(99)05511-3]

PACS numbers: 43.35.Bf [HEB]

NOMENCLATURE

U_∞	far field velocity of the flow
a	radius of the sphere
r	radial coordinate
R	Reynolds number
M	frequency parameter
ψ	outer stream function
Ψ	stream function inside the shear-wave layer
$\hat{\psi}$	stream function inside the sphere
k	wave number
θ	polar angle

τ	dimensionless time (ωt)
$\bar{\mu}$	$\cos \theta$
ν	kinematic viscosity
ω	angular frequency of the wave
$\tau_{r\theta}$	shear stress
μ	dynamic viscosity
η	inner variable inside the shear layer
κ	$\mu/\hat{\mu}$
$(\hat{\quad})$	liquid phase quantities
∞	far field (subscript)

INTRODUCTION

Containerless technology has been well recognized as a method for obtaining liquid undercooling. Particularly in microgravity, a liquid can remain undisturbed and experience deep undercooling. As pointed out by Zhao *et al.*,¹ a high degree of undercooling can promote certain types of crystal growth and at the same time provide homogeneity of the product. These kinds of new materials have higher performance and can lead to improved fuel efficiency in automotive and gas-turbine technology at high temperature. They also have very high melting points and can be used as semiconductors, refractory materials, high performance coatings and excellent infrared optical materials. In this regard, there is a strong interest in understanding the basic thermodynamics of such processes as well as the measurement of properties such as thermal conductivity and thermal diffusivity.^{2,3}

One of the major recent advances for experiments in containerless processing is acoustic levitation. The acoustic

field provides the radiation pressure necessary to levitate a liquid drop in a gravitational field. This phenomenon has been known for a long time (see Refs. 4 and 5) and many subsequent investigations on spherical particles^{6–10} have been carried out. There has also been some interest for non-spherical particles, such as disks, which experience a torque¹¹ when the orientation is oblique to the standing wave. In addition, there have some investigations on the thermal effects.^{12,13} In relation to microgravity, acoustic levitation is a good way to stabilize a levitated drop. There has been considerable experimental activity in this area,^{14,15} together with analytical studies such as that by Lee and Wang.¹⁶ While there have been several earlier analytical studies dealing with streaming flows about oscillating spherical particles,^{17–20} little has been considered about a liquid sphere being placed at the velocity antinode of a standing wave in acoustic flow.

The purpose of the present investigation is to obtain the

flow fields inside and outside of the liquid drop placed at the velocity antinode of a standing wave with velocity $u_z = U_\infty \cos kze^{i\omega t}$. Such an analysis will provide information about the characteristics of the levitation process. The theory will be very useful for overcoming some of the experimental problems by providing suitable new directions. For levitation under zero-gravity conditions, the drop takes on an equilibrium position at the velocity antinode when the external medium is a gas. In the present development, the focus is on the analysis of such a problem. The results from this analysis will be useful for dealing with a more complex problem concerning the streaming flows when the sphere takes a position between the velocity node and antinode of the wave. This happens when a liquid sphere is levitated in gravitational field.

There is an interesting result from the present calculations. We find that the recirculation in the shear-wave layer vanishes when the frequency parameter M increases to a critical value. This phenomenon needs to be explained and experimental verification is required.

I. PROBLEM STATEMENT

A standing wave of velocity $u_z = U_\infty \cos kze^{i\omega t}$ is considered, and a liquid drop of radius a is placed at the velocity antinode, $z=0$. The flow fields in the drop and the external gaseous medium are considered to be at low Mach number and may be described by the Navier-Stokes equation for incompressible flow. With the scales of the velocities, stream function, radial distance and time chosen as U_∞ , $U_\infty a^2$, a and ω^{-1} , respectively, the dimensionless equation satisfied by the stream function, $\psi(r, \theta)$, in a spherical coordinate system is (see Ref. 21)

$$|M|^2 \frac{\partial(D^2\psi)}{\partial\tau} + \hat{R} \left\{ \frac{1}{r^2} \frac{\partial(\psi, D^2\psi)}{\partial(r, \bar{\mu})} + \frac{2}{r^2} D^2\psi L\psi \right\} = D^4\psi, \quad (1)$$

where

$$D^2 = \frac{\partial^2}{\partial r^2} + \frac{(1-\bar{\mu}^2)}{r^2} \frac{\partial^2}{\partial \bar{\mu}^2}$$

and

$$L = \frac{\bar{\mu}}{(1-\bar{\mu}^2)} \frac{\partial}{\partial r} + \frac{1}{r} \frac{\partial}{\partial \bar{\mu}}.$$

Here, the Reynolds number R and frequency parameter M are defined as

$$R = U_\infty a / \nu \quad \text{and} \quad M^2 = i\omega a^2 / \nu,$$

where we have chosen $M\sqrt{2}/|M| = (1+i)$ and ε is defined as

$$\varepsilon = R|M|^{-2} = U_\infty / (\omega a) \ll 1.$$

Since the particle phase is a liquid, full tangential mobility at the liquid-gas interface is allowed. This is implemented with tangential velocity and shear-stress continuity conditions. Parameters with $(\hat{\quad})$ represent properties of liquid inside the drop, and parameters without $(\hat{\quad})$ correspond to the

gas outside the drop. The differential equation for the stream function in the drop phase, $\hat{\psi}(r, \theta)$, is

$$|\hat{M}|^2 \frac{\partial(D^2\hat{\psi})}{\partial\tau} + \hat{R} \left\{ \frac{1}{r^2} \frac{\partial(\hat{\psi}, D^2\hat{\psi})}{\partial(r, \bar{\mu})} + \frac{2}{r^2} D^2\hat{\psi} L\hat{\psi} \right\} = D^4\hat{\psi}, \quad (2)$$

where the following relationships hold:

$$\frac{R}{\hat{R}} = \frac{\hat{\nu}}{\nu}, \quad \frac{|M|^2}{R} = \frac{|\hat{M}|^2}{\hat{R}}, \quad \text{and} \quad \frac{|M|^2}{|\hat{M}|^2} = \frac{\hat{\nu}}{\nu}.$$

Long wavelength is assumed such that $ka \ll 1$ in which k is the wave number and a is the radius of the sphere. Under this approximation, the far-field flow around the drop may be approximated as a Taylor series about $z=0$ in the form

$$u_z = U_\infty \cos kze^{i\omega t} \approx U_\infty e^{i\omega t} [1 - \frac{1}{2}(kz)^2 + \dots]. \quad (3)$$

In the present investigation, only the leading term is used and for sufficiently small values of ka , the region around the drop can be regarded as having a spatially uniform velocity, $u_z \approx U_\infty e^{i\omega t}$.

The interface conditions are as follows:

(i) Velocity continuity:

$$\hat{\psi}|_{r=1} = \psi|_{r=1} = 0, \quad (4)$$

$$\frac{\partial\hat{\psi}}{\partial r} \Big|_{r=1} = \frac{\partial\psi}{\partial r} \Big|_{r=1}. \quad (5)$$

(ii) Shear stress continuity:

$$[\tau_{r\theta} - \hat{\tau}_{r\theta}]_{r=1} = - \left[\mu \frac{r}{\sin\theta} \frac{\partial}{\partial r} \left(\frac{1}{r^2} \frac{\partial\psi}{\partial r} \right) - \hat{\mu} \frac{r}{\sin\theta} \frac{\partial}{\partial r} \left(\frac{1}{r^2} \frac{\partial\hat{\psi}}{\partial r} \right) \right]_{r=1} = 0, \quad (6)$$

where μ is the dynamic viscosity for gas and $\hat{\mu}$ is that for liquid.

Here, it is assumed that surface oscillations are negligible and therefore there is zero normal velocity at the interface.

In addition to these interface conditions, a finite velocity is required within the drop, and, therefore, we have

$$\frac{1}{r^2} \hat{\psi} < \infty \quad \text{as} \quad r \rightarrow 0. \quad (7)$$

For the small parameter ε , the perturbation method is applicable to this problem and the following expansions are used:

$$\begin{aligned} \psi &= \psi_0 + \varepsilon\psi_1 + O(\varepsilon^2), \\ \Psi &= \Psi_0 + \varepsilon\Psi_1 + O(\varepsilon^2), \end{aligned} \quad (8)$$

$$\hat{\psi} = \hat{\psi}_0 + \varepsilon\hat{\psi}_1 + O(\varepsilon^2),$$

where Ψ represents the stream function in the inner region of the exterior phase. Detailed solutions are presented for the special case when the streaming Reynolds number, $R_s = \varepsilon R$, is also small, i.e.,

$$R_s = \varepsilon R \ll 1. \quad (9)$$

II. THE LEADING-ORDER SOLUTIONS

Here we are considering the case where $|M|^2 \gg 1$ and the liquid sphere is placed at the velocity antinode of the wave. As in Riley's¹⁹ problem, the stream function of outer flow ψ_0 should satisfy the inviscid flow equation,

$$\frac{\partial(D^2\psi_0)}{\partial\tau} = 0,$$

from which we find that

$$\psi_0 = (Ar^2 + Br^{-1})(1 - \bar{\mu}^2)e^{i\tau}, \quad (10)$$

where A, B are constants to be determined. The far field corresponding to the first term in Eq. (17) may be written in the stream-function form, $\psi_0 \rightarrow \frac{1}{2}r^2(1 - \bar{\mu}^2)e^{i\tau}$ [see also Riley,¹⁹ Eq. (3)]. With this condition, we obtain

$$\psi_0 = (\frac{1}{2}r^2 + Br^{-1})(1 - \bar{\mu}^2)e^{i\tau}. \quad (11)$$

Near the surface, in the shear-wave layer, the streaming field is described by

$$\frac{\partial^3\Psi_0}{\partial\tau\partial\eta^2} = \frac{1}{2}\left(\frac{\partial^4\Psi_0}{\partial\eta^4}\right), \quad (12)$$

where the inner variable η is defined as

$$\eta = \frac{1}{2}(r-1)|M|\sqrt{2}, \quad (13)$$

and Ψ_0 is related to ψ_0 in the matching region as

$$\Psi_0 \sim \frac{1}{2}\sqrt{2}|M|\psi_0. \quad (14)$$

The two sides are equal as $\eta \rightarrow 0$. Upon matching with the outer stream function (11), we may deduce that Ψ_0 is of the form

$$\Psi_0 = g(\eta)(1 - \bar{\mu}^2)e^{i\tau}. \quad (15)$$

By inserting Eq. (15) into Eq. (12), we obtain

$$\Psi_0 = (C + D\eta + Fe^{-(1+i)\eta})(1 - \bar{\mu}^2)e^{i\tau}, \quad (16)$$

where C, D , and F are constants to be determined.

We now expand Eq. (11) in Taylor series about $r=1$ and replace $(r-1)$ with $\sqrt{2}\eta/|M|$, and obtain

$$\psi_0 \sim \left\{ \left(\frac{1}{2} + B \right) + (1-B)\frac{\sqrt{2}}{|M|}\eta \right\} (1 - \bar{\mu}^2)e^{i\tau}. \quad (17)$$

Matching this to $O(\varepsilon^0)$ with Eq. (16) yields

$$\frac{1}{2} + B = C\frac{\sqrt{2}}{|M|}, \quad \text{and} \quad 1 - B = D.$$

Assuming that $C = O(1)$, and noting that $|M| \gg 1$, we can conclude that

$$\frac{1}{2} + B = 0$$

to the leading order. Thus, we obtain

$$B = -\frac{1}{2} \quad \text{and} \quad D = \frac{3}{2}.$$

Thus,

$$\psi_0 = \frac{1}{2}\left(r^2 - \frac{1}{r}\right)(1 - \bar{\mu}^2)e^{i\tau} \quad (18)$$

and

$$\Psi_0 = (C + \frac{3}{2}\eta + Fe^{-(1+i)\eta})(1 - \bar{\mu}^2)e^{i\tau}. \quad (19)$$

With $\varepsilon \ll 1$ or, equivalently, $|M|^2 \gg R$, we may deduce that

$$|\hat{M}|^2 \gg \hat{R}.$$

Thus the leading-order stream function of the liquid inside the sphere should satisfy

$$|\hat{M}|^2 \frac{\partial(D^2\hat{\psi}_0)}{\partial\tau} = D^4\hat{\psi}_0. \quad (20)$$

By comparison with the flow field in the exterior gas phase, we can surmise that the liquid flow inside the sphere is of the form,

$$\hat{\psi}_0 = \hat{\phi}_0(1 - \bar{\mu}^2)e^{i\tau}. \quad (21)$$

Upon inserting Eq. (21) into (20), and using the relationship

$$|\hat{M}|^2 = \left[\frac{\hat{M}\sqrt{2}}{(1+i)} \right]^2 = \frac{\hat{M}^2}{i}, \quad (22)$$

we obtain

$$\hat{M}^2 D^2[\hat{\phi}_0(1 - \bar{\mu}^2)] = D^4[\hat{\phi}_0(1 - \bar{\mu}^2)]. \quad (23)$$

This equation may be solved by letting

$$D^2[\hat{\phi}_0(1 - \bar{\mu}^2)] = G(r, \bar{\mu}),$$

so that

$$D^2G(r, \bar{\mu}) = \hat{M}^2G(r, \bar{\mu}). \quad (24)$$

Since $(1 - \bar{\mu}^2)$ is an eigenfunction of D^2 , $G(r, \bar{\mu})$ should be of the form

$$G(r, \bar{\mu}) = G_0(r)(1 - \bar{\mu}^2),$$

and the function $G_0(r)$ satisfies

$$\frac{d^2G_0}{dr^2} - \frac{2}{r^2}G_0 = \hat{M}^2G_0. \quad (25)$$

By substituting $x = \hat{M}r$, we have

$$x^2 \frac{d^2G_0}{dx^2} - 2G_0 = x^2G_0. \quad (26)$$

The solution is in the form of spherical Bessel functions given by

$$G_0 = \left(\frac{e^x}{x} - e^x \right) + B_1 \left(\frac{e^{-x}}{x} + e^{-x} \right). \quad (27)$$

With

$$D^2[\hat{\phi}_0(1 - \bar{\mu}^2)] = G(r, \bar{\mu}) = G_0(1 - \bar{\mu}^2), \quad (28)$$

it is not difficult to see that $\hat{\phi}_0$ satisfies

$$x^2 \frac{d^2 \hat{\phi}_0}{dx^2} - 2 \hat{\phi}_0 = x^2 C_1 \left(\frac{e^x}{x} - e^x \right) + x^2 D_1 \left(\frac{e^{-x}}{x} + e^{-x} \right), \quad (29)$$

which, upon integration, yields

$$\hat{\phi}_0 = ax^2 + bx^{-1} + c \left(\frac{1}{x} - 1 \right) e^x + d \left(\frac{1}{x} + 1 \right) e^{-x}, \quad (30)$$

or, equivalently,

$$\begin{aligned} \hat{\phi}_0 = & A^* r^2 + B^* r^{-1} + C^* \left(\frac{1}{\hat{M}r} - 1 \right) e^{\hat{M}r} \\ & + D^* \left(\frac{1}{\hat{M}r} + 1 \right) e^{-\hat{M}r}, \end{aligned} \quad (31)$$

where A^* , B^* , C^* , and D^* are constants to be determined. In order to have finite velocity at the origin, we need to satisfy

$$\frac{1}{r^2} \hat{\psi} < \infty \quad \text{as } r \rightarrow 0.$$

This requires

$$B^* = 0 \quad \text{and} \quad C^* = -D^*.$$

The solution for $\hat{\psi}_0$ takes the form

$$\begin{aligned} \hat{\psi}_0 = & \left[A^* r^2 + C^* \left(\frac{1}{\hat{M}r} - 1 \right) e^{\hat{M}r} - C^* \left(\frac{1}{\hat{M}r} + 1 \right) e^{-\hat{M}r} \right] \\ & \times (1 - \bar{\mu}^2) e^{i\tau}. \end{aligned} \quad (32)$$

The interface conditions (4)–(6), to the leading order, may be written as

(i) velocity continuity:

$$\hat{\psi}_0|_{r=1} = \Psi_0|_{\eta=0} = 0, \quad (33)$$

$$\left. \frac{\partial \hat{\psi}_0}{\partial r} \right|_{r=1} = \left. \frac{\partial \Psi_0}{\partial \eta} \right|_{\eta=0}; \quad (34)$$

(ii) shear stress continuity:

$$\begin{aligned} [\tau_{r\theta} - \hat{\tau}_{r\theta}]_{r=1} = & - \left[\mu \frac{r}{\sin \theta} \frac{\partial}{\partial r} \left(\frac{1}{r^2} \frac{\partial \Psi_0}{\partial \eta} \right) \right. \\ & \left. - \hat{\mu} \frac{r}{\sin \theta} \frac{\partial}{\partial r} \left(\frac{1}{r^2} \frac{\partial \hat{\psi}_0}{\partial r} \right) \right]_{r=1} = 0. \end{aligned} \quad (35)$$

To satisfy the boundary conditions (33)–(35), we have

$$F = -C,$$

$$A^* = C^* \left[\left(\frac{1}{\hat{M}} + 1 \right) e^{-\hat{M}} - \left(\frac{1}{\hat{M}} - 1 \right) e^{\hat{M}} \right],$$

$$C^* = \frac{\frac{3}{2} + C(1+i)}{e^{-\hat{M}}[(3/\hat{M}) + 3 + \hat{M}] + e^{\hat{M}}[-(3/\hat{M}) + 3 - \hat{M}]},$$

$$C = \frac{U}{V},$$

where

$$\begin{aligned} U = & -3\kappa \left[e^{-\hat{M}} \left(\frac{3}{\hat{M}} + 3 + \hat{M} \right) + e^{\hat{M}} \left(-\frac{3}{\hat{M}} + 3 - \hat{M} \right) \right] \\ & - \frac{3}{2} \left[e^{-\hat{M}} \left(-\frac{6}{\hat{M}} - 6 - 3\hat{M} - \hat{M}^2 \right) \right. \\ & \left. + e^{\hat{M}} \left(\frac{6}{\hat{M}} - 6 + 3\hat{M} - \hat{M}^2 \right) \right], \\ V = & (1+i) \left[e^{-\hat{M}} \left(-\frac{6}{\hat{M}} - 6 - 3\hat{M} - \hat{M}^2 \right) \right. \\ & \left. + e^{\hat{M}} \left(\frac{6}{\hat{M}} - 6 + 3\hat{M} - \hat{M}^2 \right) \right] + \kappa(1+i)(2+M) \\ & \times \left[e^{-\hat{M}} \left(\frac{3}{\hat{M}} + 3 + \hat{M} \right) + e^{\hat{M}} \left(-\frac{3}{\hat{M}} + 3 - \hat{M} \right) \right], \end{aligned}$$

and $\kappa = \mu/\hat{\mu}$.

The solutions for the liquid phase flow and the inner and outer flow fields have the following expressions for the corresponding stream functions:

$$\begin{aligned} \hat{\psi}_0 = & C^* \left[r^2 \left(\frac{1}{\hat{M}} + 1 \right) e^{-\hat{M}} - r^2 \left(\frac{1}{\hat{M}} - 1 \right) e^{\hat{M}} + \left(\frac{1}{\hat{M}r} - 1 \right) e^{\hat{M}r} \right. \\ & \left. - \left(\frac{1}{\hat{M}r} + 1 \right) e^{-\hat{M}r} \right] (1 - \bar{\mu}^2) e^{i\tau}, \end{aligned} \quad (36)$$

$$\Psi_0 = [C + \frac{3}{2}\eta - Ce^{-(1+i)\eta}] (1 - \bar{\mu}^2) e^{i\tau}, \quad (37)$$

$$\psi_0 = \frac{1}{2} \left(r^2 - \frac{1}{r} \right) (1 - \bar{\mu}^2) e^{i\tau}. \quad (38)$$

The result for ψ_0 is the same as that of Riley's¹⁹ solution for solid sphere.

Usually μ is much smaller than $\hat{\mu}$, and for such cases $\kappa \ll 1$. With the assumption that $|\hat{M}| \gg 1$, we can simplify C and C^* such that

$$C \approx -\frac{3}{4}(1-i) \quad (39)$$

and

$$C^* \approx 0. \quad (40)$$

It should be noted here that Eq. (39) is consistent with the earlier assumption that $C = O(1)$. With these values of C and C^* , $\hat{\psi}_0$ and Ψ_0 can be approximated as

$$\hat{\psi}_0 \approx 0 \quad (41)$$

and

$$\Psi_0 \approx \frac{3}{2} [\eta - \frac{1}{2}(1-i)(1 - e^{-(1+i)\eta})] (1 - \bar{\mu}^2) e^{i\tau}, \quad (42)$$

respectively. Thus, for the case when $\kappa \ll 1$ and $|M| \gg 1$, the leading-order stream functions in the case of a liquid sphere at the velocity antinode are the same as those of solid sphere.

III. THE FIRST-ORDER SOLUTIONS [$O(\varepsilon)$]

To obtain the steady streaming flow, only the time-independent parts of the stream functions are required. The first-order solutions are much more complex than the leading order. By using the set of perturbation expansions (8) to $O(\varepsilon)$, we obtain the following linear equation for $\Psi_1(\eta, \bar{\mu}, \tau)$:

$$\frac{\partial^3 \Psi_1}{\partial \tau \partial \eta^2} + \frac{\partial \Psi_0}{\partial \eta} \frac{\partial^3 \Psi_0}{\partial \bar{\mu} \partial \eta^2} - \frac{\partial \Psi_0}{\partial \bar{\mu}} \frac{\partial^3 \Psi_0}{\partial \eta^3} + \frac{2\bar{\mu}}{(1-\bar{\mu}^2)} \frac{\partial^2 \Psi_0}{\partial \eta^2} \frac{\partial \Psi_0}{\partial \eta} = \frac{1}{2} \left(\frac{\partial^4 \Psi_1}{\partial \eta^4} \right). \quad (43)$$

In this case, we can express the solution for Ψ_1 as

$$\Psi_1(\eta, \bar{\mu}, \tau) = \frac{9}{2} \{ \zeta_{20}(\eta) + \zeta_{22}(\eta) e^{2i\tau} \} \bar{\mu}(1-\bar{\mu}^2). \quad (44)$$

By inserting Eqs. (42) and (44) into Eq. (43), we find

$$\frac{d^4 \zeta_{20}}{d\eta^4} = e^{-2\eta} - 2\eta e^{-\eta} \sin \eta - e^{-\eta} \cos \eta + e^{-\eta} \sin \eta. \quad (45)$$

The solution for Eq. (45) is

$$\zeta_{20} = C_1 \eta + C_2 + \frac{1}{16} e^{-2\eta} + \frac{5}{4} e^{-\eta} \cos \eta + \frac{3}{4} e^{-\eta} \sin \eta + \frac{1}{2} \eta e^{-\eta} \sin \eta \quad (46)$$

in which C_1 and C_2 are constants to be determined.

To obtain the liquid flow field inside the sphere, we also need to separate the time-independent part from the time-dependent part. Therefore, we write $\hat{\psi}_1$ as

$$\hat{\psi}_1(r, \bar{\mu}, \tau) = \{ H_1(r) + I_1(r, \tau) \} \bar{\mu}(1-\bar{\mu}^2). \quad (47)$$

A careful examination of the perturbation procedure shows that the steady-state motion inside the sphere is just Stokes flow. Thus, $H_1(r)$ satisfies

$$D^4 [H_1 \bar{\mu}(1-\bar{\mu}^2)] = 0. \quad (48)$$

The solution for $H_1(r)$ is

$$H_1(r) = A_1 r^3 + A_2 r^{-2} + A_3 r^5 + A_4, \quad (49)$$

where A_1, A_2, A_3 and A_4 are yet unknown constants. With the limit

$$\frac{1}{r^2} \hat{\psi} < \infty \quad \text{as } r \rightarrow 0,$$

it is easy to show that

$$A_2 = A_4 = 0.$$

To solve for the other constants, the procedure is quite similar to that of the leading-order solutions. The final results for the time-independent parts of Ψ_1 and $\hat{\psi}_1$ are

$$\Psi_{1(\tau\text{-indep})} = \frac{9}{2} \left[\frac{5}{8} \eta - \frac{|M|\kappa\eta}{\sqrt{2}(20+8\kappa)} - \frac{21}{16} + \frac{1}{16} e^{-2\eta} + \frac{5}{4} e^{-\eta} \cos \eta + \frac{3}{4} e^{-\eta} \sin \eta + \frac{1}{2} \eta e^{-\eta} \sin \eta \right] \bar{\mu}(1-\bar{\mu}^2), \quad (50)$$

and

$$\hat{\psi}_{1(\tau\text{-indep})} = \frac{9|M|\kappa}{\sqrt{2}(80+32\kappa)} (r^3 - r^5) \bar{\mu}(1-\bar{\mu}^2). \quad (51)$$

Similarly, we can write the first-order outer flow stream function as

$$\psi_1 = F_1(r, \mu) + G_1(r, \mu, \tau). \quad (52)$$

The equation for F_1 , obtained from the perturbed form of Eq. (1), to $O(\varepsilon^2)$ is

$$\frac{R_s}{r^2} \left\{ \frac{\partial F_1}{\partial r} \frac{\partial}{\partial \mu} (D^2 F_1) - \frac{\partial F_1}{\partial \mu} \frac{\partial}{\partial r} (D^2 F_1) + 2L F_1 D^2 F_1 \right\} = D^4 F_1. \quad (53)$$

According to our assumption in Eq. (9), $R_s \ll 1$, and we can simplify Eq. (53) as

$$D^4 F_1 = 0. \quad (54)$$

The appropriate solution, taking account of the matching requirement and the condition $F_1 = o(r^2)$ as $r \rightarrow \infty$, is

$$F_1 = \left(\frac{B_1}{r^2} + B_2 \right) \bar{\mu}(1-\bar{\mu}^2). \quad (55)$$

If we let $r \rightarrow 1$ in (55) and introduce the inner variable η defined in Eq. (13), we have

$$F_1 \sim \left\{ B_1 + B_2 - \frac{2\sqrt{2}}{|M|} B_1 \eta \right\} \bar{\mu}(1-\bar{\mu}^2). \quad (56)$$

By matching to this order with (50),

$$B_1 + B_2 = 0 \quad \text{and} \quad -2B_1 = \frac{9}{2} \left(\frac{5}{8} - \frac{|M|\kappa}{\sqrt{2}(20+8\kappa)} \right).$$

The final result for the time-independent part of ψ_1 is

$$\psi_{1(\tau\text{-indep})} = \left(-\frac{45}{32} + \frac{9|M|\kappa}{\sqrt{2}(80+32\kappa)} \right) \left(\frac{1}{r^2} - 1 \right) \bar{\mu}(1-\bar{\mu}^2). \quad (57)$$

IV. DISCUSSION

From Eqs. (50) and (57), we can see that both inner and outer flow fields for the exterior depend on the frequency parameter M . Based on the value of M in relation to the fluid properties, the flow field is predicted to have two distinct classifications. This is seen after examining the parameter

$$B_1 = -\frac{45}{32} + \frac{9|M|\kappa}{\sqrt{2}(80+32\kappa)}, \quad (58)$$

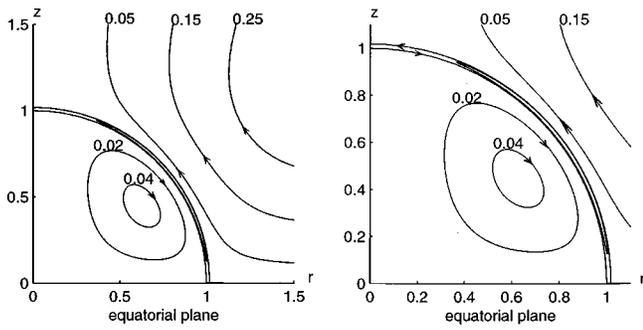


FIG. 1. Streaming about a drop at the velocity antinode when $B_1 < 0$, for $|M| = 200$ and $\kappa = 0.0156$. On the right, the flow near the interface is shown in detail.

which appears as a factor in the steady part of the outer-region flow field [Eq. (57)]. For negative values of B_1 , the steady-state flow pattern outside the liquid sphere is quite similar to that of the solid sphere obtained by Riley.¹⁹ The shear layer in this case has recirculation (see Fig. 1). The recirculation is shown with a much-amplified radial scale in Fig. 2. The liquid motion inside the sphere consists of two toroidal vortices. A quadrant showing the flow pattern in one of these vortices can be seen in Fig. 1.

With increasing $|M|$, the parameter B_1 decreases in magnitude until it vanishes, and then reverses sign. At that point, the recirculation in the shear layer ceases and the streamlines inside the layer merge with the outer ones. The liquid motion inside the sphere is stronger but the streamlines are similar to what is shown in Fig. 1 (see also Fig. 3). This theory that predicts the cessation of recirculation in the shear-wave layer is presently under thorough examination and experimental verification is necessary. For many of our levitation experiments, the wave frequency is 20–40 kHz (or angular frequency $\omega = 125 - 250 \times 10^3$ rad/s), and the diameter of the sphere is 3–8 mm. For a gaseous medium outside the liquid drop, κ is $O(10^{-2})$. In this case, $|M|$ ranges from 140 to 600, and we find that $B_1 < 0$, i.e.,

$$\frac{9|M|\kappa}{\sqrt{2}(80+32\kappa)} < \frac{45}{32} \text{ or } |M| < \frac{5\sqrt{2}(5+2\kappa)}{2\kappa},$$

whereupon recirculation is expected. For a given set of materials, the transition is predicted to take place at a drop radius,

$$a_{\text{trans}} = \frac{5(5+2\kappa)}{\kappa} \left(\frac{\nu}{2\omega}\right)^{1/2} \approx \frac{25}{\kappa} \left(\frac{\nu}{2\omega}\right)^{1/2}.$$

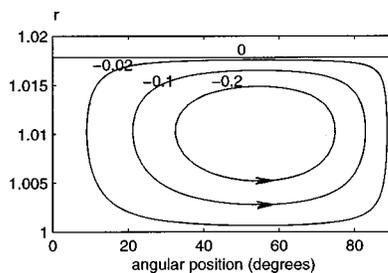


FIG. 2. Detailed recirculating flow in the Stokes layer shown with a stretched radial scale for $|M| = 200$ and $\kappa = 0.0156$.

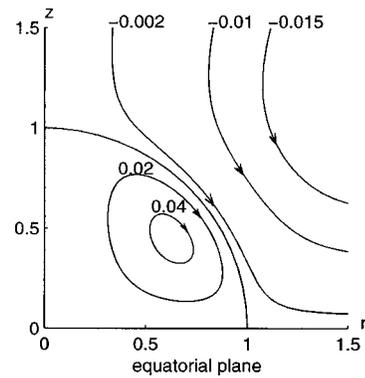


FIG. 3. Streaming about a drop at the velocity antinode for $B_1 > 0$ with $|M| = 1200$ and $\kappa = 0.0156$.

With a 20-kHz levitator, for a water drop in air at 15 °C, recirculation is predicted to not take place at radii above $a = 12$ mm, and with 40 kHz, at radii above 8.7 mm. These are large drops, and experimentation at these frequencies, without significant shape deformation, may be feasible only under microgravity conditions. Even then, the surface oscillations may play a role. Furthermore, the validity of the various approximations ($R \ll 1$ and $a \ll \lambda$) would also break down. However, possibilities of transition at higher temperature and pressure do exist for smaller-sized drops.

While detailed explanations for the cessation of recirculation in the Stokes layer are being sought, we recall that the vorticity generated by the acoustic field interacting with an interface is manifested in the form of recirculation (see Ref. 20). This takes place when an acoustic wave interacts with a solid surface. In the case of a fluid surface, interfacial mobility is likely to reduce this effect. With decreasing drop-phase viscosity, the strength of the shear-wave layer recirculation diminishes and could vanish when the parameter B_1 in Eq. (57) is zero. It should be noted that at that point, ψ_1 also vanishes, and higher-order solutions are needed for a valid description. It should also be noted that at the transition point, besides the cessation of recirculation, the outer-region streaming also reverses sign. However, this transition needs to be studied in greater detail because in the region where $B_1 \approx 0$, as it has just been stated, higher-order terms would be dominant.

ACKNOWLEDGMENTS

The authors are very grateful to the U.S. National Aeronautics and Space Administration for the support of this work under Grant No. NAG3-1842.

¹H. Zhao, S. S. Sadhal, and E. H. Trinh, "Singular perturbation analysis of an acoustically levitated sphere: Flow about the velocity node," *J. Acoust. Soc. Am.* **106**, 589–595 (1999).

²F. Shen and J. Khodadadi, "Containerless thermal diffusivity measurement of levitated spherical specimen by an extended flash method," in *Heat Transfer in Microgravity Systems—1993, Proc. 29th National Heat Transfer Conference*, edited by S. S. Sadhal and A. Hashemi, Atlanta, August 1993 (American Society of Mechanical Engineers, New York, 1993), Vol. HTD-235, pp. 33–42.

³S. S. Sadhal, E. H. Trinh, and P. Wagner, "Unsteady spot heating of a drop in a microgravity environment," *Microgravity Sci. Technol.* **9**, 80–85 (1997).

- ⁴L. V. King, "On the acoustic radiation pressure on spheres," *Proc. R. Soc. London, Ser. A* **147**, 212–240 (1934).
- ⁵L. V. King, "On the acoustic radiation pressure on circular discs: Inertia and diffraction corrections," *Proc. R. Soc. London, Ser. A* **153**, 1–16 (1935).
- ⁶T. J. Asaki and P. L. Marston, "Acoustic radiation force on a bubble driven above resonance," *J. Acoust. Soc. Am.* **96**, 3096–3099 (1994).
- ⁷L. A. Crum, "Acoustic force on a liquid droplet in an acoustic stationary wave," *J. Acoust. Soc. Am.* **50**, 157–163 (1971).
- ⁸A. I. Eller, "Force on a bubble in a standing acoustic wave," *J. Acoust. Soc. Am.* **43**, 170–171 (1968).
- ⁹C. P. Lee and T. G. Wang, "Acoustic radiation force on a bubble," *J. Acoust. Soc. Am.* **93**, 1637–1640 (1993).
- ¹⁰K. Yosioka and Y. Kawasima, "Acoustic radiation pressure on a compressible sphere," *Acustica* **5**, 167–173 (1955).
- ¹¹F. H. Busse and T. G. Wang, "Torque generated by orthogonal acoustic waves—Theory," *J. Acoust. Soc. Am.* **69**, 1634–1638 (1981).
- ¹²A. Gopinath and A. F. Mills, "Convective heat transfer from a sphere due to acoustic streaming," *J. Heat Transfer* **115**, 332–341 (1993).
- ¹³C. P. Lee and T. G. Wang, "Acoustic radiation force on a heated sphere including the effects of heat transfer and acoustic streaming," *J. Acoust. Soc. Am.* **83**, 1324–1331 (1988).
- ¹⁴E. H. Trinh, "Compact acoustic levitation device for studies in fluid dynamics and material science in the laboratory and microgravity," *Rev. Sci. Instrum.* **56**, 2059–2065 (1985).
- ¹⁵E. H. Trinh, "Levitation studies of the physical properties and nucleation of undercooled liquids," in *Proc. VIIIth Euro. Symp. on Mat. and Fluid Sci. in Microgravity*, Oxford, UK, September, 1989, January 1990, ESA Publication SP-295.
- ¹⁶C. P. Lee and T. G. Wang, "Outer acoustic streaming," *J. Acoust. Soc. Am.* **88**, 2367–2375 (1990).
- ¹⁷A. Gopinath, "Steady streaming due to a small-amplitude torsional oscillations of a sphere in a viscous fluid," *Q. J. Mech. Appl. Math.* **46**, 501–520 (1993).
- ¹⁸A. Gopinath, "Steady streaming due to small-amplitude superposed oscillations of a sphere in a viscous fluid," *Q. J. Mech. Appl. Math.* **47**, 461–480 (1994).
- ¹⁹N. Riley, "On a sphere oscillating in a viscous liquid," *Q. J. Mech. Appl. Math.* **19**, 461–472 (1966).
- ²⁰H. Schlichting, "Berechnung ebener periodischer Grenzschichtströmungen," *Phys. Z.* **23**, 327–335 (1932).
- ²¹S. Goldstein, *Modern Developments in Fluid Dynamics* (Clarendon, Oxford, 1938), Vol. I, pp. 114–115.

Shear interfacial waves in piezoelectrics

A. N. Darinskii and V. N. Lyubimov

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

(Received 12 May 1998; accepted for publication 30 August 1999)

An analysis was made of the SH-interfacial wave propagation on an interface between semi-infinite piezoelectric media. In particular, a sufficient condition for the existence of the SH-wave on the metallized rigid contact was derived. It has been shown that on the metallized rigid contact between two completely identical piezocrystals the SH-localized wave always exists, with the exception of at most two orientations of the interface. An investigation was performed of specific features of the sound reflection—transmission at incident angles corresponding to the resonant excitation of the leaky wave originating from the “supersonic” interfacial wave due to misorientation of crystals that constitute the bicrystal. Also, it has been established that in 6 mm crystals an interfacial “supersonic” wave may remain purely localized even when the direction of propagation is not perpendicular to the axis 6 or when the interface is not parallel to this axis. © 1999 Acoustical Society of America. [S0001-4966(99)05012-2]

PACS numbers: 43.35.Cg [HEB]

INTRODUCTION

A manifestation of the piezoelectric effect in acoustics is the possibility of existence of shear horizontally polarized (SH) localized waves on a flat interface between semi-infinite piezoelectric media. As is well known, such solutions do not arise in purely elastic solids, either isotropic and anisotropic; on the other hand, “nonpiezoelectric” SH-localized waves can travel in layered composites, on corrugated surfaces, and under arrays of periodically deposited strips.^{1,2}

The analytical investigation of the SH-interfacial wave problem in a piezoelectric bicrystal generally is not possible. However, a number of particular cases allow a rather simple explicit treatment. For example, in Ref. 3 the SH-wave propagation on the “180°-domain wall” in 6 mm and 4 mm piezoelectrics has been investigated. A detailed analysis of localized SH-waves on the “180°-domain wall” in crystals of various symmetry groups has been performed in Ref. 4, and in Ref. 5 domain wall waves have been observed experimentally. In Ref. 6 the authors have considered localized SH-waves propagating along the rigid contact between two different piezocrystals of symmetry 6 mm.

The first objective of the present paper is to consider the existence problem for SH-interfacial waves in crystals of symmetry 2 without resorting to explicit computations. We are concerned with the SH-wave propagation on two types of interface: (1) The perfect rigid contact; (2) The metallized perfect rigid contact. The latter simulates the case when an intermediate metallic layer of infinitesimal thickness is sandwiched between crystals.

The second objective of the paper is to study peculiarities of the coefficients of mode conversion caused by the resonant excitation of leaky waves. These leaky waves originate from the SH-interfacial wave due to misorientation of the bicrystal.

The paper is organized as follows. We begin with a brief review of basic concepts of the multidimensional formalism for piezoelectrics.^{7,8} Afterward we shall discuss the SH-

interfacial wave problem and the resonant reflection near a leaky wave branch. In the last section, a specific feature of the localized wave propagation in crystals of symmetry 6 mm is considered.

I. BASIC RELATIONS

We specify the coordinate system as shown in Fig. 1 and introduce the following notation for parameters of a plane mode α . \mathbf{A}_α is the polarization vector associated with the mechanical displacement $\mathbf{u}_\alpha(\mathbf{r}, t)$, Φ_α is the amplitude of the electrical potential $\varphi_\alpha(\mathbf{r}, t)$, \mathbf{L}_α and Q_α signify the amplitudes of the projections of the mechanical stress tensor $\hat{\boldsymbol{\sigma}}_\alpha(\mathbf{r}, t)$ and of the induction $\mathbf{D}_\alpha(\mathbf{r}, t)$ onto the unit normal \mathbf{n} to the plane boundary of the medium, respectively. $\mathbf{L}_\alpha = ik^{-1}\hat{\boldsymbol{\sigma}}_\alpha\mathbf{n}$ and $Q_\alpha = ik^{-1}\mathbf{D}_\alpha\cdot\mathbf{n}$, where k is the projection of the wave vector onto the unit vector \mathbf{m} pointing in the direction of propagation along the surface.

As usual, c_{ijkl} , e_{ijk} , and ϵ_{ij} stand for the elastic and piezoelectric moduli and the permittivity, respectively, ρ is the density. The symbol (ab) , where \mathbf{a} and \mathbf{b} is a pair of three-component vectors, denotes a 4×4 matrix with elements $(ab)_{IJ} = a_k E_{kIJI} b_I$, $I, J = 1, \dots, 4$, where $E_{kIJI} = c_{kIJI} - \rho v^2 m_k m_I \delta_{IJ}$, $I, J = 1, 2, 3$, $E_{k4JI} = e_{kJI}$, $J = 1, 2, 3$, $E_{kI4I} = e_{Ikk}$, $I = 1, 2, 3$, $E_{k44I} = -\epsilon_{kl}$, and $v = \omega/k$ is the trace speed, ω is the frequency.

Substitution of $(\mathbf{u}_\alpha, \varphi_\alpha) \propto \exp ik[(\mathbf{m} + p_\alpha \mathbf{n}) \cdot \mathbf{r} - vt]$ into the equations of motion and the appropriate constitutive relations, see, e.g., Ref. 9, yields

$$\{p_\alpha^2(nn) + p_\alpha[(nm) + (mn)] + (mm)\} \mathbf{U}_\alpha = \mathbf{0}, \quad (1)$$

$$-[(nm) + p_\alpha(nn)] \mathbf{U}_\alpha = \mathbf{V}_\alpha, \quad (2)$$

where $\mathbf{U}_\alpha = (\mathbf{A}_\alpha, \Phi_\alpha)^t$ and $\mathbf{V}_\alpha = (\mathbf{L}_\alpha, Q_\alpha)^t$ are four-component column vectors; the symbol $(\dots)^t$ signifies the transposition. Equations (1) and (2) can be recast into the form^{7,8}

$$\hat{\mathbf{N}} \boldsymbol{\xi}_\alpha = p_\alpha \boldsymbol{\xi}_\alpha, \quad (3)$$

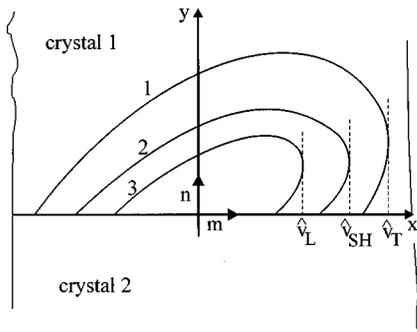


FIG. 1. Geometry of the problem and cross-sections of the slowness surfaces of bulk waves in the upper crystal (one half). (1) quasi transverse branch; (2) SH-branch; (3) quasi longitudinal branch. \hat{v}_L , \hat{v}_{SH} , and \hat{v}_T are the limiting speeds of quasi transverse, shear, and quasi longitudinal waves, respectively.

where \hat{N} is an 8×8 real matrix,

$$\hat{N} = - \begin{Bmatrix} (nn)^{-1}(nm) & (nn)^{-1} \\ (mn)(nn)^{-1}(nm) - (mm) & (mn)(nn)^{-1} \end{Bmatrix}, \quad (4)$$

$(nn)^{-1}$ is the inverse of the matrix (nn) , ξ_α is an eight-component column vector,

$$\xi_\alpha = (\mathbf{U}_\alpha, \mathbf{V}_\alpha)^t = (\mathbf{A}_\alpha, \Phi_\alpha, \mathbf{L}_\alpha, \mathbf{Q}_\alpha)^t. \quad (5)$$

It will be assumed that vectors Eq. (5) obey the normalization condition^{7,8}

$$\xi_\alpha \cdot \hat{\mathbf{T}} \xi_\beta = \mathbf{U}_\alpha \cdot \mathbf{V}_\beta + \mathbf{V}_\alpha \cdot \mathbf{U}_\beta = \delta_{\alpha\beta}, \quad \hat{\mathbf{T}} = \begin{pmatrix} \hat{\mathbf{O}} & \hat{\mathbf{I}} \\ \hat{\mathbf{I}} & \hat{\mathbf{O}} \end{pmatrix}, \quad (6)$$

where $\hat{\mathbf{I}}$ and $\hat{\mathbf{O}}$ are the unit and zero 4×4 matrices, respectively. Then

$$\sum_{\alpha=1}^8 \xi_\alpha \otimes \hat{\mathbf{T}} \xi_\alpha = \begin{pmatrix} \hat{\mathbf{I}} & \hat{\mathbf{O}} \\ \hat{\mathbf{O}} & \hat{\mathbf{I}} \end{pmatrix}, \quad (7)$$

where the symbol \otimes stands for dyadic multiplication.

The parameters p_α occur either in pairs of complex conjugates or are pairwise real. We shall use the numbering where, for complex p_α 's,

$$p_\alpha = p_{\alpha+4}^*, \quad \text{Im}(p_\alpha) > 0, \quad \xi_\alpha = \xi_{\alpha+4}^*, \quad \alpha \leq 4 \quad (8)$$

(the asterisk means the complex conjugation) and, for real p_α 's,

$$\text{a pair of subscripts } \alpha, \alpha+4 \text{ is assigned to a pair} \\ \text{reflected } (\alpha)\text{-incident } (\alpha+4) \text{ waves, } \alpha \leq 4. \quad (9)$$

Note that all p_α 's are complex in the so-called subsonic speed range, i.e., for v smaller than the limiting speed \hat{v} . The latter is the trace speed of the slowest bulk wave the group velocity of which is parallel to the surface of the medium ($\hat{v} = \hat{v}_T$ in Fig. 1). In the theory of surface waves, this wave is known as the limiting wave.^{7,8}

Note also that in terms of the vector ξ_α , the projection E_α of the energy flux of the mode α onto the normal \mathbf{n} to the surface of the medium reads

$$E_\alpha = - \frac{\omega k}{4} \xi_\alpha \cdot \hat{\mathbf{T}} \xi_\alpha^*, \quad (10)$$

where it follows that the normalized vectors ξ_α associated with reflected and incident waves are purely imaginary and purely real, respectively,

$$\xi_\alpha = - \xi_\alpha^* \quad \text{and} \quad \xi_{\alpha+4} = \xi_\alpha^* \quad (11)$$

provided \mathbf{n} is the inward normal (or vice versa if \mathbf{n} is the outward normal).

II. LOCALIZED SH-WAVES

Let both media in contact be so oriented that in each of them an even-fold axis of symmetry is parallel to the coordinate axis z . Equations (1) and (2) [and, of course, Eq. (3)] split then into two independent subsystems. One of them pertains to in-plane polarized nonpiezoactive modes with $A_{\alpha z} = L_{\alpha z} = \Phi_\alpha = Q_\alpha = 0$. The second subsystem describes shear piezoactive waves having $A_{\alpha x, y} = L_{\alpha x, y} = 0$. It will be assumed that

the subscripts $\alpha = 1, 2, 5, 6$ and $\alpha = 3, 4, 7, 8$ are ascribed

$$\text{to in-plane and SH-modes, respectively.} \quad (12)$$

Our subsequent consideration of the existence problem for interfacial SH-waves will be based on the impedance method borrowed from the general theory of localized waves in piezoelectrics. This method was developed in Refs. 7 and 8 to investigate surface wave propagation in semi-infinite piezoelectric media of arbitrary anisotropy. In Refs. 10 and 11 it was applied to the localized wave problem on an interface between two bonded piezocrystals of arbitrary symmetry. According to Refs. 10 and 11 at most two localized waves may exist on both the nonmetallized and metallized rigid contact.

The general theory of subsonic localized waves involves two Hermitian 4×4 impedance matrices $\hat{\mathbf{Z}}_F$ and $\hat{\mathbf{Z}}_\Phi$. When the sagittal plane is oriented perpendicularly to an even-fold axis of symmetry, these matrices simplify to a block diagonal form. Their left-hand side upper 2×2 blocks are associated with nonpiezoactive modes $\alpha = 1, 2, 5, 6$. Of concern to us are right-hand side lower 2×2 blocks involved in SH-wave problems; we keep the notation $\hat{\mathbf{Z}}_F$ and $\hat{\mathbf{Z}}_\Phi$ for these blocks. The 2×2 impedances

$$\hat{\mathbf{Z}}_X = \begin{pmatrix} Z_{X,33} & Z_{X,34} \\ Z_{X,34}^* & Z_{X,44} \end{pmatrix}, \quad X = F, \Phi, \quad (13)$$

where $Z_{X,ij}$ are elements of the full 4×4 matrices, interrelate the two-component vectors $\mathbf{V}_{F,\alpha} = (L_{\alpha z}, \Phi_\alpha)^t$ and $\mathbf{U}_{F,\alpha} = (A_{\alpha z}, Q_\alpha)^t$, $\mathbf{V}_{\Phi,\alpha} = (L_{\alpha z}, Q_\alpha)^t$ and $\mathbf{U}_{\Phi,\alpha} = (A_{\alpha z}, \Phi_\alpha)^t$, $\alpha = 3, 4, 7, 8$, as

$$\mathbf{V}_{X,\alpha} = -i \hat{\mathbf{Z}}_X \mathbf{U}_{X,\alpha} \quad \text{if } \text{Im}(p_\alpha) > 0, \\ \mathbf{V}_{X,\alpha} = i (\hat{\mathbf{Z}}_X)^t \mathbf{U}_{X,\alpha} \quad \text{if } \text{Im}(p_\alpha) < 0. \quad (14)$$

By Eq. (8), the modes $\alpha = 3, 4$ in the top part, $(\mathbf{n} \cdot \mathbf{r}) > 0$, and the modes $\alpha = 7, 8$ in the bottom part of the bicrystal can be used for interfacial wave construction. The medium in which

the limiting speed of SH-bulk waves \hat{v}_{SH} is the largest one will be called the fast medium; let it be the upper medium. The lower part is then the slow medium. The superscripts (f) and (s) will label quantities pertaining to the fast and slow media, respectively. In what follows, we discuss the existence of interfacial waves in the speed range $v < \hat{v}_{SH}^{(s)}$.

As we shall see, the existence of SH-interfacial waves is connected with the existence of a Bleustein–Gulyaev wave^{12–14} in each of the media constituting the bicrystal on the mechanically free metallized surface parallel to the plane of interface (it is this Bleustein–Gulyaev wave to which we refer below). This connection is due to the fact that, on the one hand, equations for the interfacial wave speed involve the matrix $\hat{\mathbf{Z}}_F$ and the element $Z_{\Phi,33}$. On the other hand, the equation for the Bleustein–Gulyaev wave speed v_{BG} can be written in the form $\|\hat{\mathbf{Z}}_F\|=0$, where the symbol $\|\dots\|$ means the determinant, or $Z_{\Phi,33}=0$. (Note that $Z_{\Phi,33}=Z_{F,44}^{-1}\|\hat{\mathbf{Z}}_F\|$ and $Z_{F,44}$ is positive and finite in the range $0 \leq v \leq \hat{v}_{SH}$; see Refs. 7, 8, 10.)

A. The nonmetallized rigid contact

The boundary conditions require continuity of displacements and electrical potentials as well as of the normal projection of the mechanical stress tensor and of the electrical induction. In terms of the vectors $\mathbf{V}_{F,\alpha}$ and $\mathbf{U}_{F,\alpha}$, they can be written, accounting for Eq. (8), as

$$\sum_{\alpha=3,4} b_{\alpha} \mathbf{V}_{F,\alpha}^{(f)} = \sum_{\alpha=3,4} c_{\alpha+4} \mathbf{V}_{F,\alpha}^{(s)*}, \quad (15)$$

$$\sum_{\alpha=3,4} b_{\alpha} \mathbf{U}_{F,\alpha}^{(f)} = \sum_{\alpha=3,4} c_{\alpha+4} \mathbf{U}_{F,\alpha}^{(s)*},$$

where b_{α} and $c_{\alpha+4}$ are the amplitudes of modes $\alpha=3,4$ and $\alpha=7,8$, respectively. We can take advantage of Eq. (14) to obtain the equation for the interfacial wave speed v_I in the form

$$\|\hat{\mathbf{Z}}_F^{(f)} + (\hat{\mathbf{Z}}_F^{(s)})^t\| = 0. \quad (16)$$

It appears that analytical properties of the 2×2 impedance $\hat{\mathbf{Z}}_F$ coincide with those of the 4×4 impedance $\hat{\mathbf{Z}}_F$; see Refs. 7, 8, 10, 11. Therefore the analysis of solvability of Eq. (16) repeats the discussion performed in Ref. 10 in relation to the localized wave problem on the nonmetallized interface in piezoelectrics of unrestricted anisotropy [the equation for the interfacial wave speed generally has the form Eq. (16) with 4×4 impedances $\hat{\mathbf{Z}}_F$ replacing 2×2 impedances]. We confine ourselves to quoting the result.

- At most one SH-wave may exist in the interval $v < \hat{v}_{SH}^{(s)}$.
- If the slow medium does not admit the Bleustein–Gulyaev wave and this wave either is not possible in the fast medium or $v_{BG}^{(f)} > \hat{v}_{SH}^{(s)}$, then no SH-interfacial wave exists.

In the other cases the SH-wave may or may not appear. One can also add that if the interfacial wave appears, then its speed is larger than the smaller of the speeds $v_{BG}^{(s)}$, $v_{BG}^{(f)}$.

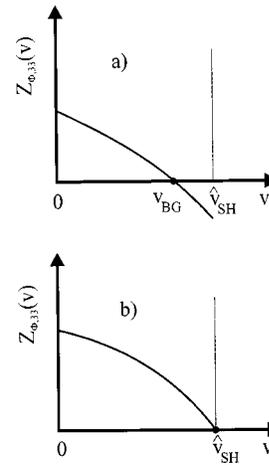


FIG. 2. Possible dependencies of $Z_{\Phi,33}$ on v . $Z_{\Phi,33}(v)$ vanishes at the Bleustein–Gulyaev wave speed v_{BG} (a). The Bleustein–Gulyaev wave does not exist in case (b).

B. The metallized rigid contact

In this instance, the electrical potential vanishes. Displacements and tractions are continuous across the interface. Using the vectors $\mathbf{U}_{\Phi,\alpha}$ and $\mathbf{V}_{\Phi,\alpha}$ in order to write the boundary conditions, we arrive at the following equation for v_I :

$$Z_{\Phi,33}^{(f)}(v) + Z_{\Phi,33}^{(s)}(v) = 0. \quad (17)$$

To analyze the solvability of Eq. (17) we must know how the element $Z_{\Phi,33}$ of the 2×2 matrix $\hat{\mathbf{Z}}_{\Phi}$ depends on v . Taking into account the properties of the 4×4 impedances $\hat{\mathbf{Z}}_F$ and $\hat{\mathbf{Z}}_{\Phi}$ discussed in detail in Refs. 7, 8, 10 as well as the existence theorems for surface waves in semi-infinite piezoelectrics,^{7,8} one can prove that the function $Z_{\Phi,33}(v)$ exhibits behavior as sketched in Fig. 2 depending on whether the limiting SH-wave does not satisfy (a) or satisfies (b) the boundary conditions of mechanically free metallized surface.

With regard for Eq. (17) and Fig. 2 we conclude that

- a sufficient condition for the existence of the SH-interfacial wave is $v_{BG}^{(f)} < \hat{v}_{SH}^{(s)}$.

When the Bleustein–Gulyaev wave appears only in the slow medium or this wave appears in both media, but $v_{BG}^{(f)} > \hat{v}_{SH}^{(s)}$, the interfacial wave may or may not exist. The SH-interfacial wave does not exist if

- neither medium admits the Bleustein–Gulyaev solution on the metallized free surface; the Bleustein–Gulyaev wave does not exist in the slow medium and exists in the fast one with $v_{BG}^{(f)} > \hat{v}_{SH}^{(s)}$.

So, we have analyzed the existence of SH-localized waves in the range $v < \hat{v}_{SH}^{(s)}$. When $v > \hat{v}_{SH}^{(s)}$, a pair of bulk SH-modes appears in the slow medium. The properties of the impedances $\hat{\mathbf{Z}}_{F,\Phi}^{(s)}$ then break down. For example, $\hat{\mathbf{Z}}_{F,\Phi}^{(s)}$ cease to be Hermitian matrices. The existence problem for SH-interfacial waves demands special consideration.

Now let us discuss the following case: The metallized interface between two identical piezoelectric media. Introducing the four-component vectors $\boldsymbol{\zeta}_{M\alpha} = (A_{\alpha z}, \Phi_{\alpha}, L_{\alpha z}, 0)^t$ and $\boldsymbol{\lambda}_{M\alpha} = (A_{\alpha z}, 0, L_{\alpha z}, \Phi_{\alpha})^t$ associated with SH-modes in

the upper and lower media, we can write the boundary conditions as $\sum_{\alpha=3,4} b_{\alpha} \zeta_{\alpha} = \sum_{\alpha=3,4} c_{\alpha+4} \lambda_{\alpha}^*$. The speed v_I can be found from Eq. (17). In this case, $Z_{\Phi,33}^{(f)}(v) \equiv Z_{\Phi,33}^{(s)}(v)$ so that the SH-interfacial wave problem becomes equivalent to the Bleustein–Gulyaev wave problem on the mechanically free metallized surface, i.e., the fulfillment of conditions at the metallized interface leads to the vanishing of the total traction on the interface.

With the traction and potential equal to zero on the interface, using Eq. (6), one can show that the relation $(b_4/b_3)^2 = (c_8/c_7)^2 = -1$ holds. Hence, at $v = v_I = v_{BG}$ one has

$$\zeta_{M3} + B_S \zeta_{M4} = B(\lambda_{M3}^* - B_S \lambda_{M4}^*), \quad (18)$$

where $B_S = \pm i$, and B is a phase factor.

From Fig. 2 we infer, accounting for specific properties of SH-waves, that the interfacial wave does not appear if the SH-limiting wave does not produce potential. It can be shown that the wave vector of this bulk wave is directed along $\mathbf{m} + \mathbf{n} \tan \theta$, where

$$\tan \theta = - \frac{e_{14} + e_{25} \pm \sqrt{(e_{14} + e_{25})^2 - 4e_{24}e_{15}}}{2e_{24}}. \quad (19)$$

Computing the x - and y -projections of the group velocity, s_x and s_y , we find the inclination angle ψ of the interface to the coordinate axis x ,

$$\tan \psi = \frac{s_y}{s_x} = \frac{c_{44} \tan \theta + c_{45}}{c_{55} + c_{45} \tan \theta}. \quad (20)$$

Thus, there may be at most two orientations of the interface for which the localized solution does not exist. The conclusion applies to Bleustein–Gulyaev waves on a metallized mechanically free surface as well.

Note that the interfacial SH-wave on an interface of the type “180° domain wall,” Refs. 3–5, can also be viewed as formed of two Bleustein–Gulyaev waves propagating above and below the interface with different phases. Let the bicrystal be made up of identical piezoelectrics of symmetry 2 mm. The axes 2 lie anti-parallel to each other in the interface; the latter is a plane of symmetry. Taking into account the structure of the tensors of material constants¹⁵ one can check that “domains” will then have alike elastic and dielectric constants with respect to a common coordinate system. However, their piezomoduli will be of opposite sign. So, the partial modes above and below the interface produce displacements and tractions of like sign, but potentials and normal components of inductions have opposite signs. The boundary conditions can be written in terms of the vectors $\zeta_{\alpha} = (A_{\alpha z}, \Phi_{\alpha}, L_{\alpha z}, Q_{\alpha})^t$ and $\lambda_{\alpha} = (A_{\alpha z}, -\Phi_{\alpha}, L_{\alpha z}, -Q_{\alpha})^t$ associated with modes in the upper and lower medium, respectively. The fulfillment of the boundary conditions yielding the vanishing of traction and potential on the “180°-domain wall,” one can show that at $v_I = v_{BG}$ these vectors obey a relation similar to Eq. (18),

$$\zeta_3 + B_S \zeta_4 = B(\lambda_3^* - B_S \lambda_4^*), \quad (21)$$

where $B_S = \pm i$, and B is a phase factor.

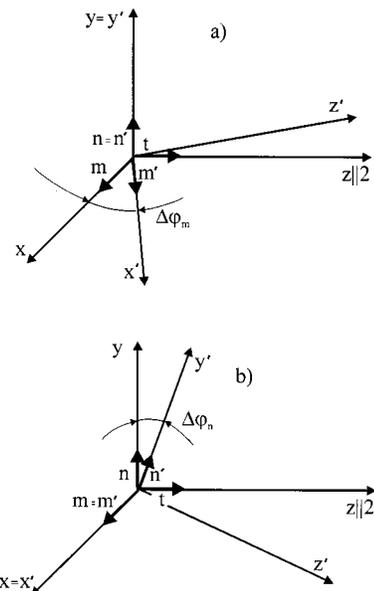


FIG. 3. Orientations of the crystal. In (a) the symmetry axis 2 lies in the plane $x'z'$ and makes an angle $\Delta\varphi_m$, $|\Delta\varphi_m| \ll 1$, with the axis z' (m-rotated crystal). In (b) the axis 2 is perpendicular to the axis x' but does not lie in the plane $x'z'$ making an angle $\Delta\varphi_n$, $|\Delta\varphi_n| \ll 1$, with the axis z' (n-rotated crystal).

III. LEAKY INTERFACIAL WAVES AND THE RESONANCE OF REFLECTION

This section deals with wave propagation in piezoelectric bicrystals in the case when the sagittal plane is not perpendicular to even-fold axes of symmetry. We introduce the coordinate system $x'y'z'$ in such a way that the plane $x'z'$ will be the interface and the axis x' will be the direction of the wave propagation. The axis y' is directed upward. Each of two parts of the bicrystal is oriented either as shown in Fig. 3(a) or as shown in Fig. 3(b). To characterize such a nonsymmetric configuration, we shall use a pair of angles $\Delta\Phi = (\Delta\varphi_{1I}, \Delta\varphi_{2J})$, where $I, J = n, m$. These angles specify the orientation of the upper, $\Delta\varphi_{1I}$, and lower, $\Delta\varphi_{2J}$, crystals. We assume that if $\Delta\varphi_{1I} = \Delta\varphi_{2J} = 0$, then the structure falls into the type either (a) a metallic layer of vanishingly small thickness is inserted into a piezoelectric parallel to an even-fold symmetry axis or (b) the “180°-domain wall” described at the end of the previous section.

Our attention will be focused on the analysis of specific features of the reflection under conditions of resonant excitation of a leaky wave which originates from the SH-interfacial wave. Note that the resonant reflection on the “180°-domain wall” in 6 mm crystals for one of the possible configurations, $\Delta\Phi = (0, \Delta\varphi_{2m})$ in our notation, has earlier been discussed in Ref. 16. Expressions for the coefficients of mode conversion have been derived using approximate solutions of Eqs. (1) and (2) found explicitly. In the present work, we shall employ a different method based on application of the perturbation theory to the eigenvalue problem Eq. (3). Expressions to be obtained apply to all piezocrystals having an even-fold symmetry axis in case (a) and to media of symmetry 2 mm, 4 mm, 6 mm, and ∞ mm in case (b).

Misorientation of crystals can cause transformation of

the interfacial SH-wave into a leaky one when the interfacial wave is “supersonic,” i.e., its speed v_I is larger than the limiting speed \hat{v}_T of in-plane waves. If $v_I > \hat{v}_T$, two situations are possible. There are two bulk waves in each crystal, an incident wave and a reflected one, for v close to v_I . There may also be two pairs of incident and reflected waves in crystals. For example, in crystals of symmetry 6 mm only the first case can be realized. On the other hand, both cases are admissible in tetragonal and orthorhombic media.

We want to derive approximate expressions describing the behavior of the coefficients of mode conversion of the quasi in-plane wave $\alpha=5$ incident from the upper medium in the plane $x'y'$ at angles which provide closeness of its trace speed to v_I . The boundary equations can be written in the form

$$\sum_{\alpha=5}^8 c'_\alpha \boldsymbol{\eta}'_\alpha = \sum_{\alpha=1}^4 b'_\alpha \boldsymbol{\xi}'_\alpha + b'_5 \boldsymbol{\xi}'_5, \quad \text{case (a),} \quad (22)$$

$$\sum_{\alpha=5}^8 c'_\alpha \boldsymbol{\eta}'_{M\alpha} = \sum_{\alpha=1}^4 b'_\alpha \boldsymbol{\xi}'_{M\alpha} + b'_5 \boldsymbol{\xi}'_{M5}, \quad \text{case (b),} \quad (23)$$

where $\boldsymbol{\xi}'_\alpha$ and $\boldsymbol{\eta}'_\alpha$ are constructed following Eq. (5), while

$$\boldsymbol{\xi}'_{M\alpha} = (\mathbf{A}'_\alpha, \Phi'_\alpha, \mathbf{L}'_\alpha, 0)^t \quad \text{and} \quad \boldsymbol{\eta}'_{M\alpha} = (\tilde{\mathbf{A}}'_\alpha, 0, \tilde{\mathbf{L}}'_\alpha, \tilde{\Phi}'_\alpha)^t. \quad (24)$$

The vectors $\boldsymbol{\xi}'_\alpha$, $\boldsymbol{\xi}'_{M\alpha}$ and $\boldsymbol{\eta}'_\alpha$, $\boldsymbol{\eta}'_{M\alpha}$ are associated with the modes in the upper and lower media, respectively. According to Eqs. (9) and (12), the mode $\alpha=5$ in the bottom crystal plays the role of the transmitted wave, the mode $\alpha=1$ in the upper medium is the reflected wave. The mode $\alpha=2$ in the upper crystal and the mode $\alpha=6$ in the lower medium are either inhomogeneous or the second reflected wave and the second transmitted wave, respectively.

So, in both cases the coefficients of mode conversion $R_{5 \rightarrow \alpha} = b'_\alpha / b'_5$ and $T_{5 \rightarrow \alpha} = c'_\alpha / b'_5$ initially have the form of ratios of determinants of 8×8 matrices constructed from $\boldsymbol{\xi}'_\alpha$ and $\boldsymbol{\eta}'_\alpha$ or $\boldsymbol{\xi}'_{M\alpha}$ and $\boldsymbol{\eta}'_{M\alpha}$. To evaluate these coefficients, we must express the eigenvectors $\boldsymbol{\xi}'_\alpha$ of the matrix $\hat{\mathbf{N}}(\mathbf{m}', \mathbf{n}')$ Eq. (4), where either $\mathbf{m}' = \mathbf{m} \cos(\Delta\varphi_m) + \mathbf{t} \sin(\Delta\varphi_m)$ and $\mathbf{n}' = \mathbf{n}$ [Fig. 3(a)] or $\mathbf{m}' = \mathbf{m}$ and $\mathbf{n}' = \mathbf{n} \cos(\Delta\varphi_n) + \mathbf{t} \sin(\Delta\varphi_n)$ [Fig. 3(b)], in terms of its “unperturbed” eigenvectors $\boldsymbol{\xi}_\alpha$ such that the components of the vectors \mathbf{A}'_α and \mathbf{L}'_α involved in $\boldsymbol{\xi}'_\alpha$ will be specified with respect the coordinate frame $x'y'z'$, while the material constants of the medium are specified with respect to the crystallographic coordinate system xyz (Fig. 3). The result is

$$\boldsymbol{\xi}'_\alpha = (\hat{\mathbf{I}} + \hat{\mathbf{F}}_J) \boldsymbol{\xi}_\alpha, \quad \hat{\mathbf{F}}_J = \sum_{k=1}^{\infty} \hat{\mathbf{F}}_J^{(k)} (\Delta\varphi_J)^k, \quad (25)$$

where the 8×8 real matrices $\hat{\mathbf{F}}_J^{(k)}$, $k=1, 2, \dots$, interrelate $\boldsymbol{\xi}_\alpha$ and the corrections $\boldsymbol{\xi}_\alpha^{(k)'}$ of the order $(\Delta\varphi_J)^k$, $\boldsymbol{\xi}_\alpha^{(k)'} = [\hat{\mathbf{F}}_J^{(k)} \boldsymbol{\xi}_\alpha] (\Delta\varphi_J)^k$, the suffix J indicates the type of rotation. Considering the “unperturbed” vectors $\boldsymbol{\xi}_\alpha$, $\alpha=1, \dots, 8$ as a set of linearly independent vectors referred to the coordinate system $x'y'z'$, we can decompose the vector $\boldsymbol{\xi}'_\alpha$ Eq. (25) with respect to $\boldsymbol{\xi}_\alpha$ to obtain

$$\boldsymbol{\xi}'_\alpha = \boldsymbol{\xi}_\alpha + \sum_{k=1}^{\infty} \left[\sum_{\beta=1}^8 (\hat{\mathbf{F}}_J^{(k)})_{\beta\alpha} \boldsymbol{\xi}_\beta \right] (\Delta\varphi_J)^k, \quad (26)$$

where the symbols $(\hat{\mathbf{F}}_J^{(k)})_{\beta\alpha}$ denote the contractions $\boldsymbol{\xi}_\beta \cdot \hat{\mathbf{T}}\hat{\mathbf{F}}_J^{(k)} \boldsymbol{\xi}_\alpha$.

The matrix $\hat{\mathbf{F}}_J$ has the form $\hat{\mathbf{F}}_J = \hat{\mathbf{X}}_J + \hat{\mathbf{M}}_J + \hat{\mathbf{X}}_J \hat{\mathbf{M}}_J$. Here $\hat{\mathbf{X}}_J$ is an 8×8 real block matrix,

$$\hat{\mathbf{X}}_J = \begin{pmatrix} \hat{\mathbf{R}}_J & \hat{\mathbf{O}} \\ \hat{\mathbf{O}} & \hat{\mathbf{R}}_J \end{pmatrix}, \quad J=n, m, \quad (27)$$

and $\hat{\mathbf{R}}_J$ are 4×4 “rotation” matrices with nonzero elements

$$\begin{aligned} R_{m,11} &= R_{m,33} = -2 \sin^2(\Delta\varphi_m/2), \\ R_{m,13} &= -R_{m,31} = \sin(\Delta\varphi_m), \\ R_{n,22} &= R_{n,33} = -2 \sin^2(\Delta\varphi_n/2), \\ R_{n,23} &= -R_{n,32} = \sin(\Delta\varphi_n). \end{aligned} \quad (28)$$

for deflections shown in Fig. 3(a) and (b), respectively.

The real 8×8 matrix $\hat{\mathbf{M}}_J = \sum_{k=1}^{\infty} \hat{\mathbf{M}}_J^{(k)} (\Delta\varphi_J)^k$, like the matrix $\hat{\mathbf{F}}_J$ Eq. (25), interrelates $\boldsymbol{\xi}_\alpha$ and the eigenvector $\boldsymbol{\xi}'_\alpha$ of the matrix $\hat{\mathbf{N}}(n', m')$. The difference is that the vectors \mathbf{A}'_α and \mathbf{L}'_α involved in $\boldsymbol{\xi}'_\alpha$ are specified with respect to the frame xyz (Fig. 3). The real matrix $\hat{\mathbf{M}}_J^{(k)}$ interrelates $\boldsymbol{\xi}_\alpha$ and the correction $\boldsymbol{\xi}_\alpha^{(k)''}$,

$$\boldsymbol{\xi}_\alpha^{(k)''} = [\hat{\mathbf{M}}_J^{(k)} \boldsymbol{\xi}_\alpha] (\Delta\varphi_J)^k = \sum_{\beta=1}^8 [(\hat{\mathbf{M}}_J^{(k)})_{\beta\alpha} \boldsymbol{\xi}_\beta] (\Delta\varphi_J)^k, \quad (29)$$

where $(\hat{\mathbf{M}}_J^{(k)})_{\beta\alpha} = \boldsymbol{\xi}_\beta \cdot \hat{\mathbf{T}}\hat{\mathbf{M}}_J^{(k)} \boldsymbol{\xi}_\alpha$. Explicit expressions for the first and second order terms are given in the Appendix.

Taking into account the symmetry of the medium and Eq. (12), one can prove that

$$\begin{aligned} (\hat{\mathbf{F}}_J^{(2k+1)})_{\alpha\beta} &= 0 \\ &\text{for } \alpha, \beta = 1, 2, 5, 6 \text{ and for } \alpha, \beta = 3, 4, 7, 8; \\ (\hat{\mathbf{F}}_J^{(2k)})_{\alpha\beta} &= (\hat{\mathbf{F}}_J^{(2k)})_{\beta\alpha} = 0 \\ &\text{for } \alpha = 1, 2, 5, 6 \text{ and } \beta = 3, 4, 7, 8. \end{aligned} \quad (30)$$

Also, it can be shown that $(\hat{\mathbf{F}}_J^{(k)})_{\beta\alpha}$ are unchanged under the change of sign of the piezomoduli. Hence, as applied to the “180°-domain wall” problem, letting Eq. (26) be vectors attached to the top “domain,” we can also write the vectors $\boldsymbol{\eta}'_\alpha$ attached to the bottom “domain” in the form Eq. (26) with $\boldsymbol{\eta}'_\alpha$ and $\boldsymbol{\eta}_\alpha$ replacing $\boldsymbol{\xi}'_\alpha$ and $\boldsymbol{\xi}_\alpha$, respectively; the expressions will involve alike contractions $(\hat{\mathbf{F}}_J^{(k)})_{\beta\alpha}$.

Introducing the vectors $\boldsymbol{\xi}'_\alpha$, $\boldsymbol{\eta}'_\alpha$, $\boldsymbol{\xi}'_{M\alpha}$, and $\boldsymbol{\eta}'_{M\alpha}$ taken in the form Eq. (26) into the appropriate determinants, making use of the aforementioned properties, taking into account the fact that the top and the bottom parts of the bicrystal are mechanically identical in the absence of misorientation and therefore $\boldsymbol{\xi}_\alpha \equiv \boldsymbol{\eta}_\alpha \equiv \boldsymbol{\xi}_{M\alpha} \equiv \boldsymbol{\eta}_{M\alpha}$, $\alpha=1, 2, 5, 6$, we obtain

$$b'_5(v, \Delta\varphi) = f(v) + \sum_{k,l} f_{k,l}^{(5)}(v) (\Delta\varphi_{1l})^k (\Delta\varphi_{2J})^l,$$

$$c'_5(v, \Delta\varphi) = f(v) + \sum_{k,l} g_{k,l}^{(5)}(v) (\Delta\varphi_{1l})^k (\Delta\varphi_{2l})^l, \quad (31)$$

$$b'_\alpha(v, \Delta\varphi) = \sum_{k,l} f_{k,l}^{(\alpha)}(v) (\Delta\varphi_{1l})^k (\Delta\varphi_{2l})^l, \quad \alpha = 1, 2,$$

$$c'_6(v, \Delta\varphi) = \sum_{k,l} g_{k,l}^{(6)}(v) (\Delta\varphi_{1l})^k (\Delta\varphi_{2l})^l,$$

where the sum $k+l$ has even values starting from 2, and

$$b'_\alpha(v, \Delta\varphi) = \sum_{k,l} f_{k,l}^{(\alpha)}(v) (\Delta\varphi_{1l})^k (\Delta\varphi_{2l})^l, \quad (32)$$

$$c'_{\alpha+4}(v, \Delta\varphi) = \sum_{k,l} g_{k,l}^{(\alpha+4)}(v) (\Delta\varphi_{1l})^k (\Delta\varphi_{2l})^l, \quad \alpha = 3, 4,$$

where the sum $k+l$ has odd values. In Eq. (31) $f(v) = \|\xi_1 \xi_2 \xi_5 \xi_6\| f_I(v)$, where $f_I(v)$ is the interfacial wave speed function and $\|\dots\|$ means the determinant of the 4×4 matrix the columns of which are the vectors $\xi_\alpha = (A_{\alpha x}, A_{\alpha y}, L_{\alpha x}, L_{\alpha y})^t$.

Allowing only for quadratic and linear terms in Eqs. (31) and (32), letting $v = v_I$ in $f_{k,l}^{(\alpha)}(v)$ and $g_{k,l}^{(\alpha)}(v)$, using Eqs. (11), (18), and (21) together with the equalities

$$(\hat{F}_J^{(1)})_{\alpha\beta} = -(\hat{F}_J^{(1)})_{\beta\alpha},$$

$$(\hat{F}_J^{(2)})_{\alpha\beta} + (\hat{F}_J^{(2)})_{\beta\alpha} = \sum_{\gamma=1}^8 (\hat{F}_J^{(1)})_{\alpha\gamma} (\hat{F}_J^{(1)})_{\gamma\beta}, \quad (33)$$

which follow from the relation $\xi'_\alpha \cdot \hat{\mathbf{T}} \xi'_\beta = \delta_{\alpha\beta}$, one can bring $R_{5 \rightarrow \alpha}$ and $T_{5 \rightarrow \alpha}$ into the form

$$R_{5 \rightarrow \alpha}(v) = \frac{b'_\alpha}{b'_5} \approx \frac{i v_\alpha}{v - v_I}, \quad \alpha = 1, 2, \quad (34)$$

$$T_{5 \rightarrow 5}(v) = \frac{c'_5}{b'_5} \approx \frac{v - \bar{v}}{v - v_I}, \quad T_{5 \rightarrow 6}(v) = \frac{c'_6}{b'_5} \approx \frac{i v_6}{v - v_I}, \quad (35)$$

$$R_{5 \rightarrow 3}(v) = \frac{b'_3}{b'_5} \approx \frac{i v_5}{v - v_I}, \quad R_{5 \rightarrow 4}(v) = \frac{b'_4}{b'_5} \approx B_S R_{5 \rightarrow 3}, \quad (36)$$

$$T_{5 \rightarrow 7}(v) = \frac{c'_7}{b'_5} \approx B_S T_{5 \rightarrow 3}, \quad T_{5 \rightarrow 8}(v) = \frac{c'_8}{b'_5} \approx -B_S T_{5 \rightarrow 3}.$$

In Eqs. (34)–(36) the quantities v_5 and v_α , $\alpha = 1, 2, 6$ are linear and quadratic functions of deviation angles, respectively,

$$v_5 = \frac{G_5}{D}, \quad v_\alpha = \frac{G_\alpha G_5}{D}, \quad \alpha = 1, 2, 6, \quad (37)$$

$$G_\alpha = (\hat{F}_I^{(1)})_{S\alpha} \Delta\varphi_{1l} - B (\hat{F}_J^{(1)})_{S*\alpha} \Delta\varphi_{2l}, \quad \alpha = 1, 2, 5, 6, \quad (38)$$

$$(\hat{F}_K^{(1)})_{S\alpha} = \mathbf{S} \cdot \hat{\mathbf{T}} \hat{\mathbf{F}}_K^{(1)} \xi_\alpha, \quad (\hat{F}_K^{(1)})_{S*\alpha} = \mathbf{S}^* \cdot \hat{\mathbf{T}} \hat{\mathbf{F}}_K^{(1)} \xi_\alpha, \quad (39)$$

$$\mathbf{S} = \xi_3 + B_S \xi_4,$$

the indices I and J indicate the type of rotation of the upper and lower crystal, respectively, B is given by Eq. (18) and by Eq. (21) for the metallized interface and for the “180° domain wall,” respectively,

$$D = i4\rho v_I \left\{ \frac{|\mathbf{A}_{z3}|^2}{p_3 - p_3^*} + \frac{|\mathbf{A}_{z4}|^2}{p_4 - p_4^*} + B_S \left[\frac{\mathbf{A}_{z4} \mathbf{A}_{z3}^*}{p_4 - p_3^*} - \frac{\mathbf{A}_{z3} \mathbf{A}_{z4}^*}{p_3 - p_4^*} \right] \right\}$$

$$= 4\rho\omega \int_0^\infty |\mathbf{A}_{z3} e^{ikp_3 y} + B_S \mathbf{A}_{z4} e^{ikp_4 y}|^2 dy > 0. \quad (40)$$

To within a factor canceling out in the course of evaluations, D equals the derivative $\partial f_I / \partial v$ of the speed function $f_I(v)$ at the point $v = v_I$.

The complex speeds $v_I = v'_I - i v''_I$ and $\bar{v} = v'_I - i \bar{v}''_I$ have alike real parts and different imaginary ones,

$$v'_I = v_I + \frac{\Xi_1 + \Xi_2}{2D}, \quad v''_I = \frac{|G_1|^2 + |G_5|^2}{2D},$$

$$\bar{v}''_I = \frac{|G_1|^2 - |G_5|^2}{2D} \quad (41)$$

in the case when $\alpha = 2, 6$ are inhomogeneous modes, and

$$v'_I = v_I + \frac{\Xi_1}{2D}, \quad v''_I = \frac{\sum_{\alpha=1,2,5,6} |G_\alpha|^2}{2D},$$

$$\bar{v}''_I = \frac{\sum_{\alpha=1,2,6} |G_\alpha|^2 - |G_5|^2}{2D} \quad (42)$$

if $\alpha = 1, 2, 5, 6$ are homogeneous modes; Ξ_1 and Ξ_2 are real and equal to

$$\Xi_1 = i \left[\{(\hat{F}_I^{(2)})_{S*S} - (\hat{F}_I^{(2)})_{S*S^*}\} (\Delta\varphi_{1l})^2 + \{(\hat{F}_J^{(2)})_{S*S} - (\hat{F}_J^{(2)})_{S*S^*}\} (\Delta\varphi_{2l})^2 + \sum_{\alpha=1,2,5,6} \{B^* (\hat{F}_I^{(1)})_{S\alpha} (\hat{F}_J^{(1)})_{S\alpha} - B (\hat{F}_I^{(1)})_{S*\alpha} (\hat{F}_J^{(1)})_{S*\alpha}\} \Delta\varphi_{1l} \Delta\varphi_{2l} \right], \quad (43)$$

$$\Xi_2 = i[G_2 \bar{G}_2 - G_6 \bar{G}_6], \quad (44)$$

where $\bar{G}_\alpha = (\hat{F}_I^{(1)})_{S*\alpha} \Delta\varphi_{1l} - B^* (\hat{F}_J^{(1)})_{S\alpha} \Delta\varphi_{2l}$, $\alpha = 2, 6$, and $(\hat{F}_K^{(2)})_{S*S} = [(\hat{F}_K^{(2)})_{S*S^*}]^* = \mathbf{S} \cdot \hat{\mathbf{T}} \hat{\mathbf{F}}_K^{(2)} \mathbf{S}^*$.

Let us analyze the expressions obtained. Note first of all that the complex pole $v = v_I$ is associated with the leaky solution. The leaky wave is the initial interfacial wave to which the modes $\alpha = 1, 2$ and $\alpha = 5, 6$ are added in the upper and lower media, respectively, with amplitudes small in the measure of smallness of deviation angles. At least two of the adjacent modes ($\alpha = 1$ and $\alpha = 5$) are bulk “reflected,” i.e., they carry energy away from the interface, causing wave attenuation.

“Mode mixing” appears due to the characteristic anisotropy of the crystal. The values of $(\hat{F}_K^{(1)})_{S\alpha}$ show to what extent the modes incorporated into the interfacial wave become mixed with the other modes when the sagittal plane is not perpendicular to the symmetry axis. Of interest is the fact that losses caused by radiation into the bulk of the upper and lower parts of the structure add up in an “interference” manner rather than additively [one might think that the expres-

sion for v_l'' would involve not $|(\hat{F}_l^{(1)})_{S\alpha}\Delta\varphi_{1l} - B(\hat{F}_j^{(1)})_{S*\alpha}\Delta\varphi_{2j}|^2$ but $|(\hat{F}_l^{(1)})_{S\alpha}\Delta\varphi_{1l}|^2 + |(\hat{F}_j^{(1)})_{S*\alpha}\Delta\varphi_{2j}|^2$.

Equations (41) and (42) allow one to estimate the first term in the expansion of v_l with respect to deviation angles. In view of Eq. (31) v_l involves only terms proportional to $(\Delta\varphi_{1l})^k(\Delta\varphi_{2j})^l$, where $k+l$ has even values, so that the second correction to v_l is of the order of $(\Delta\varphi)^4$. One can also show that v_l'' appears to be proportional to the relevant coefficient of electromechanical coupling, i.e., to a quantity of the type $e^2/\epsilon c$.

Consider now the behavior of the coefficients of mode conversion. Note that the coefficients as given by Eqs. (34)–(36) satisfy the law of energy conservation. Due to Eqs. (10) and (6) the latter implies that $|R_{5\rightarrow 1}|^2 + |T_{5\rightarrow 5}|^2 = 1$ and $\sum_{\alpha=1,2}[|R_{5\rightarrow\alpha}|^2 + |T_{5\rightarrow\alpha+4}|^2] = 1$ when $\alpha=2,6$ are inhomogeneous and homogeneous modes, respectively.

Assuming the modes $\alpha=2,6$ to be inhomogeneous we infer from Eqs. (34) and (35) that the magnitude of the reflection coefficient $R_{5\rightarrow 1}$ increases starting from values of the order of $(\Delta\varphi)^2$ up to the value $|v_1/v_1''| \propto 1$ as v approaches v_1' . The coefficient of transmission $T_{5\rightarrow 5}$ decreases from values close to unity until the value $|\bar{v}''/v_1''|$ at $v = v_1'$. These changes occur within the interval $|v - v_1'|$ of width v_1'' . The parameter \bar{v}'' may vanish, and then $|R_{5\rightarrow 1}| = 1$ and $|T_{5\rightarrow 5}| = 0$. One can show that $\bar{v}'' = 0$ in 6 mm crystals (the interface is rotated arbitrarily about the axis 6) and in 2 mm (4 mm) crystals (the interface is a plane of symmetry) for configurations of the type $\Delta\Phi = (\Delta\varphi_{1m}, \Delta\varphi_{2m})$ or $\Delta\Phi = (\Delta\varphi_{1n}, \Delta\varphi_{2n})$ and for either of the two types of interface being discussed in this section.

Let $\alpha=2,6$ be bulk modes. When v approaches v_1' , the transmission coefficient $T_{5\rightarrow 5}$ decreases to $|\bar{v}''/v_1''|$ whereas two reflection coefficients, $R_{5\rightarrow 1}$ and $R_{5\rightarrow 2}$, together with the transmission coefficient $T_{5\rightarrow 6}$ rise up to $|v_\alpha/v_1''| \propto 1$, $\alpha = 1,2,6$, at $v = v_1'$. So, the amplitudes of all bulk waves generally become comparable near v_1' . In principle \bar{v}'' may be equal to zero, and then the incident wave hardly excites the mode into which it converts except within the range $|v - v_1'| \propto v_1''$ ($\alpha=5$ in the bottom medium). However, in this case \bar{v}'' does not vanish so “easily” as it does in the previous case. One cannot formulate general conditions for $\bar{v}'' = 0$.

The coefficients $R_{5\rightarrow\alpha}$ and $T_{5\rightarrow\alpha+4}$, $\alpha=3,4$, suffer the strongest growth in magnitude. By Eqs. (36)–(38), they become of the order $|\Delta\varphi|^{-1} \gg 1$ at $v = v_1'$.

Note that the phase of the coefficients also changes abruptly in the vicinity of v_1' . Using Eqs. (34)–(36) we deduce that near v_1' the phase ϕ of the coefficients except $T_{5\rightarrow 5}$ depends upon v as $\phi(v) = -\cotan^{-1}[(v - v_1')/v_1'']$ and shows a “jump” of $\pi/2$ as v passes from $v_1' + v_1''$ to $v_1' - v_1''$. The phase of $T_{5\rightarrow 5}$ exhibits a somewhat different behavior, but it also changes by approximately π within this interval.

It should be said that Eqs. (34)–(36) correctly describe the behavior of the coefficients until attenuation of the leaky wave due to the radiation is far larger than that due to dissipation processes.

This damping can be characterized by an additional imaginary component, $-i v_d''$, to the leaky wave speed. The condition $v_l'' \gg v_d''$ will certainly break down as $\Delta\varphi_{1,2} \rightarrow 0$. To describe, at least qualitatively, the behavior of the coefficients when $v_l'' \leq v_d''$, one can just replace v_l' by $v_l' - i v_d''$ in Eqs. (34)–(36). It becomes then clear that, for example, the maximum of $R_{5\rightarrow\alpha}$ and $T_{5\rightarrow\alpha+4}$, $\alpha=3,4$, is limited by a value of the order $(v_1'/v_d'')^{1/2}$, and their magnitude tends to zero as $\Delta\varphi_{1,2} \rightarrow 0$.

IV. WAVES IN 6 mm CRYSTALS

Using Eq. (41) one can evaluate explicitly the quadratic term of v_l'' for 6 mm (∞ mm) media to discover that it becomes zero for the configurations

$$\begin{aligned} \Delta\Phi &= (\Delta\varphi_{1m}, \Delta\varphi_{2m} = \Delta\varphi_{1m}), \\ \Delta\Phi &= (\Delta\varphi_{1n}, \Delta\varphi_{2n} = \Delta\varphi_{1n}) \end{aligned} \quad (45)$$

in either of the structures being considered.

We cannot establish whether, on the domain wall in 6 mm crystals, the interfacial wave remains localized or whether it transforms into the leaky solution for configurations Eq. (45). However, one can show that if the leaky wave appears, then the imaginary component of its speed varies as $(\Delta\varphi)^6$.

Indeed, by Eqs. (31) and (32), at $v = v_l$ the coefficients b_1' and c_5' are of order $(\Delta\varphi)^{2m+2}$, and the coefficients $b_{3,4}'$ and $c_{7,8}'$ are of order $|\Delta\varphi|^{2k+1}$, where $m, k = 1, 2, \dots$, since due to Eqs. (34)–(39) and (41) the quadratic terms in b_1' and c_5' as well as the linear terms in $b_{3,4}'$ and $c_{7,8}'$ become zero simultaneously with v_l'' Eq. (41). Therefore the amplitude b_l of bulk waves incorporated into the leaky solution is of the order $|b_1'/b_3'| \propto |\Delta\varphi|^{2(m-k)+1}$. The leaky wave attenuates because of the radiation of the energy, so that v_l'' is proportional to the normal component of the energy flux, i.e., $v_l'' \propto |b_l|^2 \propto (\Delta\varphi)^{4(m-k)+2}$. Hence, when the quadratic term vanishes, the next possible dependence is $v_l'' \propto (\Delta\varphi)^6$ which completes the proof.

At the same time, when Eq. (45) holds, one can prove that the interfacial wave on a metallic layer remains purely localized and incorporates the inhomogeneous modes $\alpha = 2, 3, 4$ and $\alpha = 6, 7, 8$ in the top and bottom medium, respectively. In this instance, the “domains” remain completely identical. Accordingly, there is the same set of plane modes above and below the interface. The localized wave then exists provided that the conditions

$$\begin{aligned} \sum_{\alpha=2}^4 b_\alpha' \mathbf{L}'_\alpha &= \sum_{\alpha=6}^8 c_\alpha' \mathbf{L}'_\alpha, & \sum_{\alpha=2}^4 b_\alpha' \mathbf{A}'_\alpha &= \sum_{\alpha=6}^8 c_\alpha' \mathbf{A}'_\alpha, \\ \sum_{\alpha=2}^4 b_\alpha' \Phi'_\alpha &= \sum_{\alpha=6}^8 c_\alpha' \Phi'_\alpha & &= 0 \end{aligned} \quad (46)$$

are obeyed. In 6 mm crystals, the parameters of modes $\alpha = 1, 5$ (quasi transverse modes) can be found analytically for arbitrary values of $\Delta\varphi_m$ and $\Delta\varphi_n$. These waves do not produce an electrical potential (but they produce the induction having nonzero normal projection). Letting $b'_\alpha = \Phi'_\alpha$, $\alpha = 2, 3, 4$, and $c'_\alpha = -\Phi'_\alpha$, $\alpha = 6, 7, 8$, we get that due to Eq. (7)

and $\Phi'_{1,5}=0$ the first two equalities in Eq. (46) hold identically and $\sum_{\alpha=2}^4(\Phi'_\alpha)^2 \equiv -\sum_{\alpha=6}^8(\Phi'_\alpha)^2 \equiv -\sum_{\alpha=2}^4(\Phi'^*_\alpha)^2$.

So, if the real function $f_I(\nu, \Delta\varphi) = i\sum_{\alpha=2}^4(\Phi'_\alpha)^2$ vanishes, all conditions Eqs. (46) will be satisfied. Obviously $f_I(\nu_I, 0) = 0$ and it is easily checked that $\partial f_I / \partial \nu \neq 0$ at this point. Hence, in view of analyticity of $f_I(\nu, \Delta\varphi)$ the equation $f_I(\nu, \Delta\varphi) = 0$ is solvable with respect to ν within an interval of deviation angles. Thus, for orientations Eq. (45), there appears a branch of localized ‘‘supersonic’’ interfacial waves rather than one of leaky waves. Note that for the nonsymmetric geometry of propagation the interfacial wave is not ‘‘equivalent’’ to a wave on a metallized mechanically free surface as it does not leave the interface traction free.

V. CONCLUDING REMARKS

In the present paper, we have performed a general analysis of the possibility of existence of SH-localized waves on two types of interface between rigidly bonded piezoelectrics. When the interface is not metallized, one can establish the admissible number of solutions and formulate a condition ruling out the appearance of the interfacial wave. Remembering results obtained in Ref. 10 we see that in this case the conclusions do not provide more information than those made in Ref. 8 regarding localized waves of the generic type on the nonmetallized rigid contact between piezoelectrics of unrestricted symmetry. On the other hand, as applied to the SH-wave problem on the metallized contact, the impedance method permits derivation of a sufficient condition for the existence of the interfacial wave. In addition it has been found that SH-interfacial waves on the metallized interface between identical media exists for all orientations of the interface, with the exception of not more than two orientations. It has also been shown that the ‘‘supersonic’’ interfacial wave on a metallic layer in crystals of symmetry 6 mm may remain purely localized even if the sagittal plane is not perpendicular to a common direction of the six-fold axes of symmetry.

However, misorientation of the bicrystal commonly causes the transformation of ‘‘supersonic’’ interfacial waves into leaky waves. The appearance of a leaky wave branch results in specific peculiarities of the resonant type in the behavior of the coefficients of mode conversion. For example, it may be that a bulk wave experiences total reflection at the incident angle for which its trace speed is equal to the real part of the leaky wave speed. But the same wave will almost totally transmit through the interface if the incident angle slightly differs from the resonant value. Using expressions obtained one can estimate the resonant angle and the width of the resonance.

We considered that the layer inserted into the crystal was of vanishingly small thickness. An isotropic layer of finite thickness directly does not affect ‘‘the mode mixing’’ between SH- and in-plane modes, but the real part of the leaky wave speed becomes then frequency dependent. As a result, using a layer of finite thickness one can control the value of the resonant angle by varying the frequency of the incident wave. Note that if the inserted layer is anisotropic, then it can additionally ‘‘mix’’ modes owing to its charac-

teristic anisotropy. The leaky wave can then appear without misorientation of crystals.¹⁷

ACKNOWLEDGMENTS

We are grateful to V. I. Alshits and A. L. Shuvalov for fruitful discussions and to the Russian Foundation for Basic Research (Grant No. 98-02-16077) for partial financial support. We also thank the reviewers for correcting language mistakes and pointing out a number of misprints in the initial text.

APPENDIX

Explicitly $(\hat{M}_J^{(k)})_{\beta\alpha}$, $k=1,2$, read as

$$(\hat{M}_J^{(1)})_{\beta\alpha} = \frac{(\partial \hat{N} / \partial \varphi_J)_{\alpha\beta}}{p_\alpha - p_\beta}, \quad \alpha \neq \beta, \quad (\text{A1})$$

$$(\hat{M}_J^{(2)})_{\beta\alpha} = \frac{1}{p_\alpha - p_\beta} \left[\sum_{\gamma \neq \alpha} \left(\frac{\partial \hat{N}}{\partial \varphi_J} \right)_{\beta\gamma} (\hat{M}_J^{(1)})_{\gamma\alpha} - (\hat{M}_J^{(1)})_{\beta\alpha} \left(\frac{\partial \hat{N}}{\partial \varphi_J} \right)_{\alpha\beta} + 0.5 \left(\frac{\partial^2 \hat{N}}{\partial \varphi_J^2} \right)_{\alpha\beta} \right], \quad \alpha \neq \beta, \quad (\text{A2})$$

and $(\hat{M}_J^{(1)})_{\alpha\alpha} = 0$, $(\hat{M}_J^{(2)})_{\alpha\alpha} = -0.5 \sum_{\beta \neq \alpha} (\hat{M}_J^{(1)})_{\alpha\beta}^2$. Substituting the expressions for \mathbf{m}' and \mathbf{n}' in Eq. (4) and using Eq. (2) gives

$$\left(\frac{\partial \hat{N}}{\partial \varphi_m} \right)_{\alpha\beta} = \mathbf{U}_\alpha [p_\alpha(nt) + p_\beta(tn) + (tm) + (mt)] \mathbf{U}_\beta, \quad (\text{A3})$$

$$\left(\frac{\partial \hat{N}}{\partial \varphi_n} \right)_{\alpha\beta} = \mathbf{U}_\alpha [p_\alpha p_\beta [(nt) + (tn)] + p_\alpha(tm) + p_\beta(mt)] \mathbf{U}_\beta, \quad (\text{A4})$$

$$\left(\frac{\partial^2 \hat{N}}{\partial \varphi_m^2} \right)_{\alpha\beta} = 2 \mathbf{U}_\alpha [(tt) - (tn)(nn)^{-1}(nt)] \mathbf{U}_\beta + (p_\alpha - p_\beta) \mathbf{U}_\alpha \mathbf{V}_\beta, \quad (\text{A5})$$

$$\left(\frac{\partial^2 \hat{N}}{\partial \varphi_n^2} \right)_{\alpha\beta} = (p_\alpha - p_\beta) \mathbf{U}_\alpha \mathbf{V}_\beta + 2 \mathbf{U}_\alpha [p_\alpha p_\beta [(tt) + [(nt) + (tn)](nn)^{-1}[(nt) + (tn)]] - p_\alpha [(nt) + (tn)](nn)^{-1}(tm) - p_\beta(mt) \times (nn)^{-1}[(nt) + (tn)] - (mt)(nn)^{-1}(tm)] \mathbf{U}_\beta. \quad (\text{A6})$$

¹S. V. Biryukov, Yu. V. Gulyaev, V. V. Krylov, and V. P. Plessky, *Surface Acoustic Waves in Inhomogeneous Media* (Springer, Berlin, 1995).

²A. H. Nayfen, *Wave Propagation in Layered Media* (North-Holland, Amsterdam, 1995).

³C. Maerfeld and P. Tournois, ‘‘Pure shear wave guided by the interface of two semi-infinite media,’’ *Appl. Phys. Lett.* **19**, 117–118 (1971).

⁴G. G. Kessenikh, V. N. Lyubimov, and D. G. Sannikov, ‘‘Surface elastopolarization waves at domain boundaries in a ferroelectric,’’ *Sov. Phys. Crystallogr.* **17**, 512–514 (1972).

⁵J. C. Peusin and F. C. Lissalde, ‘‘Domain wall electroacoustic wave in 6

- mm and 4 mm ferroelectrics," *Ferroelectrics* **7**, 327–33 (1974).
- ⁶V. V. Filippov and V. N. Lyubimov, "Surface shear elastic waves on an dielectric-piezoelectric interface," *Vestzi Akademii navuk BSSR, seria fizika-matematichnih navuk* **3**, 69–73 (1982) (in Russian).
- ⁷J. Lothe and D. M. Barnett, "Integral formalism for surface waves in piezoelectric crystals. Existence considerations," *Appl. Phys.* **47**, 1799–1807 (1976).
- ⁸J. Lothe and D. M. Barnett, "Further development of the theory for surface waves in piezoelectric crystals," *Phys. Norv.* **8**, 239–254 (1976).
- ⁹G. A. Maugin, *Continuum Mechanics of Electromagnetic Solids* (North-Holland, Amsterdam, 1988).
- ¹⁰M. Abbudi and D. M. Barnett, "On the existence of interfacial (Stoneley) waves in bonded piezoelectric half-spaces," *Proc. R. Soc. London, Ser. A* **429**, 587 (1990).
- ¹¹V. I. Alshits, D. M. Barnett, A. N. Darinskii, and J. Lothe, "On the existence problem for localized acoustic waves on the interface between two piezocrystals," *Wave Motion* **20**, 233–244 (1994).
- ¹²J. L. Bleustein, "A new surface wave in piezoelectric materials," *Appl. Phys. Lett.* **13**, 412–413 (1968).
- ¹³Yu. V. Gulyaev, "Electroacoustic Surface Waves in Solids," *JETP Lett.* **9**, 37–38 (1969).
- ¹⁴G. Koerber and R. F. Vogel, "Generalized Bleustein modes," *IEEE Trans. Sonics Ultrason.* **SU-19**, 3–8 (1972).
- ¹⁵Yu. I. Sirotn and M. P. Shaskolskaya, *Fundamentals of Crystal Physics* (Mir, Moscow, 1982).
- ¹⁶V. I. Alshits, A. N. Darinskii, and A. L. Shuvalov, "Acoustoelectric waves in bicrystal media in conditions of a rigid contact or a vacuum gap at an interface," *Crystallogr. Rep.* **38**, 147–158 (1993).
- ¹⁷A. N. Darinskii and G. Maugin, "Features of elastic wave reflection in the piezoelectric—thin solid layer—piezoelectric structure," *Acust. Acta Acust.* **84**, 455–464 (1998).

Bjerknes forces between two bubbles in a viscous fluid

Alexander A. Doinikov

Institute of Nuclear Problems, Byelorussian State University, 11 Bobruiskaya St., Minsk 220050, Belarus

(Received 7 December 1998; accepted for publication 1 September 1999)

A refined expression for the secondary Bjerknes force between two spherical gas bubbles in a viscous incompressible fluid is derived, allowing for the dipole oscillations of the bubbles and acoustic streaming around them. The investigation is motivated by the fact that the existing theory [see, for example, E. A. Zabolotskaya, *Sov. Phys. Acoust.* **30**, 365–368 (1984)] neglects both of these processes, taking into account only linear viscous effects due to the radial oscillations of the bubbles. More correct calculations reveal that, provided the viscous penetration depth in the surrounding fluid is comparable to the bubble size, the forces on the bubbles differ noticeably from values given by the former theory, and, in addition, are no longer equal and opposite to each other. Such conditions are shown by numerical examples to be quite realistic ones, even for low-viscosity fluids like water. © 1999 Acoustical Society of America. [S0001-4966(99)05612-X]

PACS numbers: 43.35.Ej [MAB]

INTRODUCTION

There are two types of time-averaged radiation forces exerted by a sound wave field on gas bubbles in a fluid. Both of them are of the same nature and discriminated only for reasons of convenience. The force on an individual bubble is called the acoustic radiation pressure or the primary Bjerknes force, and that between bubbles (usually pairwise interactions are considered) is called the secondary or mutual or merely Bjerknes force. It is the secondary force that is the subject matter of this paper.

The Bjerknes forces are named after C. A. Bjerknes and his son V. F. K. Bjerknes, who were the first to report on this effect.¹ These forces are an integral part of many acoustic phenomena and applications, such as acoustic cavitation, multibubble sonoluminescence, acoustic degassing, medical ultrasonics, etc.^{2–7} In the literature, there are a number of detailed reviews of the secondary Bjerknes forces, including the most recent publications.^{8–12} There is little point in copying them here. But works directly relevant to the present investigation will be mentioned below.

The purpose of this paper is to study how the viscosity of the surrounding fluid influences the Bjerknes force of two gas bubbles. The Bjerknes formula of that force is known not to allow for dissipative processes at all, assuming that the surrounding medium is an inviscid incompressible fluid and the gas within the bubbles obeys the adiabatic law. To take the dissipation of energy into account, Zabolotskaya¹³ has used the standard phenomenological technique: The equations of radial pulsations of two interacting bubbles, which had initially been obtained for an ideal fluid, were supplemented with damping terms that allowed for radiation, heat, and viscous losses. That approach, however, takes no account of acoustic streaming around the bubbles as well as viscous effects due to the translational motion of the bubbles in each other's scattered fields. Therefore it is correct only when the viscosity of the surrounding fluid is low. To put it more exactly, that approach is correct only if the viscous penetration depth in the fluid [which is defined as $\delta_v = (2\nu/\omega)^{1/2}$, where ν is the kinematic viscosity and ω is the

angular driving frequency] is much smaller than the equilibrium radii of the bubbles. In many situations of interest that condition is not met. As an example, for a bubble with a radius R_0 of 1 μm in water at a frequency of 20 kHz (parameters typical of bubble sonoluminescence), the ratio δ_v/R_0 is equal to about 4.

The present author has tried to take into account the acoustic streaming, neglecting the translational motion of the bubbles.¹⁴ This, however, gives a wrong result since the translational motion is found to produce extra acoustic streaming of the same order as that due to the radial bubble pulsations alone. Moreover, it gives rise to the vorticity of the linear scattered field, which contributes to the interaction force as well. The present paper takes proper account of all these effects, rectifying thus the oversights of the previous work.

I. THEORY

Consider two gas bubbles with mean radii R_{10} and R_{20} and a distance L between their equilibrium centers, as shown in Fig. 1. It is assumed that the bubbles are far apart so that $L \gg R_{10}, R_{20}$. This allows us to restrict our consideration to the monopole (radial) and dipole (translational) oscillations of the bubbles and to neglect higher-order oscillations (shape deformation modes). Note, however, that if the above condition is violated, then the impact of the shape oscillations on the Bjerknes force can be very significant.¹¹ We will also neglect nonlinear bubble oscillations, assuming the imposed acoustic field to be weak enough. The surrounding medium is presumed to be a viscous incompressible fluid subject to the following equations:

$$\nabla \cdot \mathbf{v} = 0, \quad (1)$$

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v} \cdot \nabla) \mathbf{v} = \eta \Delta \mathbf{v} - \nabla p, \quad (2)$$

where \mathbf{v} is the fluid velocity, ρ is the fluid density, η is the fluid viscosity, and p is the fluid pressure.

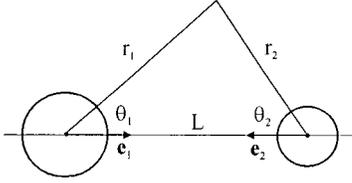


FIG. 1. Geometry of the problem considered.

The radiation interaction force exerted on the j th bubble ($j=1,2$) by a sound wave field is given by¹⁵

$$\mathbf{F}_j = \int_{S_{j0}} \langle 2\eta(\mathbf{n}_j \cdot \nabla)\mathbf{v}^{(2)} + \eta\mathbf{n}_j \times (\nabla \times \mathbf{v}^{(2)}) - \mathbf{n}_j p^{(2)} - \rho\mathbf{v}^{(1)}(\mathbf{v}^{(1)} \cdot \mathbf{n}) \rangle dS_{j0}, \quad (3)$$

where S_{j0} is the equilibrium surface of the j th bubble, \mathbf{n}_j is the unit outward normal to S_{j0} , the superscripts (1) and (2) denote quantities of first and second order in the incident field amplitude, and $\langle \rangle$ means an average over the incident field cycle.

Equation (3) shows that in a viscous fluid the calculation of the force requires the velocity $\langle \mathbf{v}^{(2)} \rangle$ and the pressure $\langle p^{(2)} \rangle$ of acoustic streaming in addition to the linear fluid velocity $\mathbf{v}^{(1)}$. These quantities are calculated in the next two subsections.

A. Linear equations

In this subsection the linear scattered fields of the bubbles are worked out with an accuracy up to leading terms in the small parameters R_{j0}/L .

By linearizing Eqs. (1) and (2), one has

$$\nabla \cdot \mathbf{v}^{(1)} = 0, \quad (4)$$

$$\rho \frac{\partial \mathbf{v}^{(1)}}{\partial t} = \eta \Delta \mathbf{v}^{(1)} - \nabla p^{(1)}. \quad (5)$$

The velocity $\mathbf{v}^{(1)}$ can be represented as

$$\mathbf{v}^{(1)} = \nabla \varphi^{(1)} + \nabla \times \boldsymbol{\psi}^{(1)}, \quad (6)$$

where $\varphi^{(1)}$ and $\boldsymbol{\psi}^{(1)}$ are the scalar and vorticity velocity potentials, respectively. These can be written as

$$\varphi^{(1)} = \varphi_I + \varphi_1^{(1)} + \varphi_2^{(1)}, \quad (7)$$

$$\boldsymbol{\psi}^{(1)} = \boldsymbol{\psi}_1^{(1)} + \boldsymbol{\psi}_2^{(1)}, \quad (8)$$

where $\varphi_I = A \exp(-i\omega t)$ is the incident potential and $\varphi_j^{(1)}$ and $\boldsymbol{\psi}_j^{(1)}$ are the scattered potentials of the j th bubble. Substituting Eq. (6) into Eq. (4), one finds that $\varphi_j^{(1)}$ is to obey the Laplace equation $\Delta \varphi_j^{(1)} = 0$. As $L \gg R_{j0}$, $\varphi_j^{(1)}$ can be taken to be

$$\varphi_j^{(1)} = A e^{-i\omega t} a_j R_{j0} / r_j, \quad (9)$$

where a_j is a dimensionless constant to be determined by the boundary conditions at the surfaces of the bubbles and the meaning of r_j is evident from Fig. 1.

In view of the axial symmetry of the problem involved, $\boldsymbol{\psi}_j^{(1)}$ must be of the form: $\boldsymbol{\psi}_j^{(1)} = \psi_j(r_j, \theta_j, t) \mathbf{e}_e$, where

$\psi_j(r_j, \theta_j, t)$ is some function and \mathbf{e}_e is the unit azimuth vector that is used in both spherical coordinate systems shown in Fig. 1. It follows that $\nabla \cdot \boldsymbol{\psi}_j^{(1)} = 0$ and hence

$$\nabla \times (\nabla \times \boldsymbol{\psi}_j^{(1)}) = \nabla (\nabla \cdot \boldsymbol{\psi}_j^{(1)}) - \Delta \boldsymbol{\psi}_j^{(1)} = -\Delta \boldsymbol{\psi}_j^{(1)}. \quad (10)$$

Substituting Eq. (6) into Eq. (5), subjecting the resulting equation to the curl operator, and taking into account Eq. (10), one obtains

$$(\Delta^2 + \alpha^2 \Delta) \boldsymbol{\psi}_j^{(1)} = 0, \quad (11)$$

where $\alpha^2 = i\omega/\nu$ and $\nu = \eta/\rho$. Let us refine the form of $\boldsymbol{\psi}_j^{(1)}$. The curl of $\boldsymbol{\psi}_j^{(1)}$, as a part of $\mathbf{v}^{(1)}$, must be a polar vector. Therefore $\boldsymbol{\psi}_j^{(1)}$ should be an axial vector. It follows that $\boldsymbol{\psi}_j^{(1)}$ should be equal to the curl of some polar vector. In view of the symmetry of the problem involved, that polar vector should be of the form: $f_j(r_j) \mathbf{e}_j$, where $f_j(r_j)$ is a certain function to be found and \mathbf{e}_j is the unit vector directed from the j th bubble to the other bubble, as shown in Fig. 1. So $\boldsymbol{\psi}_j^{(1)}$ can be represented as

$$\begin{aligned} \boldsymbol{\psi}_j^{(1)} &= A e^{-i\omega t} \nabla \times [f_j(r_j) \mathbf{e}_j] \\ &= A e^{-i\omega t} \nabla f_j(r_j) \times \mathbf{e}_j \\ &= -A e^{-i\omega t} f_j'(r_j) \mathbf{e}_e \sin \theta_j, \end{aligned} \quad (12)$$

with $f_j'(r_j) = df_j(r_j)/dr_j$ and $\mathbf{e}_j = \mathbf{e}_{r_j} \cos \theta_j - \mathbf{e}_{\theta_j} \sin \theta_j$, where \mathbf{e}_{r_j} and \mathbf{e}_{θ_j} are the unit vectors of the spherical coordinates related to the j th bubble.

Substitution of Eq. (12) into Eq. (11) yields

$$\nabla (\Delta^2 + \alpha^2 \Delta) f_j(r_j) \times \mathbf{e}_j = 0, \quad (13)$$

which gives

$$\nabla (\Delta^2 + \alpha^2 \Delta) f_j(r_j) = 0, \quad (14)$$

and then

$$(\Delta + \alpha^2) \Delta f_j(r_j) = \text{const}. \quad (15)$$

Since $\mathbf{v}^{(1)} \rightarrow 0$ for $r_j \rightarrow \infty$, $\Delta f_j(r_j)$ should tend to zero as well, and therefore Eq. (15) takes the form

$$(\Delta + \alpha^2) \Delta f_j(r_j) = 0. \quad (16)$$

This is the Helmholtz equation with respect to $\Delta f_j(r_j)$. Its solution is given by

$$\Delta f_j(r_j) = C \frac{e^{i\alpha r_j}}{r_j}, \quad (17)$$

where C is a constant and α is taken to be $\alpha = (1+i)/\delta_\nu$ [recall that $\delta_\nu = (2\nu/\omega)^{1/2}$] in order that the viscous wave should die out as $r_j \rightarrow \infty$. Integration of Eq. (17) yields

$$f_j'(r_j) = b_j \frac{e^{i\alpha r_j}}{(\alpha r_j)^2} (1 - i\alpha r_j) + c_j \left(\frac{R_{j0}}{r_j} \right)^2, \quad (18)$$

where b_j and c_j are dimensionless constants to be determined by the boundary conditions at the bubbles' surfaces. As is seen from Eq. (12), there is no need to calculate the function $f_j(r_j)$ itself.

We now proceed to find the pressure $p^{(1)}$. It can be represented as

$$p^{(1)} = p_I + p_1^{(1)} + p_2^{(1)}, \quad (19)$$

where

$$p_I = -\rho \frac{\partial \varphi_I}{\partial t} = i\omega\rho\varphi_I \quad (20)$$

and $p_j^{(1)}$ is the pressure produced by the j th bubble. According to Eqs. (5)–(8),

$$\nabla p_j^{(1)} = \eta\Delta(\nabla \times \boldsymbol{\psi}_j^{(1)}) - \rho \frac{\partial}{\partial t} (\nabla \varphi_j^{(1)} + \nabla \times \boldsymbol{\psi}_j^{(1)}). \quad (21)$$

From Eq. (12), one has

$$\nabla \times \boldsymbol{\psi}_j^{(1)} = A e^{-i\omega t} [(\mathbf{e}_j \cdot \nabla) \nabla f_j(r_j) - \mathbf{e}_j \Delta f_j(r_j)]. \quad (22)$$

Substituting Eqs. (9) and (22) into Eq. (21) and taking into account Eq. (16) and the fact that $p_j^{(1)} \rightarrow 0$ for $r_j \rightarrow \infty$, one obtains

$$p_j^{(1)} = A e^{-i\omega t} [i\omega\rho a_j (R_{j0}/r_j) + \eta(\mathbf{e}_j \cdot \nabla)(\Delta + \alpha^2)f_j(r_j)]. \quad (23)$$

It remains to calculate the constants a_j , b_j , and c_j . To this end, the linear boundary conditions for the velocity and the normal stress of the fluid at the surface of the j th bubble are used. The former is given by

$$\dot{R}_j(t)\mathbf{e}_{rj} + U_j(t)\mathbf{e}_j = \mathbf{v}^{(1)}|_{r_j=R_{j0}}, \quad (24)$$

where $R_j(t)$ and $U_j(t)$ are, respectively, the instantaneous radius and the linear translational velocity of the j th bubble, and the dot over $R_j(t)$ denotes the time derivative. Equation (22), after some manipulation, gives

$$\begin{aligned} \nabla \times \boldsymbol{\psi}_j^{(1)} = A e^{-i\omega t} & \left[\left(f_j''(r_j) + \frac{1}{r_j} f_j'(r_j) \right) \sin \theta_j \mathbf{e}_{\theta j} \right. \\ & \left. - \frac{2}{r_j} f_j'(r_j) \cos \theta_j \mathbf{e}_{rj} \right]. \end{aligned} \quad (25)$$

Substitution of this equation along with Eq. (9) into Eq. (6) yields an explicit expression for the total fluid velocity $\mathbf{v}^{(1)}$. That expression is then substituted into Eq. (24), keeping up to the second-order terms in R_{j0}/L ; in doing so, a_j is assumed to be of zero order and b_j and c_j of second order. (It will be seen below that this is the case.) As a result, Eq. (24) yields

$$a_j A e^{-i\omega t} = -R_{j0} \dot{R}_j(t), \quad (26)$$

$$A e^{-i\omega t} \left[\frac{R_{3-j0}}{L^2} a_{3-j} - \frac{2}{R_{j0}} f_j'(R_{j0}) \right] = U_j(t), \quad (27)$$

$$A e^{-i\omega t} \left[f_j''(R_{j0}) + \frac{1}{R_{j0}} f_j'(R_{j0}) - \frac{R_{3-j0}}{L^2} a_{3-j} \right] = -U_j(t). \quad (28)$$

Combining Eqs. (27) and (28), one obtains the following equation for the function $f_j'(r_j)$ at $r_j = R_{j0}$:

$$f_j''(R_{j0}) - f_j'(R_{j0})/R_{j0} = 0. \quad (29)$$

Differentiating Eq. (18), one finds an expression for $f_j''(r_j)$,

$$f_j''(r_j) = b_j \frac{e^{i\alpha r_j}}{\alpha^2 r_j^3} (\alpha^2 r_j^2 + 2i\alpha r_j - 2) - c_j \frac{2R_{j0}^2}{r_j^3}, \quad (30)$$

Substituting Eqs. (30) and (18) into Eq. (29), one expresses one of the sought-for constants, c_j , in terms of the other one, b_j ,

$$c_j = b_j \frac{e^{i\alpha R_{j0}}}{3\alpha^2 R_{j0}^2} (\alpha^2 R_{j0}^2 + 3i\alpha R_{j0} - 3). \quad (31)$$

To go on, we need an expression for $U_j(t)$. To find it, we use the equation of the translational motion of the j th bubble,

$$\frac{4}{3} \pi R_{j0}^3 \rho_j \frac{dU_j(t)}{dt} = \int_{S_{j0}} (\sigma_{rr}^{(1)} \cos \theta_j - \sigma_{r\theta}^{(1)} \sin \theta_j) dS_{j0}, \quad (32)$$

where ρ_j is the equilibrium density of the gas inside the j th bubble and $\sigma_{rr}^{(1)}$ and $\sigma_{r\theta}^{(1)}$ are the components of the linear stress tensor of the surrounding fluid in the spherical coordinates related to the j th bubble, given by¹⁶

$$\sigma_{rr}^{(1)} = -p^{(1)} + 2\eta \frac{\partial \mathbf{v}_{rj}^{(1)}}{\partial r_j}, \quad (33)$$

$$\sigma_{r\theta}^{(1)} = \eta \left(\frac{1}{r_j} \frac{\partial \mathbf{v}_{rj}^{(1)}}{\partial \theta_j} + \frac{\partial \mathbf{v}_{\theta j}^{(1)}}{r_j} - \frac{\mathbf{v}_{\theta j}^{(1)}}{r_j} \right). \quad (34)$$

Calculating $\mathbf{v}^{(1)}$ from Eqs. (6)–(9) and (25), $p^{(1)}$ from Eqs. (19), (20), and (23), substituting the resulting equations into Eqs. (33) and (34), taking into account Eq. (29), and keeping up to the second-order terms in R_{j0}/L , one finds

$$\begin{aligned} \sigma_{rr}^{(1)}|_{r_j=R_{j0}} = -i\omega\rho A e^{-i\omega t} & \{ 1 + a_{3-j}(R_{3-j0}/L) [1 \\ & + (R_{j0}/L) \cos \theta_j] + a_j [1 - 4(\alpha R_{j0})^{-2}] \\ & + \cos \theta_j [f_j'(R_{j0}) + \alpha^{-2} f_j'''(R_{j0})] \}, \end{aligned} \quad (35)$$

$$\sigma_{r\theta}^{(1)}|_{r_j=R_{j0}} = \eta A e^{-i\omega t} \sin \theta_j f_j'''(R_{j0}). \quad (36)$$

Substitution of these two equations into Eq. (32) results in

$$\begin{aligned} U_j(t) = \frac{\rho A e^{-i\omega t}}{\rho_j R_{j0}} & [f_j'(R_{j0}) + 3\alpha^{-2} f_j'''(R_{j0}) \\ & + a_{3-j} R_{10} R_{20} / L^2]. \end{aligned} \quad (37)$$

Differentiating Eq. (30), one obtains the following expression for $f_j'''(r_j)$:

$$f_j'''(r_j) = b_j \frac{e^{i\alpha r_j}}{\alpha^2 r_j^4} (i\alpha^3 r_j^3 - 3\alpha^2 r_j^2 - 6i\alpha r_j + 6) + c_j \frac{6R_{j0}^2}{r_j^4}. \quad (38)$$

Setting in Eqs. (18) and (38) $r_j = R_{j0}$ and taking into account Eq. (31), one finds

$$f_j'(R_{j0}) = b_j e^{i\alpha R_{j0}} / 3, \quad (39)$$

$$f_j'''(R_{j0}) = b_j e^{i\alpha R_{j0}} (i\alpha R_{j0} - 1) / R_{j0}^2. \quad (40)$$

Substitution of Eqs. (39) and (40) into Eq. (37) gives the linear translational velocity of the j th bubble, $U_j(t)$, in terms of the constant b_j ,

$$U_j(t) = Ab_j e^{i(\alpha R_{j0} - \omega t)} \frac{\alpha^2 R_{j0}^2 + 3i\alpha R_{j0} - 3}{\alpha^2 R_{j0}^3 (\lambda_j - 1)}, \quad (41)$$

with $\lambda_j = \rho_j / \rho$. Substituting Eqs. (41) and (39) into Eq. (27), one obtains b_j to be

$$b_j = \frac{3(\alpha R_{j0})^2 (\lambda_j - 1) \exp(-i\alpha R_{j0})}{9(i\alpha R_{j0} - 1) + (\alpha R_{j0})^2 (1 + 2\lambda_j)} \frac{R_{10} R_{20}}{L^2} a_{3-j}. \quad (42)$$

Finally, substitution of Eq. (42) into Eqs. (31) and (41) yields

$$c_j = \frac{(\lambda_j - 1)(\alpha^2 R_{j0}^2 + 3i\alpha R_{j0} - 3)}{9(i\alpha R_{j0} - 1) + (\alpha R_{j0})^2 (1 + 2\lambda_j)} \frac{R_{10} R_{20}}{L^2} a_{3-j}, \quad (43)$$

$$U_j(t) = \frac{3}{R_{j0}} A e^{-i\omega t} \frac{\alpha^2 R_{j0}^2 + 3i\alpha R_{j0} - 3}{9(i\alpha R_{j0} - 1) + (\alpha R_{j0})^2 (1 + 2\lambda_j)} \times \frac{R_{10} R_{20}}{L^2} a_{3-j}. \quad (44)$$

To find a_j , we use the boundary condition of continuity of the normal stress across the bubble surface,

$$P_0 - \sigma_{rr}^{(1)}|_{r_j=R_{j0}} + \frac{2\sigma}{R_j(t)} = P_{gj} \left(\frac{R_{j0}}{R_j(t)} \right)^{3\gamma}, \quad (45)$$

where P_0 is the hydrostatic pressure in the fluid, σ is the surface tension, P_{gj} is the equilibrium pressure of the gas inside the j th bubble, and γ is the polytropic exponent. The time-varying radius of the j th bubble can be written as

$$R_j(t) = R_{j0} + x_j^{(1)}(t), \quad (46)$$

where, as follows from Eq. (26), the linear change $x_j^{(1)}(t)$ is given by

$$x_j^{(1)}(t) = -iA e^{-i\omega t} a_j / (\omega R_{j0}). \quad (47)$$

Substituting Eqs. (46) and (47) along with Eq. (35) into Eq. (45), with an accuracy up to the first-order terms in R_{j0}/L , one has

$$\left(\frac{\omega_j^2}{\omega^2} - 1 - i\delta_j \right) a_j - \frac{R_{3-j0}}{L} a_{3-j} = 1, \quad (48)$$

where ω_j is the monopole resonance frequency of the j th bubble, given by

$$\omega_j = \frac{1}{R_{j0}} \left[\frac{1}{\rho} \left(3\gamma P_{gj} - \frac{2\sigma}{R_{j0}} \right) \right]^{1/2}, \quad (49)$$

and δ_j is the dimensionless damping constant due to the fluid viscosity, defined as

$$\delta_j = \frac{4\nu}{\omega R_{j0}^2} = \frac{2\delta_v^2}{R_{j0}^2}. \quad (50)$$

Solving Eq. (48) for a_j , one obtains

$$a_j = \frac{1}{D} \left(\frac{\omega_{3-j}^2}{\omega^2} - 1 - i\delta_{3-j} + \frac{R_{3-j0}}{L} \right) \quad (51)$$

with

$$D = \left(\frac{\omega_1^2}{\omega^2} - 1 - i\delta_1 \right) \left(\frac{\omega_2^2}{\omega^2} - 1 - i\delta_2 \right) - \frac{R_{10} R_{20}}{L^2}. \quad (52)$$

So, the linear equations have been solved. We can now proceed to the equations of acoustic streaming.

B. Equations of acoustic streaming

By averaging Eqs. (1) and (2) over time and keeping up to the second order in the incident field amplitude, one obtains

$$\nabla \cdot \langle \mathbf{v}^{(2)} \rangle = 0, \quad (53)$$

$$\eta \Delta \langle \mathbf{v}^{(2)} \rangle - \nabla \langle p^{(2)} \rangle = \rho \langle (\mathbf{v}^{(1)} \cdot \nabla) \mathbf{v}^{(1)} \rangle. \quad (54)$$

These equations are normally referred to as the equations of acoustic streaming.

The velocity $\langle \mathbf{v}^{(2)} \rangle$ is sought as

$$\langle \mathbf{v}^{(2)} \rangle = \nabla \times \Psi. \quad (55)$$

Then Eq. (53) is satisfied automatically, and Eq. (54), upon substitution of Eq. (55) and applying the curl operator to both sides of the resulting equation, gives

$$\Delta^2 \Psi = -\nu^{-1} \nabla \times \langle (\mathbf{v}^{(1)} \cdot \nabla) \mathbf{v}^{(1)} \rangle. \quad (56)$$

Taking into account that

$$\begin{aligned} \nabla \times [(\mathbf{v}^{(1)} \cdot \nabla) \mathbf{v}^{(1)}] &= \nabla \times \left[\frac{1}{2} \nabla (\mathbf{v}^{(1)})^2 + (\nabla \times \mathbf{v}^{(1)}) \times \mathbf{v}^{(1)} \right] \\ &= \nabla \times (\mathbf{v}^{(1)} \times \Delta \psi^{(1)}), \end{aligned} \quad (57)$$

one has instead of Eq. (56),

$$\Delta^2 \Psi = \nu^{-1} \nabla \times \langle \Delta \psi^{(1)} \times (\nabla \varphi^{(1)} + \nabla \times \psi^{(1)}) \rangle. \quad (58)$$

It follows from Eq. (58) that the leading term of Ψ in the vicinity of the surface of the j th bubble, denoted below by Ψ_j , is of second order in R_{j0}/L and adheres to the equation

$$\Delta^2 \Psi_j = \nu^{-1} \nabla \times \langle \Delta \psi_j^{(1)} \times \nabla \varphi_j^{(1)} \rangle. \quad (59)$$

Knowledge of Ψ_j will suffice to calculate the force on the j th bubble.

Using Eqs. (9) and (12), one finds

$$\begin{aligned} \nabla \times \langle \Delta \psi^{(1)} \times \nabla \varphi_j^{(1)} \rangle &= \frac{|A|^2 R_{j0} \sin \theta_j \mathbf{e}_e}{2r_j^5} \\ &\times \text{Re} \{ a_j^* [r_j^3 f_j^{IV}(r_j) + r_j^2 f_j'''(r_j) \\ &- 6r_j f_j''(r_j) + 6f_j'(r_j)] \}, \end{aligned} \quad (60)$$

where Re means ‘‘the real part of’’ and the asterisk indicates the complex conjugate. Substituting Eq. (60) along with Eqs. (18), (30), and (38) into Eq. (59), one obtains

$$\begin{aligned} \Delta^2 \Psi_j &= \frac{|A|^2 R_{j0} \sin \theta_j \mathbf{e}_e}{2\nu r_j^5} \\ &\times \text{Re} \{ a_j^* b_j e^{i\alpha r_j} (3 - 3i\alpha r_j - \alpha^2 r_j^2) \}. \end{aligned} \quad (61)$$

It follows that Ψ_j can be represented as

$$\Psi_j = \frac{|A|^2 R_{j0}}{2\nu} \sin \theta_j \mathbf{e}_e \text{Re} \{ a_j^* b_j g_j(r_j) \}, \quad (62)$$

where $g_j(r_j)$ is some function to be found. Substitution of Eq. (62) into Eq. (61) yields

$$g_j^{IV}(r_j) + 4r_j^{-1}g_j'''(r_j) - 4r_j^{-2}g_j''(r_j) = r_j^{-5}e^{i\alpha r_j}(3 - 3i\alpha r_j - \alpha^2 r_j^2). \quad (63)$$

This equation is solved by using the standard Lagrange method. That is, we first consider the associated homogeneous equation,

$$y^{IV}(r_j) + 4r_j^{-1}y'''(r_j) - 4r_j^{-2}y''(r_j) = 0. \quad (64)$$

Its solutions are of the form r_j^μ , where μ is defined from

$$\mu(\mu - 1)[(\mu - 2)(\mu - 3) + 4(\mu - 2) - 4] = 0. \quad (65)$$

The roots of Eq. (65) are equal to $-2, 0, 1,$ and 3 . As prescribed by the Lagrange method, the solution to Eq. (63) is given by

$$g_j(r_j) = C_1(r_j)r_j^{-2} + C_2(r_j) + C_3(r_j)r_j + C_4(r_j)r_j^3, \quad (66)$$

where the functions $C_n(r_j)$ ($n=1,2,3,4$) are determined by the following set of differential equations:

$$\begin{aligned} C_1' + r_j^2 C_2' + r_j^3 C_3' + r_j^5 C_4' &= 0, \\ 2C_1' - r_j^3 C_3' - 3r_j^5 C_4' &= 0, \\ C_1' + r_j^5 C_4' &= 0, \\ 24C_1' - 6r_j^5 C_4' &= e^{i\alpha r_j}(\alpha^2 r_j^2 + 3i\alpha r_j - 3). \end{aligned} \quad (67)$$

Solving these equations yields

$$C_1(r_j) = D_1 - \frac{e^{i\alpha r_j}}{30i\alpha} (8 - 5i\alpha r_j - \alpha^2 r_j^2), \quad (68)$$

$$C_2(r_j) = D_2 + \frac{i\alpha}{6} e^{i\alpha r_j} + \frac{1}{2} \int_{R_{j0}}^{r_j} r^{-2} e^{i\alpha r} (1 - i\alpha r) dr, \quad (69)$$

$$C_3(r_j) = D_3 - \frac{1}{6} \int_{R_{j0}}^{r_j} r^{-3} e^{i\alpha r} (3 - 3i\alpha r - \alpha^2 r^2) dr, \quad (70)$$

$$C_4(r_j) = D_4 + \frac{1}{30} \int_{R_{j0}}^{r_j} r^{-5} e^{i\alpha r} (3 - 3i\alpha r - \alpha^2 r^2) dr. \quad (71)$$

The constants D_n ($n=1,2,3,4$) are determined from the boundary conditions at infinity and at the surface of the j th bubble. From the condition that $\nabla \times \Psi_j \rightarrow 0$ as $r_j \rightarrow \infty$, one has

$$\begin{aligned} D_3 &= \frac{1}{6} \int_{R_{j0}}^{\infty} r^{-3} e^{i\alpha r} (3 - 3i\alpha r - \alpha^2 r^2) dr \\ &= \frac{1}{2R_{j0}^2} \left[E_3(-iz_j) - iz_j E_2(-iz_j) - \frac{1}{3} z_j^2 E_1(-iz_j) \right], \end{aligned} \quad (72)$$

$$\begin{aligned} D_4 &= -\frac{1}{30} \int_{R_{j0}}^{\infty} r^{-5} e^{i\alpha r} (3 - 3i\alpha r - \alpha^2 r^2) dr \\ &= \frac{1}{10R_{j0}^4} \left[\frac{1}{3} z_j^2 E_3(-iz_j) + iz_j E_4(-iz_j) - E_5(-iz_j) \right], \end{aligned} \quad (73)$$

where $z_j = \alpha R_{j0}$ and $E_m(z)$ is the integral exponent of order n , defined as¹⁷

$$E_m(z) = \int_1^{\infty} x^{-m} e^{-zx} dx, \quad m=0,1,2,\dots; \quad \text{Re } z > 0. \quad (74)$$

To find D_1 and D_2 , we use the boundary condition of adhesion of the fluid particles to the surface of the j th bubble. Note that the condition is applied to the perturbed surface $S_j(t)$ given by

$$\mathbf{r}_j = \xi_j(t) \mathbf{e}_j + [R_{j0} + x_j^{(1)}(t)] \mathbf{n}_j, \quad (75)$$

where the displacement of the center of the j th bubble $\xi_j(t)$ is subject to the relation: $\dot{\xi}_j(t) = U_j(t)$. Expansion of the exact fluid velocity \mathbf{v} at this surface in a Taylor series up to the second order in the incident field amplitude, followed by averaging over time, results in

$$\langle \mathbf{v} |_{S_j(t)} \rangle \approx \langle \mathbf{v}^{(2)} \rangle |_{r_j=R_{j0}} + \langle [(\xi_j \mathbf{e}_j + x_j^{(1)} \mathbf{e}_{rj}) \cdot \nabla] \mathbf{v}^{(1)} \rangle |_{r_j=R_{j0}}. \quad (76)$$

The bubbles can be assumed to be at rest in the average: Their ‘‘slow’’ motions are caused by the very forces we are searching for and hence can be neglected when the forces are, as such, calculated. As a result, the desired boundary condition, with an accuracy up to the second-order terms in R_{j0}/L , takes the form

$$\begin{aligned} (\nabla \times \Psi_j)_{r_j=R_{j0}} + \left\langle \xi_j (\mathbf{e}_j \cdot \nabla) \nabla \varphi_j^{(1)} + x_j^{(1)} \frac{\partial}{\partial r_j} (\nabla \varphi_j^{(1)} + \nabla \times \psi_j^{(1)}) \right\rangle \Big|_{r_j=R_{j0}} &= 0. \end{aligned} \quad (77)$$

From Eq. (62), one finds

$$\begin{aligned} \nabla \times \Psi_j &= \frac{|A|^2 R_{j0}}{2\nu r_j} \text{Re} \{ a_j^* b_j [2g_j(r_j) \cos \theta_j \mathbf{e}_{rj} - (g_j(r_j) + r_j g_j'(r_j)) \sin \theta_j \mathbf{e}_{\theta j}] \}. \end{aligned} \quad (78)$$

Substituting Eq. (78) along with Eqs. (9) and (25) into Eq. (77), and using Eqs. (29), (40), (41), (47), (66), and (68)–(71), one obtains

$$\begin{aligned} D_1 &= \frac{R_{j0}}{30} \left\{ 15R_{j0}^2 D_3 + 45R_{j0}^4 D_4 + \frac{e^{iz_j}}{z_j^4 (\lambda_j - 1)} [30(z_j^2 + 3iz_j - 3) + (1 - \lambda_j) z_j^2 (15 - 7iz_j + 5z_j^2 - iz_j^3)] \right\}, \end{aligned} \quad (79)$$

$$D_2 = \frac{1}{2R_{j0}} \left[\frac{e^{iz_j}}{3z_j^2} (3 - 3iz_j - iz_j^3) - 3R_{j0}^2 D_3 - 5R_{j0}^4 D_4 \right]. \quad (80)$$

It only remains to find the acoustic streaming pressure $\langle p^{(2)} \rangle$. It is sufficient to calculate the leading term of $\langle p^{(2)} \rangle$ in the vicinity of the j th bubble. According to Eq. (54), that term, denoted below by $\langle p_j^{(2)} \rangle$, is given by

$$\begin{aligned}
\nabla\langle p_j^{(2)}\rangle &= \eta\Delta(\nabla\times\Psi_j) - \rho\langle(\mathbf{v}^{(1)}\cdot\nabla)\mathbf{v}^{(1)}\rangle \\
&= \eta\Delta(\nabla\times\Psi_j) - \rho\langle\frac{1}{2}\nabla\langle(\mathbf{v}^{(1)})^2\rangle + (\nabla\times\mathbf{v}^{(1)})\times\mathbf{v}^{(1)}\rangle \\
&= \eta\Delta(\nabla\times\Psi_j) - \frac{\rho}{2}\nabla\langle(\mathbf{v}^{(1)})^2\rangle \\
&\quad - \rho\langle\nabla\varphi_j^{(1)}\times\Delta\psi_j^{(1)}\rangle, \tag{81}
\end{aligned}$$

where terms are kept accurate up to R_{j0}^2/L^2 .

For subsequent calculations it is convenient to represent Ψ_j as

$$\Psi_j = \nabla\times[\chi(r_j)\mathbf{e}_j] = \nabla\chi(r_j)\times\mathbf{e}_j = -\chi'(r_j)\sin\theta_j\mathbf{e}_\epsilon, \tag{82}$$

where, as is seen from comparison with Eq. (62),

$$\chi'(r_j) = \frac{|A|^2 R_{j0}}{2\nu} \operatorname{Re}\{a_j^* b_j g_j(r_j)\}. \tag{83}$$

From Eq. (82), one finds

$$\nabla\times\Psi_j = (\mathbf{e}_j\cdot\nabla)\nabla\chi(r_j) - \mathbf{e}_j\Delta\chi(r_j), \tag{84}$$

which, upon substitution in Eq. (81), gives

$$\begin{aligned}
\nabla\langle p_j^{(2)}\rangle &= \eta\nabla(\mathbf{e}_j\cdot\nabla)\Delta\chi(r_j) - \frac{\rho}{2}\nabla\langle(\mathbf{v}^{(1)})^2\rangle \\
&\quad - \eta\mathbf{e}_j\Delta^2\chi(r_j) + \rho\langle\Delta\psi_j^{(1)}\times\nabla\varphi_j^{(1)}\rangle. \tag{85}
\end{aligned}$$

Let us transform the last two terms on the right-hand side of Eq. (85), denoting their sum by \mathbf{T} . Using Eqs. (66) and (83), one finds

$$\Delta^2\chi(r_j) = -15\nu^{-1}|A|^2 R_{j0} \operatorname{Re}\{a_j^* b_j C_4(r_j)\}. \tag{86}$$

Substituting Eq. (86) along with Eqs. (9) and (12) into \mathbf{T} , and using Eqs. (18), (30), (38), and the fact that $\mathbf{e}_j = \cos\theta_j\mathbf{e}_{rj} - \sin\theta_j\mathbf{e}_{\theta j}$, one obtains

$$\begin{aligned}
\mathbf{T} &= 15|A|^2 R_{j0}\rho \operatorname{Re}\left\{a_j^* b_j \left[C_4(r_j)\cos\theta_j\mathbf{e}_{rj} \right. \right. \\
&\quad \left. \left. + \left(\frac{e^{i\alpha r}}{30r_j^4}(i\alpha r_j - 1) - C_4(r_j)\right)\sin\theta_j\mathbf{e}_{\theta j}\right]\right\}. \tag{87}
\end{aligned}$$

It is easy to verify that

$$\mathbf{T} = \nabla[Q(r_j)\cos\theta_j], \tag{88}$$

where

$$\begin{aligned}
Q(r_j) &= 15|A|^2 R_{j0}\rho \operatorname{Re}\left\{a_j^* b_j \left[r_j C_4(r_j) \right. \right. \\
&\quad \left. \left. + \frac{e^{i\alpha r_j}}{30r_j^3}(1 - i\alpha r_j)\right]\right\}. \tag{89}
\end{aligned}$$

Substituting Eq. (88) into Eq. (85), one gets

$$\langle p_j^{(2)}\rangle = \eta(\mathbf{e}_j\cdot\nabla)\Delta\chi(r_j) - \frac{\rho}{2}\langle(\mathbf{v}^{(1)})^2\rangle + Q(r_j)\cos\theta_j, \tag{90}$$

or, after some manipulation,

$$\begin{aligned}
\langle p_j^{(2)}\rangle &= \frac{1}{2}\rho|A|^2 R_{j0} \cos\theta_j \operatorname{Re}\{a_j^* b_j [20r_j C_4(r_j) \\
&\quad + 2r_j^{-2} C_2(r_j) + r_j^{-3} e^{i\alpha r_j} (1 - i\alpha r_j)]\} \\
&\quad - \frac{\rho}{4} \operatorname{Re}\{\nabla\varphi_j^{(1)*}(\nabla\varphi_j^{(1)} + 2\nabla\varphi_{3-j}^{(1)} \\
&\quad + 2\nabla\times\psi_j^{(1)})\}. \tag{91}
\end{aligned}$$

We are now fully prepared to arrive at the interaction force.

C. Calculation of the interaction force

With an accuracy up to the leading terms in R_{j0}/L , Eq. (3) is reduced to

$$\begin{aligned}
\mathbf{F}_j &= R_{j0}^2 \int_0^{2\pi} d\epsilon \int_0^\pi \sin\theta_j d\theta_j \left[2\eta \frac{\partial}{\partial r_j} (\nabla\times\Psi_j) - \eta\mathbf{e}_{rj} \right. \\
&\quad \left. \times \Delta\Psi_j - \mathbf{e}_{rj}\langle p_j^{(2)}\rangle - \rho\langle\mathbf{v}^{(1)}(\mathbf{e}_{rj}\cdot\nabla)\rangle \right]_{r_j=R_{j0}}. \tag{92}
\end{aligned}$$

Substituting all the necessary quantities from the two preceding subsections, after cumbersome but straightforward calculations, one obtains

$$\mathbf{F}_j = \frac{2\pi\rho|A|^2 R_{10} R_{20}}{L^2} \mathbf{e}_j \operatorname{Re}\{a_j^* a_{3-j} [1 + \tau(z_j, \lambda_j)]\}, \tag{93}$$

where the function $\tau(z, \lambda)$ is given by

$$\begin{aligned}
\tau(z, \lambda) &= \frac{1 - \lambda}{2[9(iz - 1) + z^2(1 + 2\lambda)]} \{6 - 6iz + 6z^2 \\
&\quad - 2iz^3 + z^2 e^{-iz} [3z^2 E_1(-iz) + 9iz E_2(-iz) \\
&\quad - (9 + z^2) E_3(-iz) - 3iz E_4(-iz) \\
&\quad + 3E_5(-iz)]\}. \tag{94}
\end{aligned}$$

Substitution of Eq. (51) into Eq. (93), and rearrangement of Eq. (94) by using recurrence formulas for $E_m(-iz)$,¹⁷ finally yield

$$\begin{aligned}
\mathbf{F}_j &= \frac{2\pi\rho|A|^2 R_{10} R_{20}}{|D|^2 L^2} \mathbf{e}_j [(T_1 T_2 + \delta_1 \delta_2) \\
&\quad \times (1 + \operatorname{Re}\{\tau(z_j, \lambda_j)\}) + \operatorname{Im}\{\tau(z_j, \lambda_j)\} \\
&\quad \times (\delta_j T_{3-j} - \delta_{3-j} T_j)], \tag{95}
\end{aligned}$$

$$\begin{aligned}
\tau(z, \lambda) &= \frac{1 - \lambda}{16[9(iz - 1) + z^2(1 + 2\lambda)]} [48 - 48iz + 18z^2 \\
&\quad + 14iz^3 - z^4 - iz^5 + (z^6 - 12z^4) e^{-iz} E_1(-iz)], \tag{96}
\end{aligned}$$

where

$$T_j = \frac{\omega_j^2}{\omega^2} - 1 + \frac{R_{j0}}{L} \tag{97}$$

and Im means ‘‘the imaginary part of.’’

Since $\lambda_j \ll 1$, τ is governed chiefly by the first argument, z , which is specified by the parameter $X = R_{j0}/\delta_\nu$. The real and imaginary parts of τ versus X for $\lambda = 0$ are plotted in Fig.

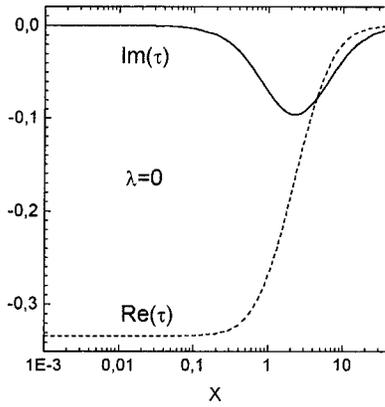


FIG. 2. The real and imaginary parts of τ as functions of the dimensionless parameter $X=R_{j0}/\delta_v$. Large values of X correspond to low viscosity and small ones to high viscosity.

2. Large values of X correspond to the limit of weak viscosity ($R_{j0} \gg \delta_v$) and small ones to the limit of strong viscosity ($R_{j0} \ll \delta_v$). For $R_{j0} \gg \delta_v$, τ is approximated by

$$\tau(z_j, 0) \approx 15z_j^{-2} = -7.5i(\delta_v/R_{j0})^2 = -3.75i\delta_j, \quad (98)$$

and in the reverse limiting case

$$\begin{aligned} \tau(z_j, 0) &\approx -\frac{1}{3} \left(1 + \frac{35}{72} z_j^2 \right) = -\frac{1}{3} \left[1 + \frac{35i}{36} \left(\frac{R_{j0}}{\delta_v} \right)^2 \right] \\ &= -\frac{1}{3} \left(1 + \frac{35i}{18\delta_j} \right). \end{aligned} \quad (99)$$

Figure 2 shows that both the real and the imaginary parts of τ are always negative. The modulus of the real part reaches maximum in the limit of high viscosity, while that of the imaginary part has its peak at medium values of X , in the neighborhood of $X=2.2$.

II. DISCUSSION

For $\tau=0$, Eq. (95) is reduced to the result of Zabolotskaya,¹³ which is mentioned in the Introduction,

$$\mathbf{F}_{jZ} = \frac{2\pi\rho|A|^2 R_{10} R_{20}}{|D|^2 L^2} (T_1 T_2 + \delta_1 \delta_2) \mathbf{e}_j. \quad (100)$$

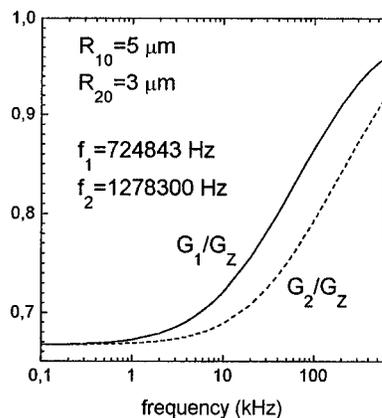


FIG. 3. The ratios G_1/G_Z and G_2/G_Z as functions of the driving frequency for two air bubbles in water in the range below the resonance frequencies f_1 and f_2 of the bubbles.

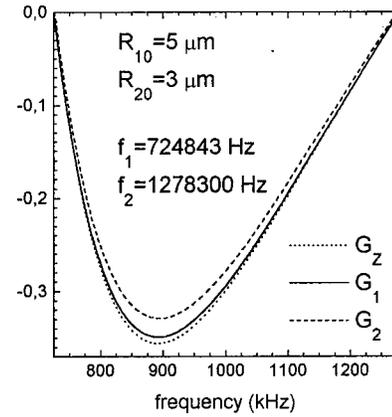


FIG. 4. G_Z , G_1 , and G_2 versus the driving frequency in the range between the resonance frequencies of the bubbles.

Comparison of Eqs. (100) and (95) reveals that the viscous effects on the Bjerknes force go beyond the term $\delta_1 \delta_2$. In the first place, the real part of τ corrects the term

$$G_Z = T_1 T_2 + \delta_1 \delta_2. \quad (101)$$

Second, the imaginary part of τ gives rise to one more term. As a result of both refinements, the force on one bubble is no longer equal and opposite to that on the other bubble. In other words, the fluid viscosity, as one would expect from a dissipative process, breaks the conservatism of the system involved. This feature is missing from Eq. (100), which is indicative of its inaccuracy.

To illustrate these inferences, some numerical calculations for two air bubbles in water have been made. The appropriate physical parameters were as follows: $\rho=1000 \text{ kg m}^{-3}$, $P_0=1.013 \times 10^5 \text{ Pa}$, $\eta=10^{-3} \text{ Pa s}$, $\sigma=0.0727 \text{ N m}^{-1}$, $\gamma=1.4$. The gas pressure P_{gj} and the gas density ρ_j within the j th bubble were calculated from $P_{gj}=P_0 + 2\sigma/R_{j0}$ and $\rho_j=\rho_{g0}P_{gj}/P_0$, where $\rho_{g0}=1.2 \text{ kg m}^{-3}$ is the density of air at the pressure P_0 . To cover either limit ($R_{j0} \gg \delta_v$ and $R_{j0} \ll \delta_v$), the radii of the bubbles were chosen to be $R_{10}=5$ and $R_{20}=3 \text{ }\mu\text{m}$. The results of the calculations are presented in Figs. 3–5.

Figure 3 depicts the ratios G_1/G_Z and G_2/G_Z against the driving frequency. The quantity G_Z is calculated from Eq. (101) and G_j from

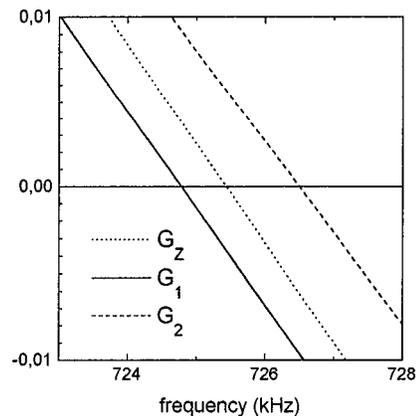


FIG. 5. The magnified region of Fig. 4 in the neighborhood of the resonance frequency of the bigger bubble (bubble 1).

$$G_j = (T_1 T_2 + \delta_1 \delta_2) [1 + \operatorname{Re}\{\tau(z_j, \lambda_j)\}] + \operatorname{Im}\{\tau(z_j, \lambda_j)\} \\ \times (\delta_j T_{3-j} - \delta_{3-j} T_j), \quad (102)$$

the small term R_{j0}/L in T_j being neglected. So, the ratio G_j/G_Z gives an idea of the discrepancy between Eqs. (95) and (100). Figure 3 refers to the frequency range below the resonance frequencies of the two bubbles, which are denoted by f_1 and f_2 in it. In this range the bubbles undergo mutual attraction. It is seen that in the general case the forces on the bubbles differ noticeably from what is given by Eq. (100) as well as each other, the smaller bubble (bubble 2) being subject to a smaller force. For low frequencies, where $R_{j0} \ll \delta_v$, both G_1 and G_2 tend to $2G_Z/3$, in compliance with Eq. (99).

Figure 4 shows G_Z , G_1 , and G_2 in the frequency range between the resonance frequencies of the bubbles, where the mutual repulsion occurs. It is seen that in this range as well there are, even if minor, departures from G_Z , especially for the smaller bubble. Their insignificance is explained by the fact that for such driving frequencies $R_{j0} \gg \delta_v$ and therefore the fluid viscosity manifests itself only slightly. Finally, Fig. 5 is a magnified region of Fig. 4 in the immediate vicinity of the resonance frequency of the bigger bubble, bubble 1. It brings out that there is a frequency range where the forces on the bubbles are of opposite sign. In particular, near f_1 , bubble 1 experiences a repulsive force while bubble 2 is still subject to an attractive one.

III. CONCLUSIONS

The aim of this paper has been to examine viscous effects on the secondary Bjerknes force of two spherical gas bubbles in a sound wave field. It has been assumed that the surrounding medium is a viscous incompressible fluid and the spacing between the bubbles is much larger than their radii. The Bjerknes force has been calculated taking into account acoustic streaming generated by both the radial and the translational oscillations of the bubbles. In this way, errors of the author's previous attempt,¹⁴ which neglected the translational motions, have been rectified. It has been found that in

the limit of high viscosity, the secondary Bjerknes forces on each bubble, first, differ noticeably from values given by the former theory¹³ ignoring acoustic streaming, and second, are not equal and opposite to each other, as the fluid viscosity breaks the conservatism of the system of two interacting bubbles.

ACKNOWLEDGMENT

This work was supported by the European Commission under the Inco-Copernicus program (Contract No. IC15-CT98-0141).

- ¹V. F. K. Bjerknes, *Fields of Force* (Columbia University Press, New York, 1906).
- ²E. A. Neppiras, "Acoustic cavitation," *Phys. Rep.* **61**, 159–251 (1980).
- ³A. J. Walton and G. T. Reynolds, "Sonoluminescence," *Adv. Phys.* **33**, 595–660 (1984).
- ⁴*Physical Principles of Ultrasonic Technology*, edited by L. D. Rozenberg (Plenum, New York, 1973).
- ⁵*Physical Principles of Medical Ultrasonics*, edited by C. R. Hill (Harcourt, London, 1986).
- ⁶T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1994).
- ⁷W. Lauterborn and C. D. Ohl, "Cavitation bubble dynamics," *Ultrasonics Sonochem.* **4**, 65–75 (1997).
- ⁸L. A. Crum, "Bjerknes forces on bubbles in a stationary sound field," *J. Acoust. Soc. Am.* **57**, 1363–1370 (1975).
- ⁹N. A. Pelekasis and J. A. Tsamopoulos, "Bjerknes forces between two bubbles. Part 1. Response to a step change in pressure," *J. Fluid Mech.* **254**, 467–499 (1993).
- ¹⁰N. A. Pelekasis and J. A. Tsamopoulos, "Bjerknes forces between two bubbles. Part 2. Response to an oscillatory pressure field," *J. Fluid Mech.* **254**, 501–527 (1993).
- ¹¹A. A. Doinikov and S. T. Zavtrak, "On the mutual interaction of two gas bubbles in a sound field," *Phys. Fluids* **7**, 1923–1930 (1995).
- ¹²A. A. Doinikov and S. T. Zavtrak, "Radiation forces between two bubbles in a compressible liquid," *J. Acoust. Soc. Am.* **102**, 1424–1431 (1997).
- ¹³E. A. Zabolotskaya, "Interaction of gas bubbles in a sound field," *Sov. Phys. Acoust.* **30**, 365–368 (1984).
- ¹⁴A. A. Doinikov, "Dissipative effects on Bjerknes forces between two bubbles," *J. Acoust. Soc. Am.* **102**, 747–751 (1997).
- ¹⁵A. A. Doinikov, "Acoustic radiation pressure on a compressible sphere in a viscous fluid," *J. Fluid Mech.* **267**, 1–21 (1994).
- ¹⁶L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, New York, 1959).
- ¹⁷M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

An acoustical helicoidal wave transducer with applications for the alignment of ultrasonic and underwater systems

Brian T. Hefner^{a)} and Philip L. Marston

Department of Physics, Washington State University, Pullman, Washington 99164-2814

(Received 25 May 1999; accepted for publication 30 August 1999)

A simple four-panel transducer capable of producing a beam with a screw dislocation along its axis was constructed and evaluated. A screw dislocation in a wavefront is characterized by a phase dependence about the dislocation axis that varies as $\exp(-im\phi)$, where m is an integer and ϕ is the angle about the axis. At the axis, the phase is indeterminate and as a result there is a corresponding null in the pressure magnitude. The screw dislocation in the transducer beam is along the beam axis and is found to exist in both the far- and near-fields of the transducer. This null then clearly indicates the axis of the beam at all distances and has the potential to be used as an aid in the alignment of objects in sonar experiments or other similar applications. The helicoidal wave is also shown to possess axial angular momentum. A related transducer was summarized previously [J. Acoust. Soc. Am. **103**, 2971 (1998)] and is also discussed here for the purposes of comparison. © 1999 Acoustical Society of America. [S0001-4966(99)04812-2]

PACS numbers: 43.35.Yb, 43.38.Hz [HEB]

INTRODUCTION

A dislocation in a wavefront is characterized by an indeterminate phase at its core with a corresponding null in the wave magnitude.^{1,2} For a monochromatic, continuous traveling wave beam, this null typically forms a fixed curve in space. In a screw dislocation, the phase about the core varies as $\exp(-im\phi)$, where m is an integer, the sign of which indicates the charge or handedness of the beam. For a beam with $m = 1$, the planes of constant phase form a corkscrew or helicoid running in the direction of propagation as shown in Fig. 1. For higher values of m , the planes of constant phase consist of m interleaved helicoids. In the following, a beam with a screw dislocation at its center will be referred to as a helicoidal wave.

Although the existence and properties of screw dislocations in a wave field were first identified in the context of acoustics, a majority of the subsequent research has been on optical screw dislocations.^{3,4} Various researchers have noted the similarities between the structure of screw dislocations and vortices in fluids, leading them to describe screw dislocations in light as *optical vortices*.^{5,6} Indeed, under certain conditions optical vortices have been observed to interact with one another in ways that are very similar to vortices in fluids.⁷ It has also been shown^{8,9} that a helicoidal wave possesses angular momentum in its spatial distribution as opposed to circularly polarized light, which possesses angular momentum by virtue of its polarization.¹⁰ This property of the helicoidal wave has found application as a means to transfer angular momentum from a light beam to very small particles, allowing one to control the rotation of objects held in an optical trap. This is sometimes referred to as an optical spanner.¹¹

This paper describes a simple transducer for use in water that is capable of generating a beam which has many of the

important properties of a helicoidal wave. Previously, we presented a different type of transducer that could also generate a helicoidal beam and discussed several properties of this wavefront.^{12,13} That transducer used a less versatile design than is presented here, as will be discussed in Sec. I. The current design uses a four-panel configuration, each panel of which is driven separately and with the appropriate phase to generate a helicoidal beam. Details of its construction and operation are presented in Sec. I. This type of beam may have several uses in acoustics, one of which is the possibility of using this beam for the purposes of alignment. This axial null exists in the near-field of the transducer as well as the far-field, and this yields a well-defined axis relative to the transducer face. This null may be used as a more accurate means of aligning sonar targets or other equipment than a typical Gaussian beam; this application is discussed in Sec. II. The helicoidal wave also carries angular momentum as was discussed previously and is shown here in Sec. II. This property has been exploited in optics and may also find similar application in acoustics.

I. HELICOIDAL BEAM TRANSDUCER

To generate a helicoidal beam, the phase of the output needs to be a function of angle about the center of the transducer. Thus, to produce a beam with $m = 1$, the phase at the transducer face needs to possess the phase dependence of the radiation. There are several ways to produce this phase shift, two of which will be discussed here. It is worth noting that to generate the optical helicoidal beam, one typically starts with a Gaussian beam, such as the output of a laser, and transforms it using either a system of lenses⁸ or a grating⁴ to produce the desired structure. These techniques could be employed in acoustics; however, they are not discussed here.

In the transducer presented previously,^{12,13} the transducer face was physically offset to induce the proper phase in the output, as is shown in Fig. 2(a). That transducer consists of a ring of marine brass and an annular piece of PVDF,

^{a)}Electronic mail: bhefner@mail.wsu.edu

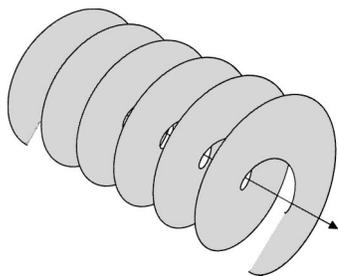


FIG. 1. Planes of constant phase for a wavefront containing a screw dislocation (helicoidal wavefront). The direction of propagation is indicated by the arrow which runs along the axis of the beam. The phase is indeterminate along the axis and there is a corresponding null in the magnitude.

which was stretched across the face of the ring. To produce the 2π phase shift about the axis of the beam, the ring and PVDF were cut at one point, which allowed the ring and transducer face to move like the coil of a spring. By offsetting one side of the cut, the height of the ring face, and hence the PVDF, becomes a continuous function of angle. If the offset width, d , at the cut is chosen such that it is equal to the wavelength of the sound produced by the PVDF, the surface height can then be described as $z_T = \lambda \phi / 2\pi$ and the generated beam should have the required angular dependence of $\exp(-i\phi)$. To vary the offset, a screw was threaded into a portion of the ring and the offset could be varied between 0 and 10 mm, beyond which the ring would begin to bend out of the required helical shape. This sets the lower limit for the operating frequency of the transducer at approximately 150 kHz.

This transducer produced a beam having many of the characteristics of a helicoidal beam; however, the design had several important limitations. Once the user set the offset distance for the operating frequency of interest, the transducer could only be used at that frequency. To change the frequency, the transducer would have to be physically removed from its position and the offset would have to be manually reset. Also, because unbacked PVDF was used, the transducer had a great deal of flexibility but it worked poorly when used to receive sound. This transducer simultaneously

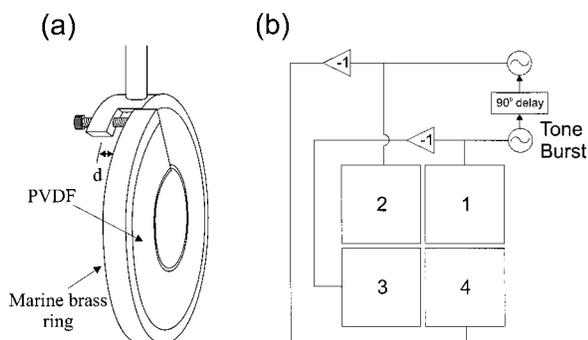


FIG. 2. Schematic of the transducers used to generate helicoidal wavefronts. (a) In the ring transducer, the phase ramp required to produce the helicoidal wave is introduced by the cut in the ring. When $d = \lambda$, the height of the surface is described as $z_T = \lambda \phi / 2\pi$, where λ is the wavelength of the beam and ϕ is the angle about the beam axis. (b) A four-panel 3-1 piezocomposite transducer was driven such that each panel was 90° out of phase with the previous panel. This gives the required 360° change in phase around the axis of the beam producing a helicoidal wavefront in the far-field.

radiated in the forward and backward directions.

To overcome these difficulties, a second transducer was built using a much simpler and more versatile design. This transducer was constructed from a 4×4 -in. 1-3 composite piezoelectric panel constructed by Material Systems, Inc. The panel is optimally designed to operate at 200 kHz but can perform from several kHz to nearly 600 kHz. The panel was etched such that there are four equal-area 2×2 -in. squares, each of which can be driven on a separate channel. An acoustic backing was attached to one side of the panel to minimize radiation in the backward direction and the entire system was encased in polyurethane. To produce the helicoidal wave, each section was driven with a tone burst and the phase was chosen such that each quadrant was driven 90° out of phase from the previous quadrant as shown in Fig. 2(b). The phasing was then 0° , 90° , 180° , and 270° , instead of the continuous phase ramp as in the previous design. In the far-field of this transducer, this should produce an approximately helicoidal wavefront.

The performance of the transducer was assessed in a redwood tank of 8 ft diameter and 8 ft depth. The transducer was placed 4 ft below the surface and 2 ft from the side of the tank. The transducer panels were driven at 60 kHz with 50-cycle tone bursts. An X - Y positioning system was used to scan a plane perpendicular to the axis of the transducer and 86 cm from the transducer's face. A hydrophone, attached to the scanning system, recorded the phase and magnitude of the acoustic field at 1-cm increments, the results of which are displayed in Figs. 3(a) and 4. Although the beam does not possess cylindrical symmetry, it does possess the required screw dislocation at its center. Close to the beam axis, the phase becomes helical in nature and at the axis it becomes indeterminate. This produces the desired near-null in the magnitude of the beam along the axis. To demonstrate that this null exists along the axis, even into the near-field of the beam, the data in Figs. 3(a) and 4 were numerically back-propagated using acoustic holography techniques.¹⁴ Figure 3(b) shows the beam at 43 cm, half the distance to the transducer, and as expected the null along the axis is still present. Although the beam may become complicated as the propagation distance is decreased, the null will still be present, as is discussed in more detail below. It is interesting to note that the screw dislocation at the center is not the only dislocation present in the beam. There are also four prominent edge dislocations,¹ as can be seen in the phase distributions. At the cores of these dislocations, the phase becomes indeterminate as well and there are four corresponding nulls in the magnitude. Unlike the dislocation along the beam axis, these dislocations are not fixed and depend on the distance from the transducer face.

With this design, the frequency of the output can be changed as well as the helicity of the beam without disrupting the alignment of the transducer. Also, the transducer can be switched from operating with $m = 1$ to $m = 0$ where there is no dislocation present in the beam. Unlike the previous design, where it was possible to set $d = m\lambda$, this transducer cannot operate with $m > 1$, although this design could be generalized to a larger number of panels or sections to produce this type of beam. The piezocomposite material also works

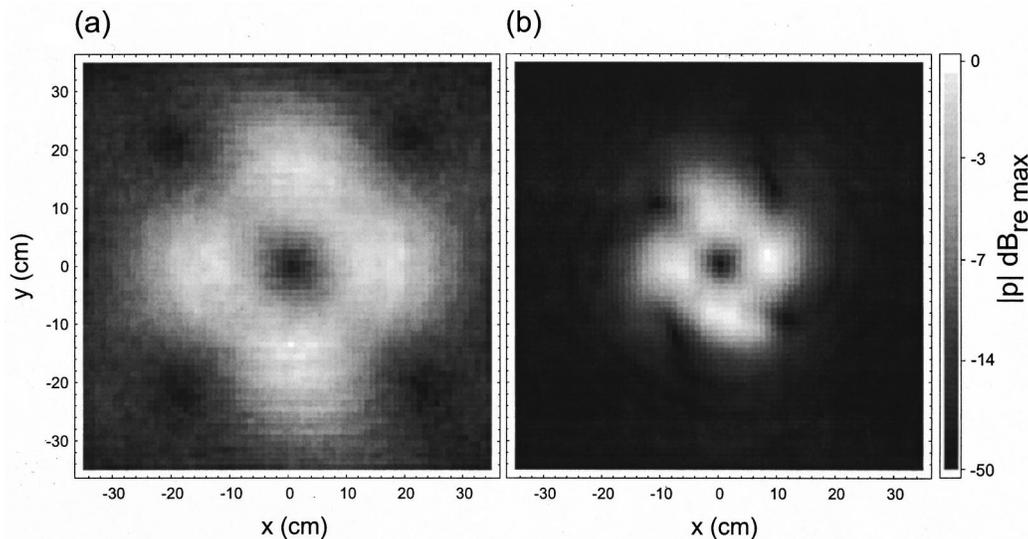


FIG. 3. Output of the panel transducer operating at 60 kHz. (a) The normalized magnitude was measured in a plane perpendicular to the beam's axis at a distance of 86 cm from the transducer. The output is displayed on a decibel scale measured relative to the maximum response of the receiver. (b) The magnitude and phase information was numerically backpropagated to 43 cm to confirm the existence of the axial null within the Rayleigh distance of the transducer, $R = 54.8$ cm.

very well for receiving as well as transmission. This particular transducer has been used for this type of application with all of the panels connected; however, it has not been used to send and receive with each panel on a separate channel. There are several applications, such as scattering, for which this would be a very useful configuration. From symmetry considerations, specular reflections of an $m = 1$ beam from axisymmetric objects may be suppressed or enhanced using the superposition of signals from appropriately phased detectors.¹³

II. DISCUSSION

As mentioned above, this helicoidal beam may be a very useful tool for the alignment of sonar experiments or other ultrasonic and underwater applications where objects need to be placed along a common axis. The helicoidal beam produced a well-defined null along the axis of the beam which exists in both the near- and far-fields. Its existence for the particular transducer discussed above can be understood by realizing that along the axis, the output of each panel interferes destructively with the panel across from it. Regardless of how complicated the overall structure of the beam may become in the near-field, this interference should be maintained producing a continuous axial null. Using a hydrophone, the location of the axis could be found at the desired distance from the transducer and then replaced by the target to insure that the target is placed on the beam axis. A more accurate method might be to use the transducer in a send and receive mode and look at the specular reflection from the target using the helicoidal beam. As noted previously, the reflected signal would have the opposite helicity and if the object is aligned with the axis, the received signal should be suppressed. Furthermore, for axisymmetric targets the back-scattered amplitude vanishes if measured on the axis.

To further confirm the presence of a continuous null into the near-field of the transducer, a scan similar to the one

presented in Fig. 3 was numerically backpropagated as in Fig. 3(b) to track the existence of the null. For this scan, the transducer was operated at 100 kHz and the sample plane was 95 cm from the transducer face. A cross section of these backpropagation results corresponding to the $x-z$ plane is given in Fig. 5. For the transducer operating at this frequency, the Rayleigh distance is taken to be $R \approx (1/2)ka^2 = 54.8$ cm, where $a = 5.1$ cm is the half-width of the transducer. As seen in Fig. 5, the null exists well within the Rayleigh distance of the transducer. Also, notice that the width of the axial null decreases as one approaches the face of the transducer. By increasing the operating frequency, because the null exists at all distances along the beam's axis regardless of the Rayleigh distance, the width of the null can be

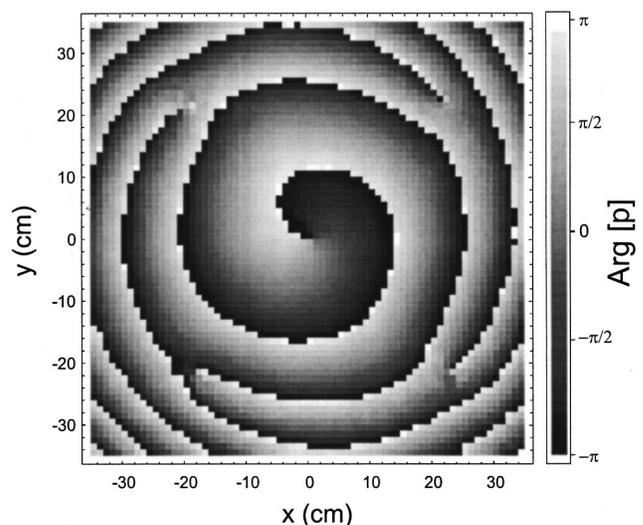


FIG. 4. The phase of the output measured in a plane perpendicular to the beam's axis at a distance of 86 cm from the transducer. At the center of the beam the phase becomes indeterminate and comparison to Fig. 5 indicates that this corresponds to the location of the null in the magnitude of the wave field.

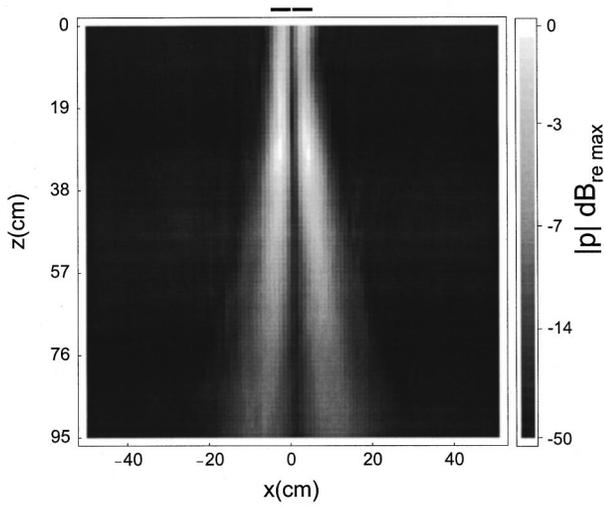


FIG. 5. Cross section of the transducer beam operating at 100 kHz. This was obtained by numerical backpropagation of the transducer output measured 95 cm from the transducer as discussed in Sec. II. The two lines at the top of the graph indicate the location of the panel transducer.

made tighter for longer distances. This could increase the accuracy in determining the axis location and hence the alignment of objects along this axis.

To examine the nature of the angular momentum transport in helicoidal waves, one can use a relatively simple mathematical description of the beam which is often used in optics for pure modes. The Laguerre–Gaussian (LG) beam is a solution to the paraxial wave equation in cylindrical coordinates and has many of the important properties of the helicoidal wave.⁸ With θ denoting the azimuthal angle the solutions for the acoustic paraxial wave equation are

$$p_{mn}(r, \theta, z) = A_{mn} \left[1 + \left(\frac{z}{z_R} \right)^2 \right]^{-1/2} e^{-(r/w)^2} e^{ikr^2/2R} e^{-i\Psi} \times e^{im\theta} ((r/w)\sqrt{2})^{|m|} L_n^{|m|}(2r^2/w^2) e^{ikz}, \quad (1)$$

where $z_R = \pi w_0^2/\lambda$ is the Rayleigh range, $w(z) = [2(z_R^2 + z^2)/kz_R]^{1/2}$ is the local beam width, $w_0 = w(0)$, $R = (z_R^2 + z^2)/z$, $\Psi = (|m| + 1 + 2n)$ is a generalized Guoy phase shift, A_{mn} is a constant, and $L_n^{|m|}$ is the associated Laguerre polynomial. The angular index, m , can take on the values $m = 0, \pm 1, \pm 2, \dots$ and the $e^{im\theta}$ term introduces the ramp in the phase about the z -axis. For $m = 0$, the solution reduces to the Gaussian beam solution with no angular variation in phase. For $m \neq 0$, the $((r/w)\sqrt{2})^{|m|}$ term insures that for these values of m , there is the required pressure null along the axis of the beam. The radial index, n , can take on values $n = 0, 1, 2, \dots$ and this index affects the number of radial nodes in the beam. For the present analysis, we consider LG beams with $n = 0$.

Let $\varphi = (ip/\omega\rho)$ denote the complex velocity potential and $\mathbf{v} = -\nabla \text{Re}[\varphi]$ be the fluid velocity where Re denotes the real part. From Eq. (1), the azimuthal velocity is $v_\theta = \text{Re}[-im\varphi/r]$. The average axial angular momentum density of the beam is $\langle (\delta\rho)r v_\theta \rangle$, where $\delta\rho = (c^{-2})\text{Re}[p]$ is the first-order change in density due to the acoustic wave and $\langle \rangle$ denotes a time average. The angular momentum flux $\langle L_z \rangle$ and power P of the beam are

$$\langle L_z \rangle = 2\pi c \int_0^\infty \langle (\delta\rho)r v_\theta \rangle r dr, \quad (2)$$

and

$$P = 2\pi \int_0^\infty \langle \text{Re}[p] \text{Re}[-ik\varphi] \rangle r dr, \quad (3)$$

where P is the integral of the local average acoustic intensity $\langle v_z \text{Re}[p] \rangle$. Inspection of Eq. (2) gives $\langle L_z \rangle/P = m/\omega$ for each value of m and n in Eq. (1). This ratio is the same as for an electromagnetic beam.⁸ Consequently, absorption of acoustic energy from beams with $m \neq 0$ will produce an axial torque on an absorber. Wang *et al.*^{15,16} have investigated torques associated with the absorption of appropriately phased standing waves; however, the torques considered here are associated with the absorption of traveling waves.

ACKNOWLEDGMENT

This research was supported by the Office of Naval Research.

- ¹J. F. Nye and M. V. Berry, Proc. R. Soc. London, Ser. A **336**, 165–190 (1974).
- ²M. V. Berry, in *Physics of Defects*, edited by R. Balian, M. Kleman, and J.-P. Poirer (North-Holland, Amsterdam, 1981), pp. 453–543.
- ³V. Yu. Bazhenov, M. S. Soskin, and M. V. Vasnetsov, J. Mod. Opt. **39**, 985–990 (1992).
- ⁴I. V. Basistiy, V. Yu. Bazhenov, M. S. Soskin, and M. V. Vasnetsov, Opt. Commun. **103**, 422–428 (1993).
- ⁵G. Indebetouw, J. Mod. Opt. **40**, 73–87 (1993).
- ⁶D. Rozas, C. T. Law, and G. A. Swartzlander, Jr., J. Opt. Soc. Am. B **14**, 3054–3065 (1997).
- ⁷D. Rozas, C. T. Law, and G. A. Swartzlander, Jr., Phys. Rev. Lett. **79**, 3399–3402 (1997).
- ⁸L. Allen, M. W. Beijersbergen, R. J. C. Speeuw, and J. P. Woerdman, Phys. Rev. A **45**, 8185–8189 (1992); L. Allen, M. J. Padgett, and M. Babiker, “The orbital angular momentum of light,” Prog. Opt. **39**, 291–372 (1999).
- ⁹S. M. Barnett and L. Allen, Opt. Commun. **110**, 670–678 (1994).
- ¹⁰P. L. Marston and J. H. Crichton, Phys. Rev. A **30**, 2508–2516 (1984).
- ¹¹N. B. Simpson, K. Dholakia, L. Allen, and M. J. Padgett, Opt. Lett. **22**, 52–54 (1997).
- ¹²B. T. Hefner and P. L. Marston, J. Acoust. Soc. Am. **103**, 2971 (1998).
- ¹³B. T. Hefner and P. L. Marston, in *Proceedings of the 16th International Congress on Acoustics*, edited by P. K. Kuhl and L. A. Crum (ASA, Seattle, 1998), pp. 1921–1922.
- ¹⁴E. G. Williams and J. D. Maynard, Phys. Rev. Lett. **45**, 554–557 (1980).
- ¹⁵F. H. Busse and T. G. Wang, J. Acoust. Soc. Am. **69**, 1634 (1981).
- ¹⁶T. G. Wang, E. H. Trinh, A. P. Croonquist, and D. D. Ellenman, Phys. Rev. Lett. **56**, 452 (1986).

One-dimensional phenomenological model of hysteresis.

I. Development of the model

Jean C. Piquette^{a)} and Stephen E. Forsythe

Naval Undersea Warfare Center, Code 2161, 1176 Howell Street, Newport, Rhode Island 02841

(Received 12 April 1999; accepted for publication 30 August 1999)

A one-dimensional phenomenological model of hysteresis is presented. The model is suitable for application both to electrostrictive materials, such as lead magnesium niobate (PMN), and to magnetostrictive materials, such as Terfenol D. The concepts of “inflation,” “field space,” and the “reference ellipse” are introduced as suitable mechanisms for transforming measured hysteretic data into corresponding anhysteretic versions. An anhysteretic model is then fitted (in the least-squares sense) to the transformed data. By applying the inverse transforms to the fitted anhysteretic model, a hysteretic model is deduced. Good agreement with the original (hysteretic) data is seen. It is shown that, when a sample is driven by a monofrequency electric field, the area of the polarization vs electric-field hysteresis loop is independent of all harmonics in the polarization but the first. A principle useful for understanding the shapes assumed by $P-E$ and $M-H$ hysteresis loops in general is described. [S0001-4966(99)03712-1]

PACS numbers: 43.38.Ar, 43.20.Px, 43.30.Yj, 43.58.Vb [SLE]

INTRODUCTION

A phenomenological model of hysteresis is given. The model is suitable for application both to electrostrictive materials, such as lead magnesium niobate (PMN) and to magnetostrictive materials, such as Terfenol D. It is emphasized that the model presented here is phenomenological in character, as opposed to a fundamental or “first-principles” approach. One goal of the work is to predict material performance under drive conditions other than those used to obtain data. Another goal is to permit the use of drive-preforming techniques, in which a drive is designed to produce a desired output trajectory, such as to reduce harmonic distortion in the output. (Such a capability is desirable to permit application of signal-processing techniques, which may require linearity in the output signal.) It is also hoped that the model can be used as a transducer design tool. Last, it is desired to identify parameters that characterize the material sufficiently accurately, so that experimenters might be guided toward those measurements that are most critical in determining material performance.

The approach used to develop the model is as follows:

- (1) Apply suitable (reversible) transforms to measured hysteretic data to produce (approximately) anhysteretic data.
- (2) Fit an anhysteretic model to the anhysteretic (transformed) data.
- (3) Apply the inverse of the transforms to the anhysteretic model.
- (4) Compare the resulting model with the original (untransformed) hysteretic data.

This procedure can be applied to any anhysteretic theory,

and is not limited to our particular theory, as considered here.

In developing transforms for step 1, it should be remarked that simply using the average of the two values of the response occurring on the lower (increasing drive field) and upper (decreasing drive field) segments of the hysteresis loop at each drive-field value would be unsatisfactory. Such a procedure is not reversible. That is, simply knowing that a given field value is the average of two quantities is not sufficient to recover the two individual quantities that were averaged to produce it. Therefore, step 3 of the process could not be carried out if such an averaging were used to produce effectively anhysteretic data.

In order to achieve the transforms required in step 1 of the process, the concepts of “inflation,” “field space,” and the “reference ellipse” are introduced. Inflation is the process of introducing an additional dimension into (i.e., “inflating”) the experimental data, thus converting the conventional two-dimensional “hysteresis loops” into three-dimensional “hysteresis hoops.” The extra coordinate introduced is that of the time derivative of the driving field. We use the time derivative of the driving field simply because it is clear that if this derivative is zero, no hysteresis can occur. That is, even under quasistatic conditions the drive field *must* change in order for hysteresis to be manifest. So, we believe this derivative likely plays some role in the physical processes that underlie the hysteresis phenomenon.

Field space is the term used to describe the resulting three-dimensional coordinate system. [For electrostrictive materials, field space is formed by the coordinates (E, P, \dot{E}) , with E the electric field and P the polarization. For magnetostrictive materials, field space is formed by the coordinates (H, M, \dot{H}) , with H the magnetic field and M the magnetization. For PMN, which is the electrostrictive material of primary interest, the permittivity is so large that the P coordi-

^{a)}Electronic mail: piquettejc@npt.nuwc.navy.mil

nate can be freely replaced by the electric displacement D . In the discussion that follows, we will indeed use the P and D fields interchangeably, although the reader should bear in mind the distinction. We will also refer explicitly only to electrostrictive materials in the following discussion, but the applicability to magnetostrictive materials should also be kept in mind.] The desired transform to produce the anhysteretic version of the D vs E data is achieved by performing suitable rotations of the D - E hysteresis hoop in field space such that the hoop's shadow (projection) in the D - E plane is as close as possible to a zero-area curve.

As a step in deriving the required rotations in field space, it is shown that when a sample is driven with a monofrequency electric field, the area of the resulting P - E hysteresis loop is independent of all harmonics of the polarization but the first. (Second, and higher, harmonics in the polarization affect only the loop's *shape*, not its area.) Hence, it is possible to discard all harmonics of the polarization above the first without affecting the net loss per cycle per unit volume of the sample. When all harmonics of the polarization but the first are discarded and the hysteresis loop is redrawn, it is an ellipse. The three-dimensional (inflated) version of this figure is also an ellipse. We term these two figures the "two-dimensional reference ellipse" and "three-dimensional reference ellipse," respectively. Two of the rotation angles required in field space to produce anhysteretic data involve rotations of the reference ellipses. One of these is a rotation of the three-dimensional reference ellipse about the \dot{E} -axis, which aligns the major axis of the associated two-dimensional reference ellipse with the E -axis. The second is a rotation of the resulting three-dimensional reference ellipse about the E -axis, which causes the area of the associated two-dimensional reference ellipse to become zero.

Differences in symmetry between the D vs E curve and the S vs E curve do not permit the same field-space rotations, which transform the D - E curve into anhysteretic form, to be applied to the problem of transforming the S - E curve. To help achieve that transformation, we postulate that the exact proportionality of strain upon the square of the electric displacement, postulated previously in our anhysteretic theory,^{1,2} extends without modification to hysteretic data. Once hysteretic D -data has been transformed into anhysteretic form by the proposed field-space rotations, it is shown that hysteretic strain data can be accommodated as well using the extended strain postulate.

We note that there are also other theories of hysteresis.³⁻⁹ However, most of these are based on modified Preisach modeling,³ and many can have difficulty in producing realistic (closed) loops for asymmetric drives. (Since asymmetric drives are required for many practical systems, such as PMN and Terfenol-D-based transducers, which require biasing, this is an important consideration.) One advantage of the present approach is that it always produces closed loops, regardless of drive conditions.

The concepts of inflation and field space are introduced in Sec. I. Field-space rotations that produce anhysteretic polarization data are described in Sec. II, while the extended strain postulate, required to achieve the anhysteretic transform of the S - E curve, is described in Sec. III. An applica-

tion of the transformations to experimental data is presented in Sec. IV, and the process of applying the inverse transformations to an anhysteretic model to produce a hysteretic model is described in Sec. V. A summary, discussion, and conclusion are given in Sec. VI. Appendix A describes a modification of our one-dimensional anhysteretic model. Appendix B describes the concept of the reference ellipse, and gives its important properties. In this Appendix a proof is given that the P - E and M - H loop area is independent of all harmonics but the first. A general principle for understanding the shapes assumed by P - E and M - H hysteresis loops in general is also given. Finally, a form of data preprocessing we use is described in Appendix C.

I. INFLATION AND FIELD SPACE

The first step in the accommodation of hysteresis is the introduction of the time rate of change of the electric field \dot{E} as an extra dimension in the usual plots of D -field vs E -field. This process is termed "inflation," and the resulting three-dimensional coordinate system is termed "field space." Inflation can be performed either on experimental data or on a theory. (The process of inflation and field-space rotations described here can be applied to *any* anhysteretic theory to produce a corresponding hysteretic version.) In the case of experimental data, inflation is performed by first doing a linear least-squares fit of the measured driving field to determine the quantities A and B in the equation

$$E = A \sin(\omega t) + B \cos(\omega t). \quad (1)$$

[Here, ω is the known angular frequency at which the experiment was conducted. We remind the reader that we are developing the model in terms of its application to electrostrictive materials for convenience of expression, but that it is equally applicable to magnetostrictive materials. For magnetostrictive measurements, of course, the electric field E of Eq. (1) is to be replaced by the magnetic field H .] The time derivative of the electric field \dot{E} is then computed analytically by differentiation of Eq. (1). The two-dimensional graph of D vs E is then inflated to three dimensions by incorporating \dot{E} as the coordinate associated with the third dimension. (The resulting E , D , \dot{E} coordinate system constitutes field space.) One data set for which this computation has been carried out is examined in Fig. 1(a)-(c). These data were acquired by the experimental group led by James Powers at the Naval Undersea Warfare Center, Division Newport. In Fig. 1(a) the original D vs E plot is shown. (The graph of strain vs E , shown in Fig. 1(b), is presented now for completeness. It will be used subsequently.) In Fig. 1(c) the three-dimensional version, which results from inflating the data of Fig. 1(a), is shown. Shadows (projections) of the three-dimensional hysteresis "hoop" are shown in each coordinate plane. The shadow in the D - E plane of Fig. 1(c) is identical to the hysteresis "loop" seen in the original two-dimensional plot of Fig. 1(a). Using the process of inflation to derive a hysteretic model from an anhysteretic model is described in a following section.

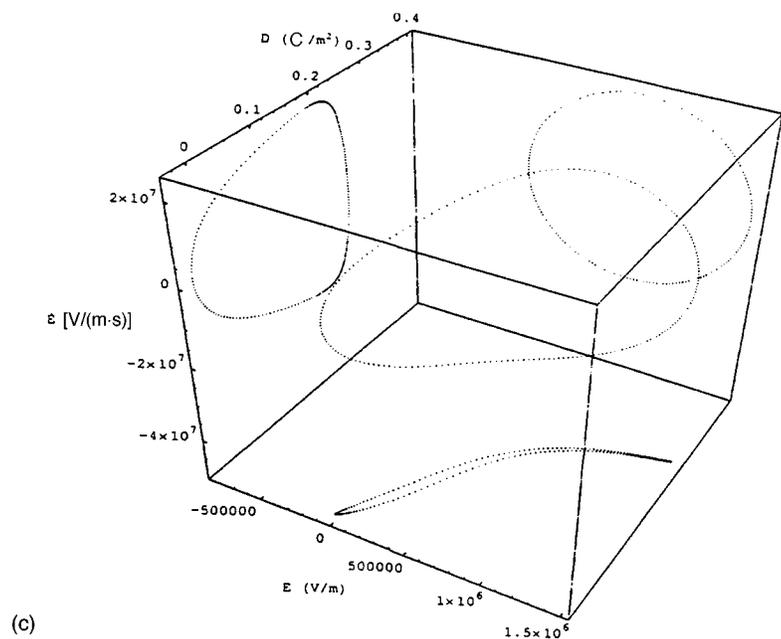
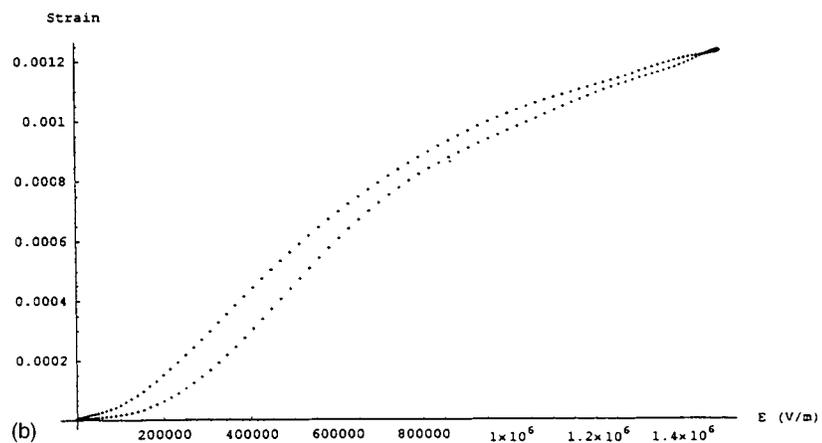
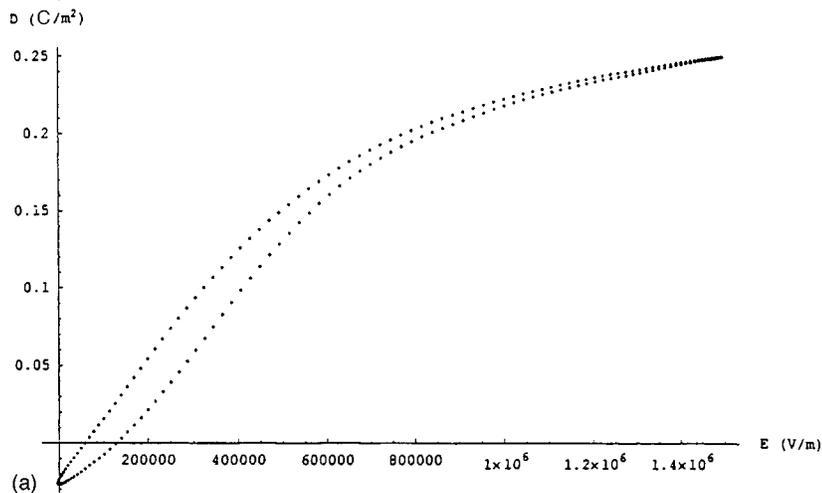


FIG. 1. (a) Experimental $D-E$ hysteresis loop for a strontium-doped sample of PMN measured at -5°C and a compressive prestress of 13.8 MPa. (b) Experimental $S-E$ hysteresis loop corresponding to the data of (a). (c) Inflated version of the data of (a). Shadows of the three-dimensional hysteresis "hoop" are shown in each coordinate plane. The shadow in the $D-E$ plane is identical to (a).

In generating the graphs of Fig. 1, the original data have been fitted to a Fourier series truncated at the tenth harmonic. (It is the series that is plotted, not the original raw data.) This is done both because it reduces the effects of the noise present in the original measurements, and because the harmonic content of the signals is required for subsequent calculations.

II. ROTATIONS IN FIELD SPACE

To convert hysteretic data into a form suitable for application of an anhysteretic theory, we propose that rotations be carried out in field space of the inflated $D-E$ hoop such that the area of its shadow in the $D-E$ coordinate plane is minimized. For a purely anhysteretic theory this area would of

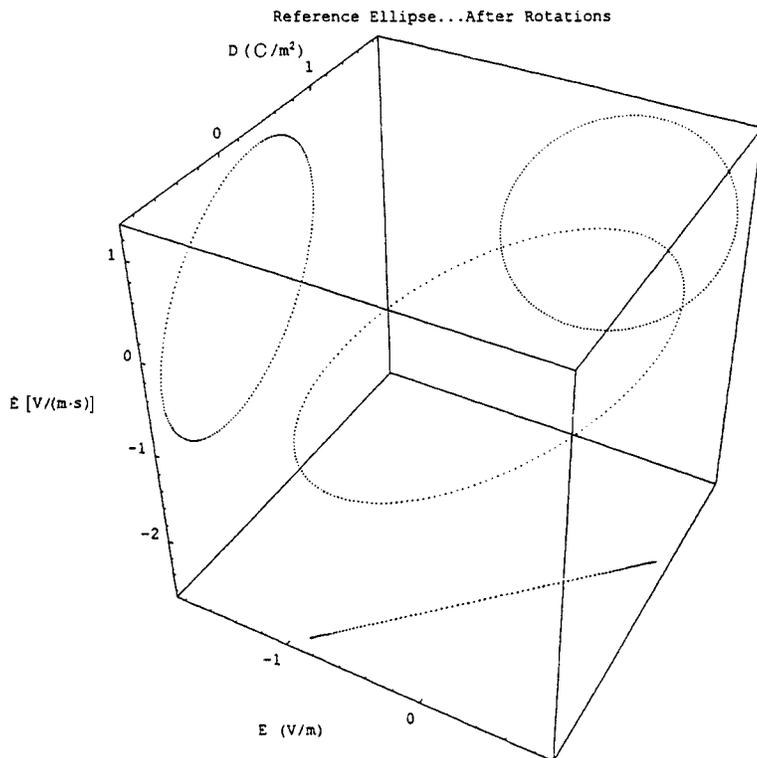


FIG. 2. Reference ellipses after the rotations of Eq. (3) have been applied.

course be zero, signifying that no losses are present. (In analyzing data newly available to us, we have found it necessary to revise slightly our earlier theory¹ to accommodate unexpected behavior near the origin. This revision is described in Appendix A. It is worthwhile noting, however, that similar deviant behavior in the vicinity of the origin has been seen by others. See, e.g., Ref. 6, Fig. 1.)

Our prescription for producing a minimal-area $D-E$ shadow is as follows:

- (1) First, perform a rotation of the three-dimensional reference ellipse about the \dot{E} -axis so that the major axis of its shadow in the $D-E$ plane is coincident with the E -axis. Denote the required rotation angle by ψ_a , and the rotation matrix that implements this rotation by Ψ_a .
- (2) Next, perform a rotation of the three-dimensional reference ellipse about the E -axis so that the area of its shadow in the $D-E$ plane is zero. Denote the required rotation angle by γ_{ab} , and the rotation matrix that implements this rotation by Γ_{ab} .
- (3) Perform another rotation of the three-dimensional reference ellipse about the \dot{E} -axis that “undoes” the rotation of step 1. That is, rotate about \dot{E} by $-\psi_a$.

(The rotation angles required in these steps can be computed analytically. Their calculation is considered in Appendix B.)

The sequence of rotations applied to the three-dimensional reference ellipse outlined in steps 1–3 can be expressed in matrix notation as $\Psi_a^{-1}\Gamma_{ab}\Psi_a$, where Ψ_a^{-1} denotes an inverse rotation matrix, achieved by the rotation described in step 3. These rotations are applied to the inflated data vector, and the process can be represented in matrix notation as

$$\begin{pmatrix} E' \\ D' \\ \dot{E}' \end{pmatrix} = \Psi_a^{-1}\Gamma_{ab}\Psi_a \begin{pmatrix} E \\ D \\ \dot{E} \end{pmatrix}. \quad (2)$$

The primed variables of Eq. (2) denote the post-rotation forms of the data.

In order to produce the anhysteretic data required for application of an anhysteretic theory, one additional rotation of the original data in field space beyond those described for the reference ellipse is required. If the rotations summarized by Eq. (2) are applied directly to the full harmonic-containing data hoop of Fig. 1(c), the resulting shadow in the $D-E$ plane, while exhibiting zero net area, achieves this zero net from the undesirable characteristic of “twists” in the shadow. That is, the zero net area is achieved from cancellations of positive- and negative-area contributions in the shadow. To limit this, one additional rotation is performed, after the Γ_{ab} rotation and prior to the Ψ_a^{-1} rotation, and is done about the D -axis. (Since the two-dimensional reference ellipse has zero area after the Γ_{ab} rotation has been applied, and is aligned with the E -axis prior to applying the Ψ_a^{-1} rotation, this additional rotation about the D -axis cannot change this area to a nonzero value. Hence, the zero area of the projected two-dimensional reference ellipse is unaffected by this additional rotation.) The required rotation angle is empirically determined as that minimizing the area of each of the twists in the shadow in the $D-E$ plane of the full harmonic-containing data hoop. Denote the required rotation angle by ξ_D , and its associated rotation matrix by Ξ_D . The full process of generating anhysteretic data from the measured hysteretic data can then be summarized by the matrix equation

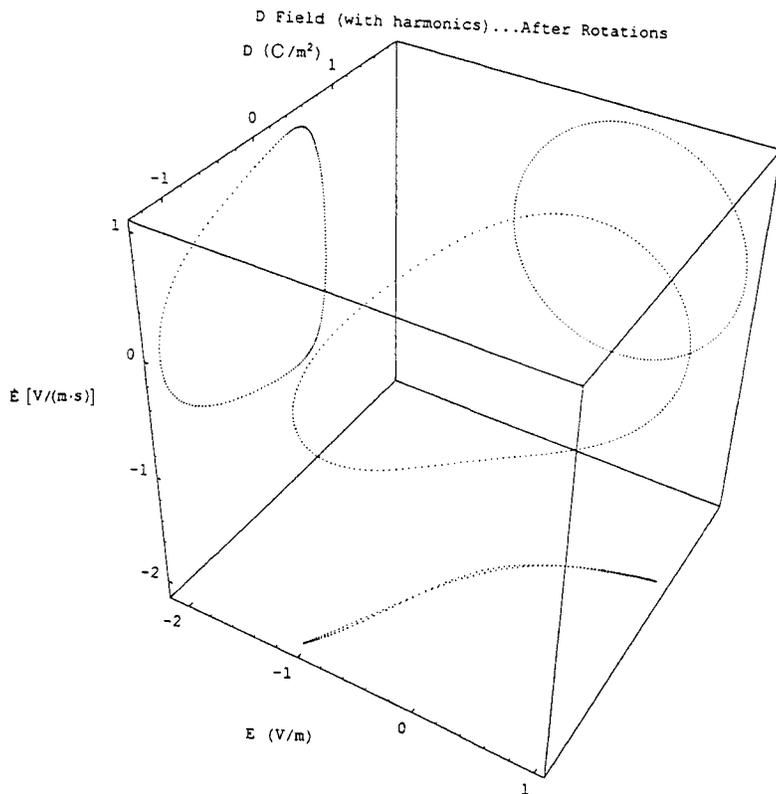


FIG. 3. Transformed version of inflated D - E data together with shadows in the coordinate planes after the rotations of Eq. (3) have been applied. Data are scaled to unity first harmonic.

$$\begin{pmatrix} E' \\ D' \\ \dot{E}' \end{pmatrix} = \Psi_a^{-1} \Xi_D \Gamma_{ab} \Psi_a \begin{pmatrix} E \\ D \\ \dot{E} \end{pmatrix}. \quad (3)$$

The result of applying the rotations summarized by Eq. (3) to the reference ellipse (introduced in Appendix B and depicted in Fig. B1) is shown in Fig. 2. The result of applying the rotations to the full harmonic-containing data hoop of Fig. 1(c) is shown in Fig. 3. Finally, the shadow of the transformed data is shown in Fig. 4.

III. TRANSFORMATION OF THE STRAIN

We next consider the transformation of the strain data into anhysteretic form. While one might suppose that applying a series of rotations to the strain that are analogous to those applied to the D -field might achieve the desired trans-

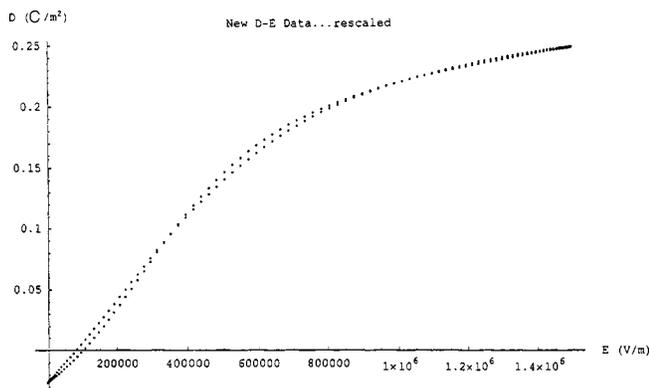


FIG. 4. Shadow in the D - E plane from Fig. 3 redrawn in the original two-dimensional coordinate system. Data are rescaled to original maximum values.

formation, differences in symmetry between the strain and the electric displacement do not permit this. Transformation of the strain data into anhysteretic form therefore requires a different approach. This transformation is accomplished with the help of postulating that the strain equation of the original anhysteretic theory applies without modification when hysteresis is present. In the current section, we discuss how this additional postulate helps achieve the required transformation.

Our one-dimensional anhysteretic theory postulates for the strain the equation

$$S = \beta_1 T + \beta_2 D^2. \quad (4)$$

Here, T is the external stress and β_1 and β_2 are constants. (β_1 is related to the Young's modulus of the material and β_2 to the electrostriction constant.) In the experimental data available to us, strains associated with the application of a prestress are subsumed within the calibration of the zero of strain. It is therefore convenient to transpose the $\beta_1 T$ term of Eq. (4) to the same side of the equation as S , introducing a new strain variable \hat{S} that includes this term. Moreover, we introduce a prime notation suggestive of the fact that the data in question are anhysteretic, hence

$$\hat{S}' = \beta_2 (D')^2. \quad (5)$$

[The prime notation is intended to imply anhysteretic data in the same manner as the prime notation of Eq. (3).] If we now let variables without primes denote hysteretic data, our new postulate assumes the form

$$\hat{S} = \beta_2 D^2. \quad (6)$$

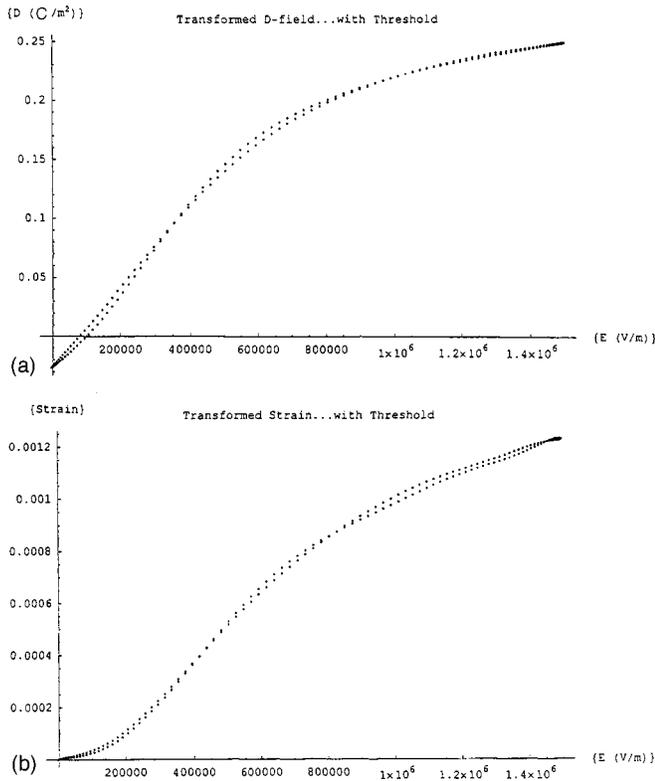


FIG. 5. (a) Transformed D -data after thresholding is applied. (b) Transformed S -data after thresholding is applied.

Forming a ratio between Eq. (5) and Eq. (6) and solving for \hat{S}' derives the anhysteretic form of the strain data. The result is

$$\hat{S}' = (D'/D)^2 \hat{S}. \quad (7)$$

The unprimed quantities of Eq. (7), i.e., hysteretic data, represent measured experimental quantities. The quantity D' on the right-hand side of Eq. (7) is available once the field-space transformation of Eq. (3) has been completed. Thus, Eq. (3) and Eq. (7) constitute a complete prescription for transforming hysteretic data into anhysteretic form.

IV. APPLICATION TO DATA

In applying Eq. (7) to actual data, which are always contaminated to some extent by noise, it is obviously necessary to avoid data for which D is close to zero. We do this by simply setting an arbitrary threshold value for exclusion. [That is, data points with values close to zero are excluded, owing to the disproportionate influence such values would have on the ratio presented in Eq. (7).]

In Fig. 5(a) and (b), respectively, are shown the transformed D vs E and S vs E data. The data shown in Fig. 5(a) result from application of Eq. (3) to the data of Fig. 1(a). The data of Fig. 5(b) result from applying Eq. (7) to the data of Fig. 1(b) [i.e., the quantity S of Eq. (7)]. Also used in the calculation are the data of Fig. 1(a) [to obtain the quantity D of Eq. (7)], and the data of Fig. 5(a) [to obtain the quantity D' of Eq. (7)]. The thresholding process has not affected the plots shown in Fig. 5(a) and (b) owing to the effects of a form of preprocessing, which has been applied to the D -field

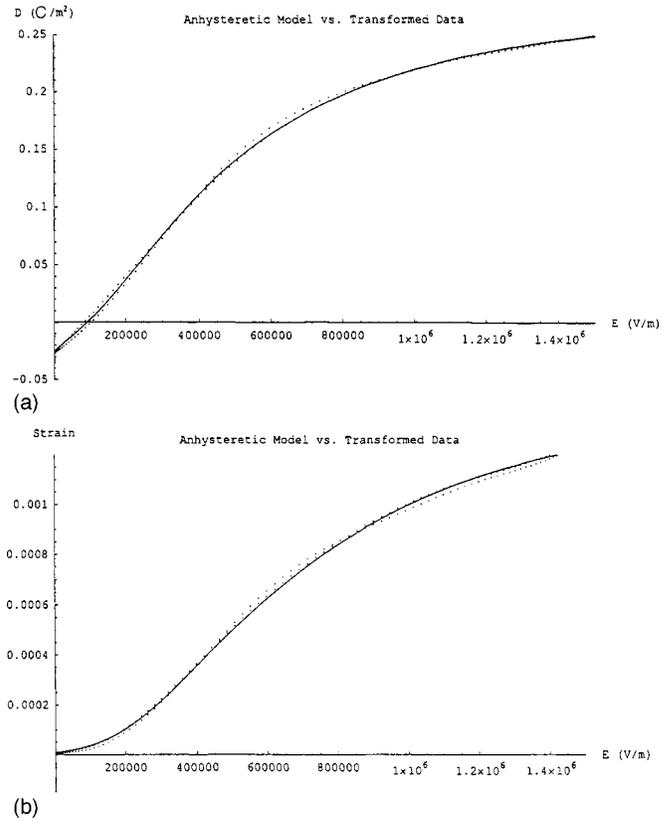


FIG. 6. (a) Anhysteretic theory (solid line) compared with transformed data (dots) for D after least-squares fitting. (b) Anhysteretic theory (solid line) compared with transformed data (dots) for S after least-squares fitting.

data. This preprocessing results, in this particular case, in the elimination of any D values near zero. The preprocessing is described in Appendix C.

V. MODEL OF HYSTERESIS

A hysteretic version of the model is now produced using the following prescription:

- (1) Apply the transformations of Eq. (3) and Eq. (7) to the measurements to produce (approximately) anhysteretic data.
- (2) Perform a least-squares fit of the anhysteretic theory to the anhysteretic data to determine “best-fit” anhysteretic model parameters in the least-squares sense.
- (3) Apply the inverses of the transformations, which have been applied to the hysteretic data, to the anhysteretic theory.

Figure 5(a) and (b) are the result of applying step 1 of this procedure. The results of applying step 2 are shown in Fig. 6(a) and (b). [Note, especially from Fig. 6(a), the change in slope of the curve in the vicinity of the origin, necessitating the revised anhysteretic model described in Appendix A. Compare also with Fig. 1 of Ref. 6.] Figure 7(a) and (b) show the results of applying step 3. In generating the theoretical curves, all the same rotation angles are applied as in generating the anhysteretic version of the data, except negative values of the angles are used. Of course, the correspond-

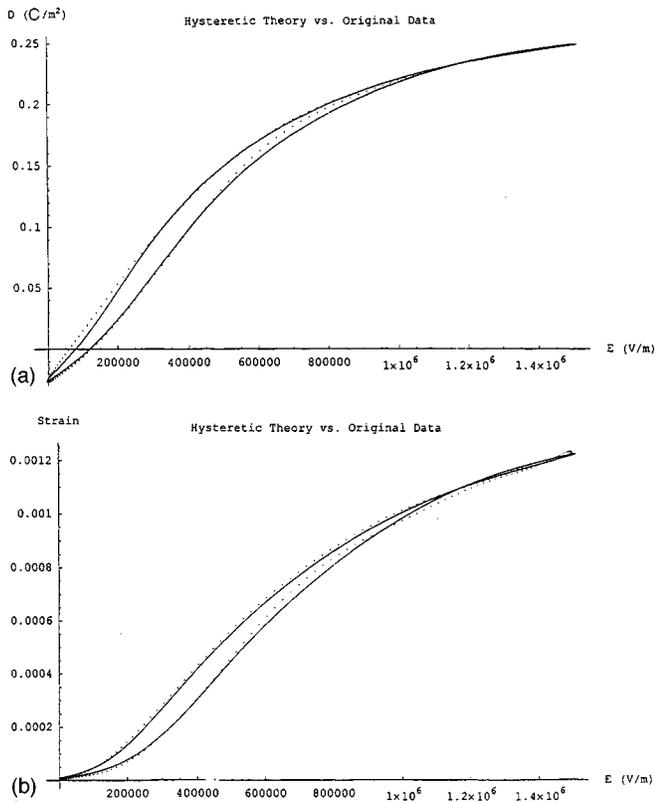


FIG. 7. (a) Hysteretic theory compared with original data for D . (b) Hysteretic theory compared with original data for S .

ing theoretical values for the electric field E and its derivative \dot{E} are used.

One difficulty with the theoretical curves of Fig. 7(a) and (b) are the apparent "twists" in the curves (i.e., the curves cross themselves). However, this is actually a consequence of the fact that the experimental D vs E curve also exhibits such a twist. [The twist may be an artifact of the truncated Fourier series used to process the data. Although this twist is not apparent in Fig. 1(a) it can be seen with sufficient magnification of the plot.] This difficulty can be corrected by performing a form of "preprocessing" of the data. In particular, if a 0.5-deg rotation in field space of the D vs E curve is performed about the E -axis, the twist in this curve is eliminated. (To maintain consistency of the D vs E curve with the S vs E curve, a similar rotation is applied to the strain measurements, also scaled to unity first harmonic.) When the procedure is applied to data that have been preprocessed in this manner, the results shown in Fig. 8(a) and (b) are obtained. As can be seen, the twists in the model are eliminated, or at least have been greatly reduced. The primary consequence of this preprocessing is that the value obtained for the loss tangent is changed with respect to the result obtained when the data are not preprocessed. The loss tangent for the present case is of a magnitude approximately 0.05 prior to preprocessing, but becomes approximately 0.06 after preprocessing.

VI. SUMMARY, DISCUSSION, AND CONCLUSION

A one-dimensional model of hysteresis has been described. The procedures described here are independent of

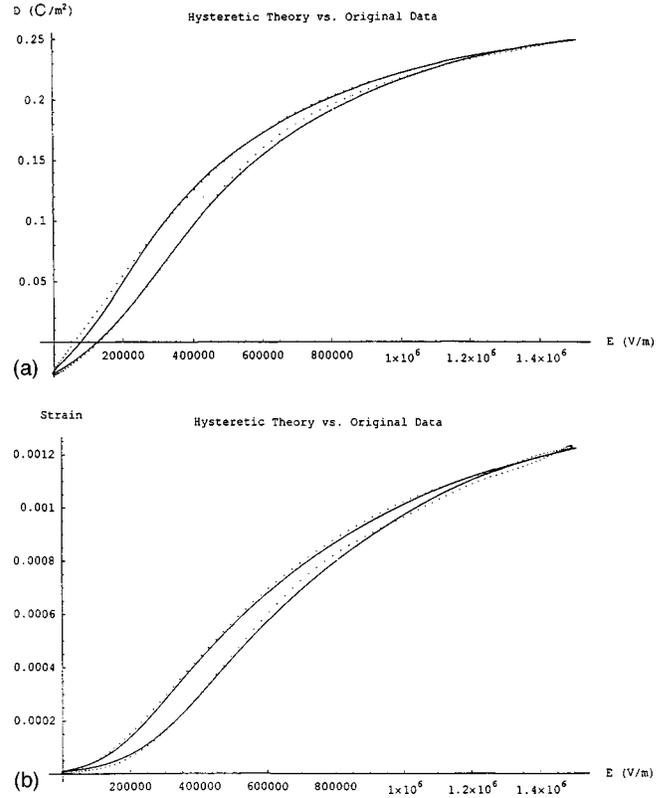


FIG. 8. (a) Hysteretic theory compared with original data for D after a 0.5-deg preprocessing rotation of the data about the E -axis in field space has been carried out. (b) Hysteretic theory compared with original data for S after a 0.5-deg preprocessing rotation of the data about the E -axis in field space has been carried out.

any specific anhysteretic theory, and could be applied with equal ease to any such theory. In summary, the rotations in field space of the D vs E data and the transformation of the S vs E data to anhysteretic form are operations directly performed on the experimental data, and are entirely independent of the anhysteretic theory of interest. Once a fit of a candidate anhysteretic theory to anhysteretic data is complete, the hysteretic version of that theory can be deduced by applying the inverse transformations. The field-space transformations described here for preprocessing the data are relevant to a classic problem in hysteresis measurements. As noted by Jaffe *et al.*,¹⁰ the presence of nonzero conductivity in a sample can lead to the observation of misshapen hysteresis loops. The experimental procedure suggested by Jaffe *et al.*, and used by most experimenters, to correct for conductivity of the sample (*viz.*, including an additional series or parallel resistance in the apparatus to produce a corrective phase shift) can produce measurement errors. As noted by Jaffe *et al.*, if the resistance of the sample is nonlinear, the problem is indeterminate. The preprocessing field-space rotations described here are an alternate way to correct the data for the twists introduced by sample resistance, and one application of this was shown here.

ACKNOWLEDGMENTS

We gratefully acknowledge the experimental group led by J. Powers, Code 2131, Naval Undersea Warfare Center, Newport, RI for graciously providing the experimental data

used herein. Responsibility for the interpretations of the data, however, rests solely with the authors. This work was supported by the Office of Naval Research.

APPENDIX A: MODIFIED ONE-DIMENSIONAL ANHYSTERETIC THEORY

In examining data newly available to us, we have seen behavior in the graphs of D vs E that is not well described by our earlier theory.¹ To accommodate this newly observed behavior, we propose modifying our original one-dimensional theory (for the D -field at zero stress) to

$$D = P_0 + \frac{1 + \alpha e^{-kE^2}}{(1 + \alpha)\sqrt{1 + \alpha E^2}} \epsilon_0 \epsilon E, \quad (\text{A1})$$

where α and k are new parameters. (The original theory is recovered if $\alpha=0$. Hence, this proposal is a perturbation of the earlier theory.) The exponential term is introduced to accommodate a change in slope of the D vs E curve, in the vicinity of the origin, which was not seen in previous data sets. (As noted elsewhere, however, Smith⁶ has observed a similar deviation in the data near the origin. The present form of Smith's theory does not accommodate that behavior.¹¹)

Denote the solution of Eq. (A1) for E by $E_{T=0}$, where T denotes applied stress. The modified anhyseretic theory is then given by the equations

$$E = E_{T=0} - 2\beta_2 TD, \quad (\text{A2})$$

and

$$S = \beta_1 T + \beta_2 D^2. \quad (\text{A3})$$

APPENDIX B: THE LOSS TANGENT AND THE REFERENCE ELLIPSE

Consider first the Fourier series representation of the polarization,

$$P = \sum_{n=1}^{n_P} P_n \sin(n\omega t + \phi_n^{(P)}). \quad (\text{B1})$$

The driving field is assumed to be multifrequency,

$$E = \sum_{n=1}^{n_E} E_n \sin(n\omega t + \phi_n^{(E)}). \quad (\text{B2})$$

In Eqs. (B1) and (B2), the quantities n_P and n_E have the obvious meanings of the number of spectral components in each field, and $\phi_n^{(E)}$ and $\phi_n^{(P)}$ are phase angles. The variables P and E are taken to include any constant component of the fields, so no explicit constant needs to be retained in the series.

It is well-known that the area of the P - E loop represents the loss per cycle per unit volume. We seek to compute this loop area in general, which we will denote by \mathcal{A} . Introduce the usual unit vectors \mathbf{i} , \mathbf{j} , \mathbf{k} , where \mathbf{i} , \mathbf{j} lie along the positive E , P axes, respectively, and $\mathbf{k} \equiv \mathbf{i} \times \mathbf{j}$. Define a vector \mathbf{C} such that $\mathbf{C} \equiv E\mathbf{j}$. Consider the integral of the curl of \mathbf{C} , evaluated in the P - E coordinate system and analyzed over the surface of the P - E loop, where the vector differential of

surface area is taken to be $d\mathbf{S} = dS \mathbf{k} = dE dP \mathbf{k}$. Since the curl of \mathbf{C} in this coordinate system is simply \mathbf{k} , we have

$$\iint \nabla \times \mathbf{C} \times d\mathbf{S} = \iint dS = \iint dE dP = \mathcal{A}, \quad (\text{B3})$$

where \mathcal{A} is the desired area of the loop. The first surface integral in Eq. (B3) can be re-expressed using Stokes' theorem, giving

$$\mathcal{A} = \oint \mathbf{C} \times d\mathbf{l}, \quad (\text{B4})$$

where the line integral is taken around the hysteresis loop in the counterclockwise sense. The vector differential line element is given by

$$d\mathbf{l} = dE \mathbf{i} + dP \mathbf{j}. \quad (\text{B5})$$

Since $\mathbf{C} \equiv E\mathbf{j}$, combining Eqs. (B4) and (B5) gives

$$\mathcal{A} = \oint E\mathbf{j} \times (dE \mathbf{i} + dP \mathbf{j}) = \oint E dP. \quad (\text{B6})$$

Introducing a new variable $\theta \equiv \omega t$ and combining Eqs. (B1), (B2), and (B6) produces

$$\begin{aligned} \mathcal{A} = & \int_0^{2\pi} \sum_{l=1}^{n_E} E_l \sin(l\theta + \phi_l^{(E)}) \\ & \times \frac{d}{d\theta} \left[\sum_{m=1}^{n_P} P_m \sin(m\theta + \phi_m^{(P)}) \right] d\theta. \end{aligned} \quad (\text{B7})$$

The integration range in Eq. (B7) corresponds to one cycle of the drive field, which is once around the loop. The integral of Eq. (B7) can be carried out directly, or using integration by parts. The result is

$$\mathcal{A} = \pi \sum_{n=1}^{n_E} n E_n P_n \sin(\phi_n^{(E)} - \phi_n^{(P)}). \quad (\text{B8})$$

In producing Eq. (B8), it has been assumed that $n_E \leq n_P$. Evaluating Eq. (B8) for the most common case, that in which the sample is driven by a monofrequency electric field (i.e., $n_E = 1$), gives the very simple expression

$$\mathcal{A} = -\pi P_1 E_1 \sin \phi_1^{(P)}, \quad (\text{B9})$$

where it has been assumed that $\phi_1^{(E)} = 0$. (The minus sign signifies that the polarization lags the electric field; hence $\phi_1^{(P)}$ is a negative angle. Taking $\phi_1^{(E)} = 0$ is equivalent to taking a particular definition of the zero of time.)

The area of the loop as given by Eq. (B9) is independent of any harmonics in the polarization except the first. We conclude that the second and higher harmonics of the polarization affect only the *shape* of the loop, not its area. (The derivation given is sufficiently general that the result applies to any non-self-crossing closed figure in the plane, provided the abscissa can be parametrized by a monofrequency sine.) This result, while it may be unexpected, is simply a consequence of the orthogonality of the higher harmonics of the polarization to the monofrequency driving electric field. (We also note that this is consistent with a similar result reported by Leary and Pilgrim.¹²) Therefore, to determine the loss per cycle per unit volume in a P - E hysteresis loop, it is suffi-

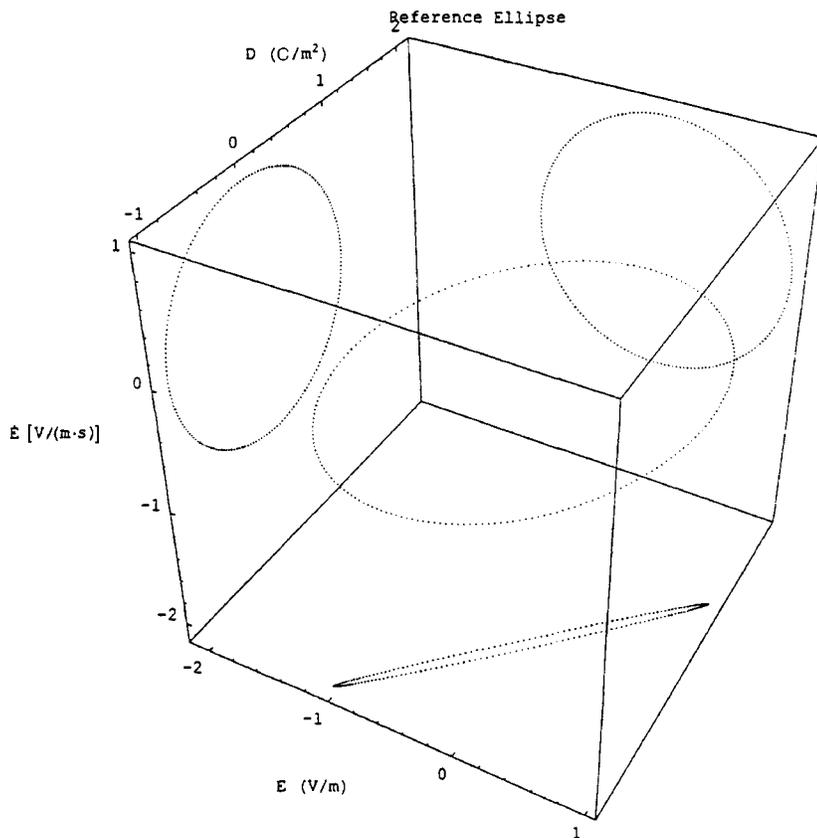


FIG. B1. Three-dimensional reference ellipse and its shadows in the coordinate planes. The shadow in the $D-E$ plane is the two-dimensional reference ellipse. Data are scaled to unity first harmonic.

cient to determine the relative phase angle between the electric field and the first harmonic polarization, as well as the first-harmonic amplitudes.

Since the area of the loop depends only upon the first harmonic of the polarization, at least for the case in which the sample is driven by a monofrequency sine, it is possible to redraw the $P-E$ hysteresis loop retaining only the first harmonic without changing the loop's area. Because the electric field and the first-harmonic polarization are both monofrequency sinusoids, the resulting figure in a plot of P vs E is an ellipse. We term this the reference ellipse. If this two-dimensional reference ellipse is inflated, a three-dimensional reference ellipse is obtained in field space. The three-dimensional reference ellipse resulting from the data of Fig. 1(c) is shown in Fig. B1, along with its shadows (projections) on each of the coordinate planes. The shadow in the $D-E$ plane is the two-dimensional reference ellipse. (In generating Fig. B1, and several other figures in this article, a form of "first-harmonic scaling" has been applied to the data. That is, the data are scaled such that the amplitudes of the first harmonic components of the data are equal to unity. This scaling is done in order that simplified analytical expressions, to be given presently, are applicable.)

The result given by Eq. (B9) can be used to understand the shapes assumed by $D-E$ hysteresis loops (and, of course, $M-H$ loops) in general. Referring once again to Fig. 1(a), notice that the loop shown exhibits essentially a "teardrop" shape, in which the area tends to be greatest near the origin and to decrease with increasing electric field. This (essentially) monotonic decrease in area away from the origin reflects the fact that the first harmonic in D is also de-

creasing. Since in the vicinity of the origin D is essentially linearly proportional to E , it follows that the first harmonic in the polarization is dominant near the origin. As the electric field increases, the first harmonic contribution to the polarization decreases while the second and higher harmonics initially increase, owing to the effects of saturation. At sufficiently large values of electric field, all harmonics are decreasing in amplitude as E increases. Thus, all harmonics but the first tend to initially grow with increasing E , then shrink with increasing E , while the first harmonic, and the loop area, both tend to decrease monotonically. In effect, there is little loop area in the regions away from the origin because there is little first harmonic content in the associated polarization.

The loss tangent $\tan \delta$ is also directly related to the phase angle ϕ_1 . (In what follows, we restrict our attention to the case of a monofrequency drive field. Hence, the shorthand notation ϕ_1 is used in place of $\phi_1^{(P)}$.) By taking the definition of the loss tangent as the ratio of the imaginary to the real part of the dielectric permittivity, and considering the connection between polarization P and electric displacement D , it is not too difficult to show that

$$\tan(\delta) = \tan(\phi_1). \quad (\text{B10})$$

The important properties of both the two-dimensional and three-dimensional versions of the reference ellipse can be determined analytically. The length of the semimajor axis of the two-dimensional reference ellipse is given by

$$a_{\text{axis}} = \sqrt{[E_1 \sin \theta_a]^2 + [P_1 \sin(\theta_a + \phi_1)]^2}, \quad (\text{B11})$$

where

$$\theta_a \equiv \cos^{-1} \left[-\sqrt{1/2 - \frac{E_1^2 + P_1^2 \cos(2\phi_1)}{2\sqrt{E_1^4 + P_1^4 + 2E_1^2 P_1^2 \cos(2\phi_1)}}} \right].$$

Here, and in what follows, ϕ_1 is assumed to lie in the interval $0 \geq \phi_1 \geq -\pi/2$. [If first-harmonic scaling is applied to the data, i.e., $E_1 = P_1 = 1$ (as in Figs. 2, 4, and B1), then $\theta_a = \pi/2 - \phi_1/2$, and $a_{\text{axis}} = \sqrt{2} \cos(\phi_1/2)$.]

The length of the semiminor axis of the two-dimensional reference ellipse is given by

$$b_{\text{axis}} = -\frac{P_1 E_1 \sin(\phi_1)}{a_{\text{axis}}}. \quad (\text{B12})$$

[First-harmonic data scaling gives $b_{\text{axis}} = -\sqrt{2} \sin(\phi_1/2)$.]

The angle between the major axis of the two-dimensional reference ellipse and the positive E -axis is

$$\psi_a = \tan^{-1} \left[\frac{P_1 \sin(\theta_a + \phi_1)}{E_1 \sin \theta_a} \right]. \quad (\text{B13})$$

(For first-harmonic data scaling, $\psi_a = \pi/4$.)

The angle between the plane of the three-dimensional reference ellipse and the \dot{E} axis is

$$\gamma_{ab} = \tan^{-1}(b_{\text{axis}}/\dot{E}_0), \quad (\text{B14})$$

where \dot{E}_0 denotes the amplitude of the time derivative of the electric field, and b_{axis} is given by Eq. (B12).

The reader should understand that Eqs. (B13) and (B14) also apply to Eq. (3). That is, these field-space rotation-angle formulas, derived from the reference ellipse in which only the first harmonic of the polarization is considered, are applicable to the original polarization data with all harmonics retained.

APPENDIX C: PREPROCESSING OF D DATA

A form of preprocessing is applied to the data. Our an hysteretic theory requires quadratic dependence of the strain upon the electric displacement D , as reflected in Eq. (A3). However, a direct application of Eq. (A8) to the measurements shows somewhat poor agreement. We find that a function of the form

$$\hat{S}' = \beta_2(D' - D_0)^2 \quad (\text{C1})$$

gives a much more satisfactory fit to the experimental data than does a direct fit to Eq. (A3). Here, D_0 is a constant determined in the fit and β_2 has the same meaning as in Eq. (A3).

Although the physical meaning of D_0 is unclear, we do not believe an equation of the form of Eq. (C1) would be a reasonable physical model, so we suspect the data require a constant shift by the amount D_0 . We believe Eq. (C1) is implausible as a physical model of strain because the subtraction of D_0 from D' in this equation suggests that the strain can somehow distinguish in its response between various sources of surface-charge density. That is, if it is indeed reasonable to suppose that the strain responds to the presence of a surface charge D [as is suggested by Eq. (A3)], it is physically unreasonable to then exclude some portion of the

surface charge. This exclusion of a portion of surface charge is effectively what is done through the subtraction of D_0 in Eq. (C1).

Consider the possibility that D_0 is the constant component of the D -field, i.e., the component arising from remanent polarization and bias voltage. In that case, the first-order Fourier series representation of D' can be written in the form

$$D' = D_0 + D_1 \sin(\omega t + \phi). \quad (\text{C2})$$

Here, D_1 is the first-harmonic amplitude of the D -field and ϕ is a phase angle. Consider a case in which the drive amplitude is sufficiently small that Eq. (C2) is a good approximation of the D -field, i.e., a case in which harmonics above the first are negligible. (This happens, for example, when an ordinary piezoceramic is driven with a modest voltage.) If the expression for D' given by Eq. (C2) is substituted into Eq. (C1), the quadratic expanded, and the result re-expressed in a harmonic series, no term in $\sin(\omega t)$ or $\cos(\omega t)$ will appear in the outcome. (The manner in which D_0 appears in these equations causes it to ‘subtract out,’ and this consequently results in the exclusion of $\sin(\omega t)$ and $\cos(\omega t)$ terms from the final expression.) But, the absence of any such first-harmonic terms in the strain would contradict the approximate linearity of strain that is known to result when either a large bias or a large remnant is present. This calculation further suggests that Eq. (C1) is an unlikely candidate for a physical law, and so we consider D_0 to be a constant shift required to correct the data. [Of course, one could still maintain that D_0 need not be the *total* constant component of the field, but simply some portion thereof, and could still maintain Eq. (C1) is a physical law. The demonstration given here does not exclude that possibility. However, such a view would again suggest that the strain exhibits a physically implausible distinction between a response to a bias and a response to a remnant.]

Prior to applying our theory (hysteretic or anhysteretic version), we first determine D_0 by fitting the data to Eq. (C1), then subtract the resulting value from the measured D -field to produce ‘shifted’ data. (It is this subtraction of D_0 that results in the thresholding process described in Sec. IV having no effect on the case considered. In that particular case D_0 is negative, so the preprocessing effectively results in an upward shift of the data; the shift is sufficiently large to avoid $D=0$.) However, in the plots that compare the theory with measurement, the fitted data and the theory are both shifted by D_0 in order that the data will appear in their original (unshifted) form.

¹J. C. Piquette and S. E. Forsythe, ‘‘A nonlinear material model of lead magnesium niobate (PMN),’’ *J. Acoust. Soc. Am.* **101**, 289–196 (1997).

²J. C. Piquette and S. E. Forsythe, ‘‘Generalized material model for lead magnesium niobate (PMN) and an associated electromechanical equivalent circuit,’’ *J. Acoust. Soc. Am.* **104**, 2763–2772 (1998).

³A. A. Adly and I. D. Mayergoz, ‘‘A new vector Preisach-type model of hysteresis,’’ *J. Appl. Phys.* **73**, 5824–5826 (1993).

⁴D. C. Jiles and D. L. Atherton, ‘‘Theory of ferromagnetic hysteresis,’’ *J. Magn. Magn. Mater.* **61**, 48–60 (1986).

⁵R. C. Smith and C. L. Hom, ‘‘A domain wall theory for ferroelectric hysteresis,’’ Center for Research in Scientific Computation Report Number CRSC-TR99-01 (January 1999).

⁶R. C. Smith, ‘‘Inverse compensation for hysteresis in magnetostrictive transducers,’’ Center for Research in Scientific Computation Report Number CRSC-TR98-36 (October 1998).

- ⁷K. Ghandi and N. W. Hagood IV, "Nonlinear modeling and characterization techniques for phase transitions in electro-mechanically coupled devices," Massachusetts Institute of Technology Report Number AMSL #98-1 (May 1998).
- ⁸F. T. Calkins, R. C. Smith, and A. B. Flatau, "An energy-based hysteresis model for magnetostrictive transducers," Institute for Computer Applications in Science and Engineering Report Number 97-60 (November 1997), p. 8.
- ⁹R. A. Guyer and P. A. Johnson, "Nonlinear mesoscopic activity: Evidence for a new class of materials," *Phys. Today* **52**, 30–36 (1999).
- ¹⁰B. Jaffe, W. R. Cooke, Jr., and H. Jaffe, *Piezoelectric Ceramics* (Academic, New York, 1971), pp. 37–47.
- ¹¹R. C. Smith, private communication, 1999.
- ¹²S. P. Leary and S. M. Pilgrim, "Harmonic analysis of the polarization response in PMN-based ceramics—A study in aging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**(1), 163–169 (1998).

One-dimensional phenomenological model of hysteresis.

II. Applications

Jean C. Piquette^{a)} and Stephen E. Forsythe

Naval Undersea Warfare Center, Code 2161, 1176 Howell Street, Newport, Rhode Island 02841

(Received 12 April 1999; accepted for publication 30 August 1999)

A model of hysteresis is applied to determine material response to multifrequency drives, and to the output control problem. Although as presented in Paper I the model is based on a monofrequency sinusoidal drive, it can readily be generalized. The generalization is based upon the fact, at least for quasistatic drives, that the shape of the hysteresis loop is independent of the shape of the drive waveform used to produce it provided that the drive is characterized by only one wave amplitude. The material response to a given arbitrarily shaped drive can be determined if the drive is first subdivided into single-amplitude regimes or epochs. Each such regime then has associated with it a unique hysteresis loop, which can be determined from the model. Each theoretical loop is generated using a monofrequency sinusoidal drive whose amplitude is equal to the single amplitude contained within the corresponding drive epoch. The material response is then determined by correlating the level of the given drive field (and the *sign* of its time derivative) with that of the sinusoidal drive used to generate the associated theoretical loop. The response to the arbitrary drive is taken to be equal to the response to the sinusoidal drive at the corresponding drive level and correspondingly signed time derivative. This process is capable of inversion. Thus, not only can the material response be determined for a drive of arbitrary waveshape, but also the drive waveshape required to produce a desired output trajectory can be determined. The procedure is illustrated by determining the drive necessary to produce a monofrequency sinusoidal magnetization response from a biased, prestressed sample of Terfenol D driven at high-amplitude magnetic field.

[S0001-4966(99)03812-6]

PACS numbers: 43.38.Ar, 43.20.Px, 43.30.Yj, 43.58.Vb [SLE]

INTRODUCTION

Our model of hysteresis¹ is generalized here to accommodate drives of arbitrary waveshape. The resulting model is applicable to magnetostrictive materials, such as Terfenol D, and to electrostrictive materials, such as lead magnesium niobate (PMN). (In Paper I, we focused on PMN. In the present part, we focus on Terfenol D. However, it should be understood that the results and methods are applicable to both.) The model of Ref. 1 is based upon incorporating the time rate of change of the drive field as a third coordinate, a process termed “inflation.” (The resulting three-dimensional coordinate system is termed “field space.”) The three-dimensional hysteresis “hoops” produced by the inflation process are collapsed to produce minimum-area two-dimensional loops (thus yielding approximately anhysteretic curves) by suitable rotations and a projection. This procedure permits the use of an anhysteretic theory² to analyze hysteretic data. By evaluating previously published data,³ it is demonstrated that minor hysteresis loops also can be satisfactorily subsumed within the model through the introduction of an additional field-space rotation and a translation in magnetization or polarization. The resulting minor-loop model is made predictive by the use of a few simple empirical relations. The model is also suitable for determining the waveshape of the drive that is needed to produce a desired output trajectory. Such a capability is desirable to permit application

of signal-processing techniques, which may require linearity in the output signal. This capability is demonstrated by determining the drive required to produce a monofrequency sinusoidal magnetization response in a prestressed sample of Terfenol D driven at high amplitude.

We note here that there are also other theories of hysteresis.⁴⁻⁷ These other theories can have difficulty in generating realistic (closed) curves when the sample is driven asymmetrically, e.g., under bias conditions such as typically would be needed to utilize PMN or Terfenol D in a practical transducer.^{8,9} One advantage of the present model is that it always produces closed loops, regardless of the drive conditions.

The concept of drive amplitude regimes, and their relation to minor hysteresis loops, is introduced in Sec. I. The empirical formulas needed to render the model predictive are provided in Sec. II. These formulas permit interpolation between and extrapolation from measured hysteresis loops. Section III describes the concept of “correlation” of a drive of arbitrary waveshape with a sinusoidal drive (or drives). The notion of correlation is developed by calculating the magnetization and strain that result from applying a biased, increasing-amplitude, triangular-wave magnetic field to a prestressed Terfenol D sample. The resulting predictions are compared with experiment. Section IV considers the problem of determining the drive waveshape required to produce a monofrequency magnetization output trajectory. The resulting drive is checked theoretically using a model based on experiment. The conclusion is given in Sec. V.

^{a)}Electronic mail: piquettejc@npt.nuwc.navy.mil

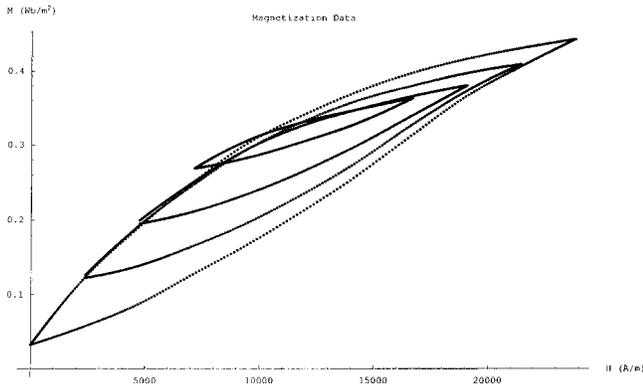


FIG. 1. Magnetization curves from the experiment reported in Ref. 3. The sample is Terfenol D, prestressed to 6.9 kPa. Each nested loop was produced by biasing the sample to the fixed field $H=11.9$ kA/m, but varying the alternating part of the drive. The largest hysteresis loop, herein termed the “semimajor” loop, was acquired by applying an alternating magnetic field of amplitude equal to the bias.

I. MINOR LOOPS AND AMPLITUDE REGIMES

In Fig. 1 an experimental³ sequence of nested minor magnetization hysteresis loops acquired from a prestressed sample of Terfenol D is shown. Each loop was acquired at the same biasing magnetic field and prestress, and thus the loops differ only in the amplitude of the varying portion of the applied magnetic field. (Although several other biases, drives, and prestresses were investigated in the research reported in Ref. 3, we selected the particular case depicted in Fig. 1 for description here because it exhibits the greatest relative amount of hysteresis, and hence is the most stringent test of the model. Specifically, the largest loop shown has associated with it a hysteretic loss of about 25% per cycle. That is, the magnitude of the loss tangent is 0.256 for this loop.) Figure 1 shows that the material response naturally divides into regimes (or epochs) determined by the amplitudes contained within the drive. As can be seen, each such amplitude regime has associated with it a unique hysteresis loop.

Also apparent from comparing the loops of Fig. 1 with each other is that as the driving magnetic field is reduced in amplitude, the resulting magnetization loop displays a rotation and a translation with respect to loops associated with higher-amplitude drives. This rotation and translation are even more apparent if the sequence of field-space rotations¹ required to collapse the area of each of the experimental loops is applied. The result of applying this sequence of rotations to each of the loops of Fig. 1 is shown in Fig. 2.

The curves of Fig. 2 are suggestive of what might be done to produce a minor-loop predictive capability. In particular, if an appropriate rotation and translation are applied to each collapsed-area minor loop shown, it can be made to overlay the collapsed-area semimajor loop. (The term “semimajor loop” is used here to denote the loop obtained from driving the sample with a field of alternating amplitude equal to the bias, so that the net applied field just barely becomes zero at one point in each cycle. The semimajor loop for the experiment currently under discussion is the largest of the loops depicted in Fig. 1. The term is introduced to distinguish it from the term “major loop,” which applies when no

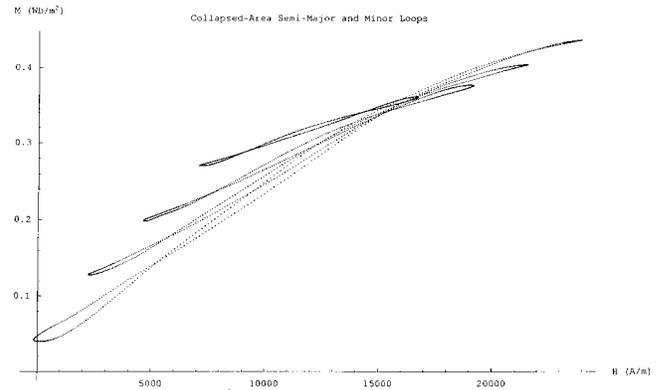


FIG. 2. Collapsed-area hysteresis loops that result from the application of Eq. (1) sequentially to each of the loops of Fig. 1. For a complete description of the procedure, see Ref. 1.

bias is used.) The result of applying such a rotation-and-shift transformation to the smallest collapsed-area minor loop of Fig. 2 is shown in Fig. 3, along with the (unmodified) collapsed-area semimajor loop. As can be seen, the transformed collapsed-area minor loop well overlays the collapsed-area semimajor loop.

This procedure suggests that the minor loop might be predicted from the anhysteretic theory by applying the inverse of the transformations that produce the overlay seen in Fig. 3. The results of this procedure, applied to theoretically generate the smallest minor loop, are shown in Fig. 4(a) (magnetization) and (b) (strain). (In these figures, the semimajor loop data are also presented for reference.) The theoretical (solid line) minor loops shown in these figures were produced by first evaluating the anhysteretic theory at the drive level present in the experimental minor-loop data, then applying the inverse of the transformations that produce the overlay seen in Fig. 3. As can be seen, the resulting theoretical minor loops are in reasonably good agreement with the measured minor loops.

II. INTERPOLATION AND EXTRAPOLATION

While the process described above for deducing minor loops is straightforward, several details must be considered

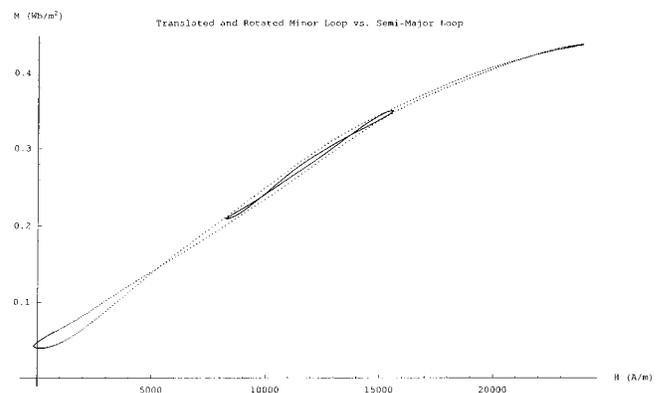


FIG. 3. Overlay of smallest collapsed-area minor loop onto the collapsed-area semimajor loop. This overlay results from applying the transformation summarized by Eq. (6) to the smallest area minor loop and to the semimajor loop depicted in Fig. 1.

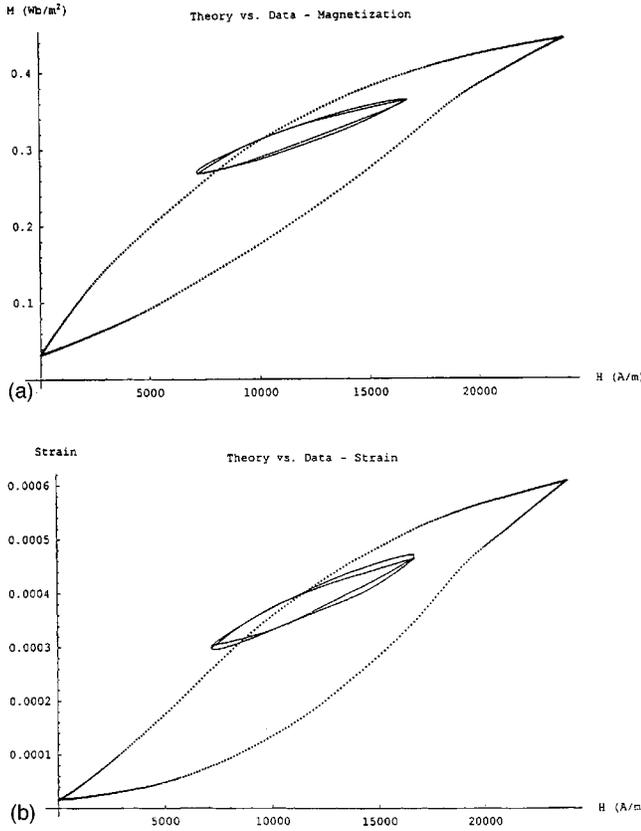


FIG. 4. Comparison of minor-loop model with data. The minor-loop model results from applying the inverses of the transformations embodied in Eq. (6) to the anhysteretic theoretical prediction resulting from a drive level appropriate to the minor loop of interest. (a) Magnetization; (b) strain.

in order to render the model predictive. Referring to Fig. 2, it is clear that each of the depicted collapsed-area loops would require a different field-space rotation and translation in order to achieve overlay of the kind seen in Fig. 3. Also, each of the minor loops requires its own set of field-space rotations to produce the requisite area collapse. To render the model predictive, it is necessary to be able to calculate each required set of rotations as a function of (arbitrarily specified) drive-field amplitude.

Applying the rotations specified by the equation

$$\begin{pmatrix} H' \\ M' \\ \dot{H}' \end{pmatrix} = \Psi_a^{-1} \Xi_D \Gamma_{ab} \Psi_a \begin{pmatrix} H \\ M \\ \dot{H} \end{pmatrix} \quad (1)$$

to each data set (where H and \dot{H} are the applied magnetic field and its time derivative, respectively, and M is the resulting magnetization) and projecting into the H - M plane produces each collapsed-area loop.¹ [It is to be understood that the quantities H , M , and \dot{H} of Eq. (1) are replaced with the quantities E , P , and \dot{E} , respectively, if the material of interest is electrostrictive instead of magnetostrictive. It is also to be understood that the transformation of Eq. (1) assumes that the data have been “correlated” with sine wave drives if the actual drives used to acquire the data are non-sinusoidal. This correlation process is detailed in Sec. III. For now, the reader may simply assume the drive field is sinusoidal, but should bear in mind that this restriction is

subsequently relaxed. Finally, the reader should realize that the transformation of Eq. (1) assumes the “first-harmonic scaling” described in Ref. 1 has been applied. In addition to scaling each data point by the amplitude of the first harmonic component of the data, this procedure entails subtracting out any constant component of the data.] In Eq. (1), the symbols Ξ_D , Γ_{ab} , Ψ_a , and Ψ_a^{-1} denote rotation matrices, which implement coordinate-axis rotations through angles ξ_D , γ_{ab} , ψ_a , and $-\psi_a$, respectively. See Ref. 1 for a detailed description of this transformation.

While the rotation angles ξ_D and ψ_a are independent of the drive amplitude, the rotation angle γ_{ab} varies with drive level, and thus must be determined for each drive of interest. Its value is predicted by the empirical equation

$$\gamma_{ab} = g^2 \gamma_{ABS} + \gamma_{REF}, \quad (2)$$

where

$$g \equiv (H_{bias} - H_{drive}) / H_{bias}, \quad H_{bias} = \text{bias field}, \quad (3)$$

and

$$H_{drive} = \text{drive field}, \quad 0 \leq H_{drive} \leq H_{bias}.$$

The term “drive field” refers to the alternating portion of the applied field. As can be seen, the smallest allowable value of the drive field H_{drive} is presumed to be zero, and the largest allowable value H_{bias} . As H_{drive} varies between its minimum and maximum values, g varies from 0 to 1. The quantity γ_{REF} is the “reference” value of the γ_{ab} field-space rotation angle, and is the angle used to collapse the semimajor loop to minimum area. (For the discussion that follows, it is helpful to note that $g=0$ when $H_{drive}=H_{bias}$ so that $g=0$ may be considered a condition to obtain the semimajor loop.) The quantity γ_{ABS} denotes the absolute largest correction to γ_{REF} required owing to the effects of decreasing drive amplitude. That is, $\gamma_{ab} \rightarrow \gamma_{REF} + \gamma_{ABS}$ as $H_{drive} \rightarrow 0$. The parameters γ_{REF} and γ_{ABS} as well as the fixed rotation angle ξ_D , are determined from experimental data. (As detailed in Ref. 1, the fixed rotation angle ψ_a is determined theoretically.)

Once a collapsed-area loop is produced through the use of Eq. (1), it is then necessary to generate a further rotation that aligns the collapsed minor loop with the collapsed semimajor loop, and to apply a translation or “shift” that produces the final overlay. Let the required rotation be performed by the matrix Θ , with the associated rotation angle¹⁰ denoted by ϑ . The angle ϑ required in the Θ matrix is predicted by the (empirical) equation

$$\vartheta = g \vartheta_{ABS}, \quad (4)$$

where g is again given by Eq. (3). It should be understood that the matrix Θ is applied by premultiplying the matrices appearing on the right-hand side of Eq. (1). The experimentally determined constant ϑ_{ABS} is the absolute largest additional field-space rotation angle required owing to the effects of decreasing drive amplitude.

Overlay of the minor and semimajor loops also requires a translation in addition to the rotation by the angle specified by Eq. (4). This translation is of an amount determined by

$$M_{SHIFT} = \sqrt{g} M_{ABS}. \quad (5)$$

The notation M_{ABS} refers to the absolute largest possible translation in magnetization required owing to the effects of decreasing drive amplitude; i.e., $M_{\text{SHIFT}} \rightarrow M_{\text{ABS}}$ as $H_{\text{drive}} \rightarrow 0$. Again, M_{ABS} is a constant determined from experiment. Owing to the appearance of the quantity g under a square root in Eq. (5), that formula does not apply should the drive level H_{drive} exceed the bias level H_{bias} . Thus, the present formulas may be used to interpolate between measured drive levels or to extrapolate below the lowest measured drive level, but may not be used to extrapolate responses for drives that exceed the bias.

In summary, the transformations that produce the collapsed-area minor loops that overlay the semimajor loop can be reduced to the single equation

$$\begin{pmatrix} H' \\ M' \\ \dot{H}' \end{pmatrix} = R \begin{pmatrix} H \\ M \\ \dot{H} \end{pmatrix} + \begin{pmatrix} 0 \\ M_{\text{SHIFT}} \\ 0 \end{pmatrix}. \quad (6)$$

Here, R is an overall rotation matrix defined by $R \equiv \Theta \Psi_a^{-1} \Xi_D \Gamma_{ab} \Psi_a$, in which the last four terms yield the loop-area collapse, and the matrix Θ produces the rotation by ϑ . The last term of Eq. (6) yields the translation that produces the final overlay of the semimajor loop. [As with the transformations presented in Ref. 1, any constant components in H or M are to be subtracted prior to application of Eq. (6).]

Finally, we remark that Eq. (2) can be combined with the results of Appendix B of Ref. 1 to produce a predictive formula for the loss tangent as a function of ac drive level. Using the results presented in that Appendix for the case of first-harmonic scaling, it is not difficult to show that

$$\tan \gamma_{ab} = -\sqrt{2} \sin \phi_1/2, \quad (7)$$

where ϕ_1 is the phase angle between the first-harmonic magnetization and the driving magnetic field. It is also shown in Ref. 1 that $\tan \delta = \tan \phi_1$. Combining this result with the solution to Eq. (7) for ϕ_1 gives

$$\tan \delta = \tan\{2 \arccos[\tan(\gamma_{ab})/\sqrt{2}]\}. \quad (8)$$

Here, γ_{ab} is evaluated using Eq. (2). Equation (8) provides a prediction of the variation of the loss tangent with ac drive level. (It has been noted elsewhere¹¹ that the loss tangent does indeed depend upon the drive level.) In the scaled coordinates for which Eq. (8) is valid, the quantity $\tan(\gamma_{ab})$ can range over the extreme values 0 to 1, as can be seen from Eq. (7) if it is recalled that these formulas assume $0 \leq \phi_1 \leq -\pi/2$. The resulting extreme values of the loss tangent range over the interval 0 to (minus) infinity, respectively.

III. "CORRELATION" OF AN ARBITRARY DRIVE WITH SINUSOIDAL DRIVES

In our model of hysteresis¹ the loops are generated using monofrequency sinusoidal drives. However, the response of a sample to a drive of arbitrary waveshape, at least for quasistatic conditions, can be deduced by "correlating" the level of the given drive field (and the *sign* of its time derivative) with that of an appropriately selected sinusoidal drive (or drives). This process of correlating a given drive with a

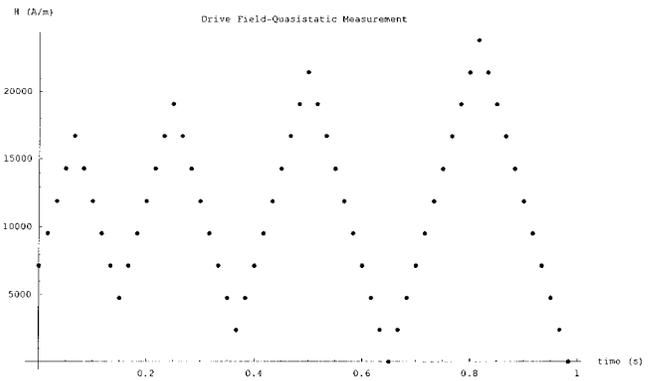


FIG. 5. Sequence of biased triangular waves applied to the Terfenol D sample in the experiment of Ref. 3. Prestress is 6.9 kPa. Biasing magnetic field is 11.9 kA/m. This drive produced the sequence of experimental magnetization responses seen in Fig. 1.

sinusoidal drive is required for Eqs. (1) and (6) to be applicable. We develop the notion of correlation through the presentation of a specific example.

Consider the biased, increasing-amplitude, triangular-wave drive depicted in Fig. 5. (The drive of Fig. 5 produced the experimental magnetization response seen in Fig. 1. These two figures clearly illustrate the amplitude regimes that form the basis of the present extension of our model. That is, it is clear that each amplitude regime seen in Fig. 5 produces a corresponding magnetization minor loop seen in Fig. 1. We also note that similar behavior is seen geologically in the response of rocks to stresses.¹² The behavior seen in Fig. 1 here in response to the drive of Fig. 5 is completely consistent with the behavior described in Ref. 12. The model described here is thus consistent with many varied "memory" phenomena.) Here, the sample is biased to a magnetic field of magnitude 11 900 Amp/m.

In order to deduce sample response, the given drive is first subdivided into regimes or epochs. A single amplitude, or excursion away from the bias level, characterizes each such regime or epoch. For the depicted drive, the first such epoch runs from $t=0$ to about $t=0.125$ s. During this interval the magnetic field rises from a low level of roughly 6900 Amp/m at $t=0$ up to a maximum level of approximately 16 900 Amp/m at around $t=0.0625$ s, then back down to roughly 6900 Amp/m at around $t=0.125$ s. Thus, the single amplitude 5000 Amp/m $[(16\,900 \text{ Amp/m} - 6\,900 \text{ Amp/m}) \div 2]$ characterizes the drive in this interval. This means that one theoretical hysteresis loop is sufficient to deduce the response in this particular amplitude regime. It is apparent from examining Fig. 5 that four amplitude regimes or epochs are required to characterize the entire drive. Hence, four unique theoretical hysteresis loops must be generated from the model in order to predict sample response in this case. (Again cf. Fig. 1. Although the experimental loops presented in Fig. 1 were obtained in response to the triangular-wave drive of Fig. 5, the *same* set of loops results from the corresponding set of theoretical *sine* wave drives. That is, the same set of loops will result from sine waves having the same amplitudes as the triangular waves, owing to the independence of the loop shapes from drive waveshape.) The required loops are completely determined from the formulas

of Sec. II when evaluated for the amplitude appropriate to each drive epoch.

The response to the given triangular-wave drive can be deduced by correlating the given drive with the appropriate sinusoidal drive. The sinusoidal drive used in the correlation process is taken to have the same amplitude as the given drive during the drive regime of interest. (Thus, for the time interval $t=0$ to $t=0.125$ s of the present example, the sine wave is taken to have an amplitude of 5000 Amp/m.) What must be correlated are the field level of the given drive and that of the sine drive. Of course, the sine drive typically has two points at which a given field level is achieved. The point on the sine drive selected for correlation with the given drive is that point which has the same sign time derivative as that in the given drive. The *response* to the given drive is taken to be equal to that resulting from the chosen point on the sine drive.

For example, consider the drive of Fig. 5 during the interval $t=0$ to $t=0.0625$ s. In this interval, as pointed out above, the triangular-wave drive has an amplitude of 5000 Amp/m. Thus, a theoretical hysteresis loop based on a sinusoidal drive of this amplitude is generated. The transformation specified by Eq. (6), with the relevant parameters suitably evaluated for this amplitude using the empirical formulas given in Sec. II, is employed in order to produce the proper orientation and positioning of the loop. Notice that during the time interval currently being considered, the given drive is increasing in level, i.e., $dH/dt > 0$. The theoretical hysteresis loop is thus examined over the region for which $dH/dt > 0$. (For the interval $t=0.0625$ s to $t=0.125$ s it is seen that $dH/dt < 0$, so for this interval correlation is done with the region of the theoretical loop for which $dH/dt < 0$.) The response to the given drive is then taken to be the theoretical response to the sinusoidal drive *at the corresponding level of drive field and correspondingly signed time derivative*. Since the theoretical drive is computed using a sine wave having the same amplitude as the given drive in the regime of interest, one is assured that a corresponding theoretical response is available. The correlation process allows for the fact that the time at which the desired field level occurs in the model is different from the time at which it occurs in the given drive. Nonetheless, a simple interpolation of the theoretical response yields the desired response.

The results of applying these ideas to the entire drive of Fig. 5 are shown in Fig. 6(a) (magnetization) and (b) (strain).¹³ The solid lines are theory, the dots experiment. As can be seen, agreement is quite good. These figures also illustrate the success of the interpolation formulas of the preceding section. The numerical parameters required in those formulas were deduced by utilizing only the data from the first and last available amplitude epochs. (That is, only the data of the semimajor loop and of the smallest minor loop were used in determining the numerical parameters required.) Hence, the curves presented for the two intermediate epochs (corresponding approximately to the time interval $t=0.125$ s to $t=0.65$ s) were deduced from hysteresis loops generated using the given empirical interpolation formulas.¹⁴

It is also interesting to evaluate Eq. (8) for this example, using Eq. (2) to evaluate γ_{ab} . The fitted parameters have the

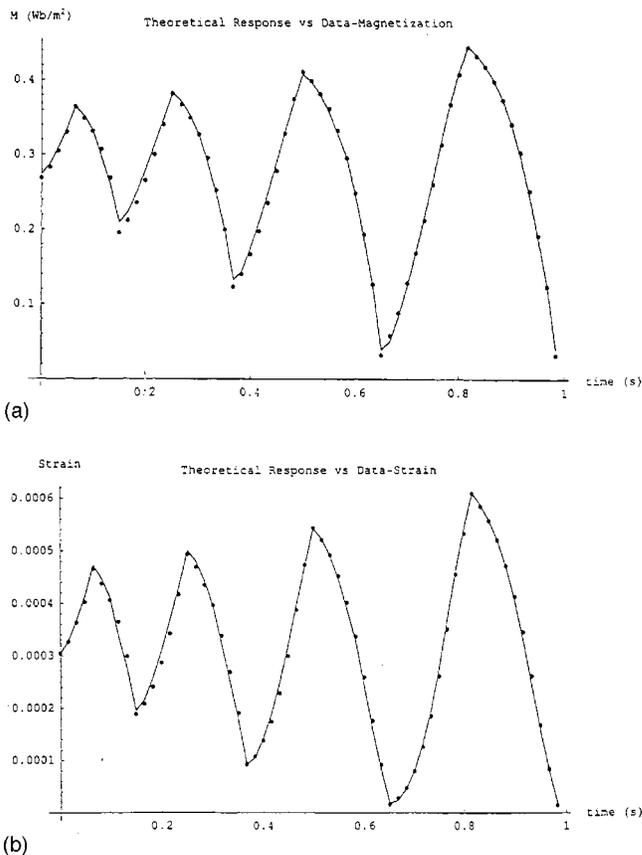


FIG. 6. Comparison with data of the theoretical response that results from the correlation process applied to the drive of Fig. 5. (a) Magnetization; (b) strain.

values $\gamma_{REF}=0.176$ and $\gamma_{ABS}=-0.0135$. Using these values, Eq. (8) suggests that $\tan \delta$ varies between about -0.258 (for the case of the drive equal to the bias) down to about -0.237 at zero drive. This is a (predicted) reduction in the loss tangent of some 8%. (The reader should not be troubled by the negative value of the loss tangent. Whether this quantity is reported as positive or negative depends essentially on the chosen convention. We choose to include the minus sign that occurs in the complex dielectric permittivity as part of the definition,¹⁵ while other authors often exclude it.)

IV. WAVEFORM INVERSION TO PRODUCE A DESIRED OUTPUT TRAJECTORY

The correlation procedure described in the previous section can also be inverted to determine the drive waveshape required to produce a desired output trajectory. As an example, we consider the production of a monofrequency sine-wave magnetization output at high drive amplitude. The sample material, prestress, and bias conditions are all the same as in the example considered in the preceding sections.

In Fig. 7(a) the theoretical magnetization time-waveform response to a high-amplitude sine wave drive is shown. [The experiment of Ref. 3 utilized triangular-wave drives of the type depicted in Fig. 5, and a sample model was determined from these measurements. The theoretical response to a sine-wave drive depicted in Fig. 7(a) was deduced from this model.] Here, the drive amplitude is equal to the bias; hence,

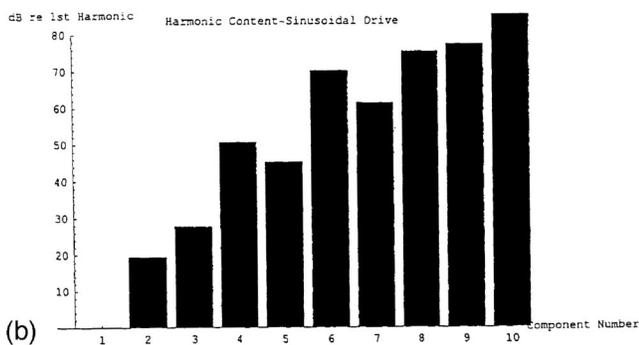
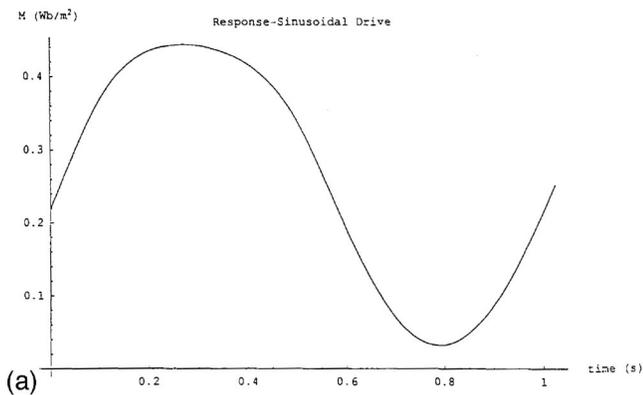


FIG. 7. Theoretical response of the Terfenol D sample to a sinusoidal drive of high amplitude. (a) Time-domain magnetization. (b) Frequency-domain magnetization, expressed in dB *re* the first-harmonic amplitude.

the prediction is based on the semimajor loop. The resulting deviation from a monofrequency response is evident. The spectrum of the response is given in Fig. 7(b). Spectral amplitudes are expressed in dB relative to the first harmonic.

The drive waveform deduced to produce a monofrequency sinusoidal response by inverting the correlation process is shown in Fig. 8. Although a single amplitude characterizes the desired response, *two* amplitudes are required to characterize the drive. This occurs due to the nonlinearity and hysteresis of the sample. The time-domain magnetization predicted to result from the drive of Fig. 8 is shown in Fig. 9(a). (This prediction was once again generated using the experimental sample model.) The spectrum of the pre-

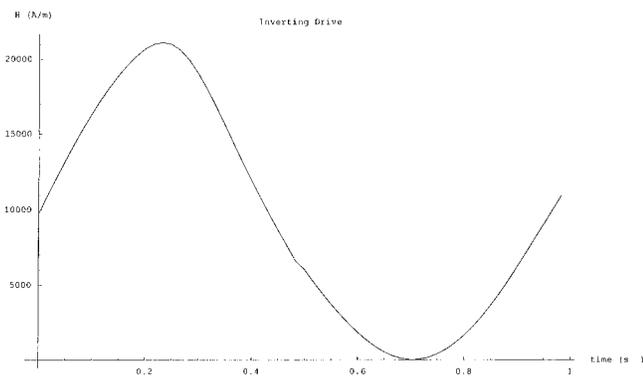


FIG. 8. Theoretical driving magnetic field required to produce a monofrequency sinusoidal output of approximately the same level as that seen in Fig. 7.

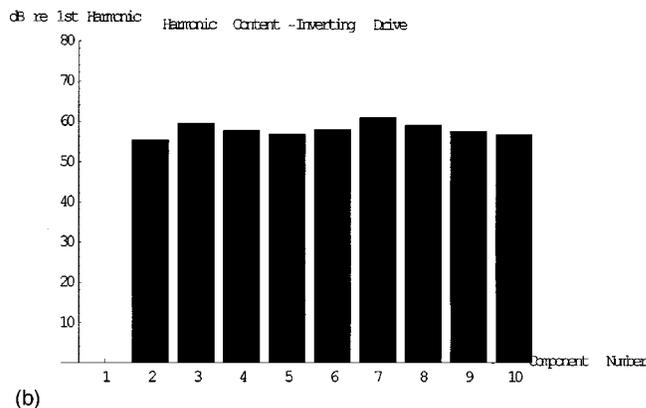
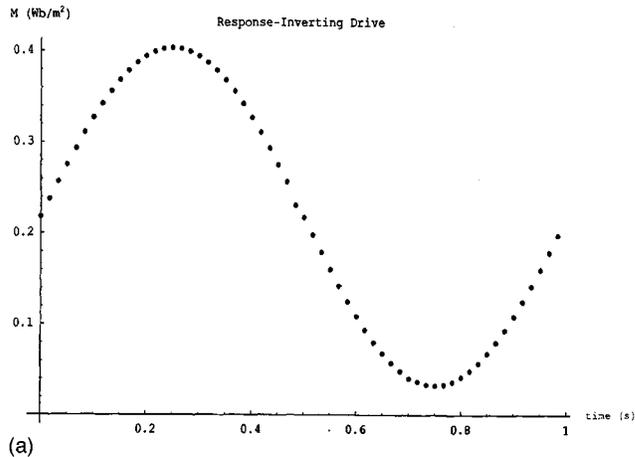


FIG. 9. Theoretical response to the drive of Fig. 8, based on the experimental sample model. (a) Time-domain magnetization. (b) Frequency-domain magnetization, expressed in dB *re* the first harmonic amplitude.

dicted response is given in Fig. 9(b). As can be seen by comparing Fig. 9(b) with Fig. 7(b) a significant reduction in harmonic distortion is predicted. The second-harmonic distortion of about 19 dB seen in Fig. 7(b) is reduced to about 55 dB in Fig. 9(b). The corresponding values for the third harmonic are 28 and 60 dB, respectively.

V. CONCLUSION

Our model of hysteresis has been generalized to accommodate drives of arbitrary waveshape for quasistatic conditions. This is done by dividing a given drive into epochs, each of which can be characterized by one amplitude. Each such epoch has associated with it a unique hysteresis loop. The characteristics of these loops are determined using parameters determined from simple empirical formulas. These empirical formulas render the model predictive. The ability to handle arbitrarily shaped drives was demonstrated by considering a drive consisting of a biased triangular wave of increasing amplitude.

The model can be inverted so that the drive required to produce a desired trajectory can be deduced. The procedure was demonstrated analytically by determining the drive required to produce a monofrequency magnetization curve from a prestressed sample of Terfenol D driven at high amplitude. A significant reduction in harmonic distortion is predicted to occur based on the model.

The model described herein is strictly applicable only to quasistatic drives. Its extension to drives of a more general character is the subject of further research.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Mark Moffett for providing for our analysis the data acquired in the experiment described in Ref. 3. We also acknowledge the suggestion of Dr. Harold Robinson that our model of electrostrictive hysteresis might be applicable to magnetostrictive materials. The Office of Naval Research supported this work.

- ¹J. C. Piquette and S. E. Forsythe, "One-dimensional phenomenological model of hysteresis. I. Development of the model," *J. Acoust. Soc. Am.* **106**, 3317–3327 (1999).
- ²J. C. Piquette and S. E. Forsythe, "A nonlinear material model of lead magnesium niobate (PMN)," *J. Acoust. Soc. Am.* **101**, 289–196 (1997).
- ³M. B. Moffett, A. E. Clark, M. Wun-Fogle, J. F. Lindberg, J. P. Teter, and E. A. McLaughlin, "Characterization of Terfenol-D for magnetostrictive transducers," *J. Acoust. Soc. Am.* **89**, 1448–1455 (1991).
- ⁴D. C. Jiles and D. L. Atherton, "Theory of ferromagnetic hysteresis," *J. Magn. Magn. Mater.* **61**, 48–60 (1986).
- ⁵R. C. Smith and C. L. Hom, "A domain wall theory for ferroelectric hysteresis," Center for Research in Scientific Computation Report Number CRSC-TR99-01 (January 1999).
- ⁶R. C. Smith, "Inverse compensation for hysteresis in magnetostrictive transducers," Center for Research in Scientific Computation Report Number CRSC-TR98-36 (October 1998).
- ⁷K. Ghandi and N. W. Hagood IV, "Nonlinear modeling and characterization techniques for phase transitions in electro-mechanically coupled devices," Massachusetts Institute of Technology Report Number AMSI #98-1 (May 1998).
- ⁸F. T. Calkins, R. C. Smith, and A. B. Flatau, "An energy-based hysteresis

model for magnetostrictive transducers," Institute for Computer Applications in Science and Engineering Report Number 97-60 (November 1997), p. 8.

- ⁹See also Ref. 5, p. 24.
- ¹⁰This rotation is carried out about an axis parallel to the \dot{H} -axis. The coordinates of the point in the H - M plane through which this axis passes are the constant values in the H and M fields, respectively. The constant component of H is simply the applied bias. The constant component of M is determined from the empirical formula $M_{\text{constant}} = M_{\text{semimajor}} + gM_{\text{minor}}$. In this equation, the quantity $M_{\text{semimajor}}$ is the constant component of M in the semimajor loop, and M_{minor} the maximum correction owing to the effects of decreasing drive amplitude. The quantity g is given by Eq. (3).
- ¹¹D. A. Berlincourt, D. R. Curran, and H. Jaffe, "Piezoelectric and piezomagnetic materials and their function in transducers," in *Physical Acoustics, I (Part A)*, edited by W. P. Mason (Academic, New York, 1964), p. 208.
- ¹²R. A. Guyer and P. A. Johnson, "Nonlinear mesoscopic activity: Evidence for a new class of materials," *Phys. Today* **52**, 30–36 (1999).
- ¹³See Ref. 1 for a discussion of how the strain curve is computed from the magnetization curve. (Although the model described there is developed for electrostrictive materials, the obvious correspondences render it suitable for magnetostrictive materials.) It was found to be necessary to extend the strain equation to the fourth-power term to accommodate the data of Ref. 3. A suitable corresponding term must also be added to the magnetization equation.
- ¹⁴This same case has also been used to check the extrapolation capability implicit in the formulas of Sec. II. This was done by deducing the numerical parameters required in those formulas using the data from the second- and fourth-amplitude epochs of Fig. 5, rather than from the first and fourth epochs as described in the text. The process of computing sample response for the first epoch then becomes an extrapolation rather than an interpolation. The agreement between theory and data seen in Fig. 6(a) and (b) is also typical of that seen for extrapolation.
- ¹⁵See, for example, *Dielectric Materials and Applications*, edited by R. Von Hippel (M.I.T. Press, Cambridge, MA, 1954), Eq. (1.9), p. 4.

A basic concept of direct converting digital microphone

Yoshinobu Yasuno^{a)}

Matsushita Communication Industrial Co., Ltd., 600 Saedo, Tsuzukiku, Yokohama, 224-8539 Japan

Yasuhiro Riko

Riko Associates, 1988-34 Kitahassaku, Midoriku, 226-0021 Japan

(Received 26 December 1998; accepted for publication 10 July 1999)

An electroacoustic system which directly converts analog acoustic signals to digital electric signals is described. The system consists of a subtractor, a sampling and holding circuit, a sigma-delta modulator as a comparator, an accumulator, and a local direct digital-to-analog converting transducer similar to a typical electronic analog-to-digital converter. The subtractor is an electrostatic device which has a diaphragm, driving electrodes, and a detecting electrode. The surface area of the driving electrodes corresponds to the significant bits in the digital signal, as an electroacoustic digital-to-analog converter. The detecting electrode produces an electrical signal proportional to the displacement of the diaphragm driven by subtracting the received acoustic signal from the electrostatically driven force. This is regarded as a subtractor. The detected signal is amplified and sampled-held and modulated by the sigma-delta procedure and generates a signal of ± 1 bit, which is added to the accumulator memory by a high clock frequency. The output of the accumulator is the digital signal output and is also fed to the driving electrodes. A 4-bit conceptual system was developed to affirm this concept. © 1999 Acoustical Society of America. [S0001-4966(99)07412-3]

PACS numbers: 43.38.Bs [SLE]

INTRODUCTION

In human acoustic-machine interfaces, an analog-to-digital converter (ADC) and digital-to-analog converter (DAC) are often employed. The traditional digital electric-to-analog acoustic signal interface is a DAC and a transducer (loudspeaker or receiver). This interface has been proposed as a device using an integrated DAC and a direct digital-to-analog converting receiver by Flanagan¹ in 1980. Another essential interface is the conversion from analog acoustic to digital electric signal which requires an acoustoelectric transducer (microphone) and an ADC. It is rather difficult to integrate this into one device.

This paper describes an experimental model of an integrated device which has the functions of an ADC and DAC, and electroacoustic transducer.² Because the direct-converting DAC has already been described by Flanagan, this paper concentrates on a direct-converting ADC. The concept is based on a replacement of the local DAC from electronic form to a newly developed electroacoustic device. Some experimental models have been fabricated to explore future possibilities.

I. SYSTEM DESCRIPTION

The fundamental concept of this direct-converting digital microphone is to replace a portion of the conventional electronic ADC. Figure 1 shows the diagram of a differential-type electronic ADC. The analog signal is fed to the input terminal.

The signal is subtracted from another analog signal which is generated in the local DAC, controlled by the ana-

log output y . The analog signal y is the result of the conversion of the digital signal $[q_0q_1q_2q_3, \dots]$ at each clock sequence as expressed by Eq. (1).

$$y = s(q_02^0 + q_12^1 + q_22^2 + \dots), \quad (1)$$

where y is the discrete analog value, s is the sign (± 1) and q_0, q_1, q_2, \dots are binary digits (0 or 1).

Equation (1) shows that the converted discrete analog signal y is the summation of each bit weighted by 2^n , where n is 0, 1, 2, ... according to the significant place.

The subtracted result generates a ± 1 -bit pulse from the comparison with a ternary reference level, at the timing of the internal clock. The $+1$ -bit is generated when this result exceeds the positive reference level and the -1 bit for exceeding the negative reference level. This ± 1 -bit signal is counted in the accumulator and creates a numerical value by accumulation, which is the raw digital output signal. The nominal digital output is sampled by the system interface clock and this digital output signal is fed to the local DAC and is converted to analog form for the subtraction signal at the subtractor. The feedback loop is completed resulting in a type of differential electronic ADC.

The fundamental concept of the direct-converting digital microphone described here is the same type of electronic ADC. Figure 2 shows the diagram of this direct-converting digital microphone. The subtractor and local DAC as illustrated in Fig. 1 is replaced by an electroacoustic transducing device. The acoustic analog input signal is applied across the diaphragm of the transducer, which is also driven by the DAC driver electrostatically. Thus, the electroacoustic transducer acts as a subtractor and also as a local DAC. The electric output of the transducer which represents the displacement of the diaphragm is modulated by sigma-delta

^{a)}Electronic mail: yyasuno@avsd.mci.mei.co.jp

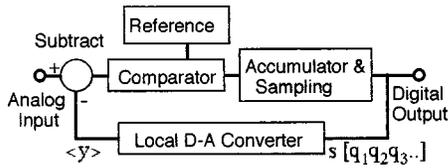


FIG. 1. Block diagram of a typical analog-to-digital converter.

modulation. This corresponds to the comparator in the electronic ADC described in Fig. 1.

The accumulator and the sampling circuit are almost the same as the electronic ADC.

II. PRINCIPAL COMPONENTS AND THEIR CHARACTERISTICS

A. Local DAC

Figure 3 embodies the structure of the electroacoustic transducer used in this experiment. The summation of each weighted term in Eq. (1) is performed by the electroacoustic transducer. The summation of each term in Eq. (1) is the synthesis of the resultant vibration of the diaphragm which is actuated by the divided driving electrodes shown in Fig. 3. The vibration of the diaphragm is converted to electrical signals by the detecting electrode, which has an electret layer to compose an electret condenser microphone with the conductive diaphragm.

The diaphragm illustrated in Fig. 3 is a fluoroethylene-propylene (FEP) film with a conductive surface as a grounded electrode of the electrostatic transducer. The diaphragm has a dielectric surface on the other side with charged electret. The driving electrode is separated into several insulated sections of which each surface area is proportional to $2^0:2^1:2^2:2^3::2^n \dots = 1:2:4:8:\dots$.

These values are proportional to the significance of each bit of the binary-coded digital signal. The digital signal is applied to each section of the divided driving electrodes in parallel and corresponding to the significant position at the same electric potential. The electric potential between each driving electrode and the grounded conductive layer of the diaphragm represents the status of each bit. As is well-known, the electrostatic transducer requires a polarizing voltage and the driving force between the electrodes is determined by adding or subtracting the potential voltage to the polarizing voltage depending on the status of each bit. Therefore, the digital signal expressed by Eq. (1) should be slightly modified to the ternary form as Eq. (2) in which each term corresponds to each section of the driving electrode.

$$y = s_0 p_0 2^0 + s_1 p_1 2^1 + s_2 p_2 2^2 + \dots, \quad (2)$$

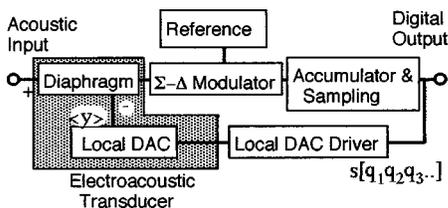


FIG. 2. Schematic diagram of direct-converting digital microphone.

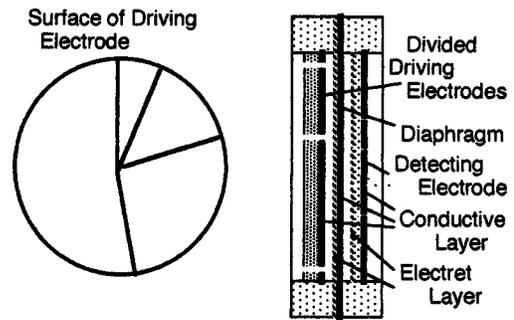


FIG. 3. Sectional view of electroacoustic transducer and surface of divided driving electrode corresponding to 4-bit system.

where s_0, s_1, s_2, \dots are the sign of each bit and p_0, p_1, p_2, \dots are 0 or 1. This modification enables each status of bit to add or to subtract independently and generates tristate signal around the polarizing voltage. Figure 4 shows the outside view of the experimental electroacoustic transducer device for the subtractor and the local DAC.

Figure 4 also shows the driving electrode which is divided into four parts that have the surface area of 1:2:4:8.

Figure 5 shows an example of the operation of this electroacoustic DAC. This figure shows the converted analog signal from the electroacoustic DAC that is the displacement signal of the diaphragm detected by the detecting electrode. Another analog signal which is electronically converted from the same binary signal by an external electronic DAC is used as a reference. In this experiment, the input signal of the sigma-delta modulator is not the displacement signal of the diaphragm, but an external sinusoidal signal generator for the purpose of confirmation of the operation of the sigma-delta modulator and digital-to-analog conversion. In other words, the data presented are for an open loop. Although some distortion is apparent in this data due to the error of the driving-force distribution, it is noted that the DA conversion is performed. The dividing technology of the driving electrode for more nearly precise and higher resolution are the critical issues for design in practice.

B. Subtractor

One of the features of this direct-converting digital microphone is that the subtractor is integrated with the electroacoustic digital-to-analog converter in one device. The diaphragm is driven by the receiving acoustic signal (the sound pressure) and also driven by the electrostatic driving force of the electroacoustic DAC. This function is the subtraction. The subtracted result is transduced to an electric signal by the detecting electrode as the displacement signal of the diaphragm.

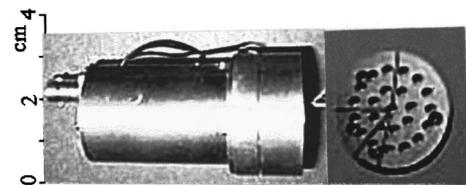


FIG. 4. Outside view of electroacoustic transducer for DAC and subtractor and its driving electrode (right).

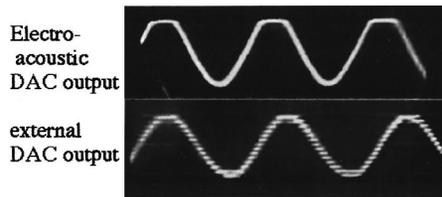


FIG. 5. Operation of electroacoustic DAC. Analog waveforms of converted from 4-bit digital signal by electroacoustic DAC and by external electronic DAC.

C. Comparator

The comparator in this digital microphone is the ± 1 -bit digitizer which is performed by the sigma–delta modulation of the output from the detecting electrode. The output voltage of the detecting electrode is compared with a bipolar reference level. When the output of the detecting electrode exceeds the positive or negative reference level, a ± 1 bit is generated and supplied to the accumulator. This procedure represents the ternary quantization. The local clock frequency of this has to be rather high. The nominal clock frequency of the system f_s and the size of quantization s_d determines the local clock frequency f_1 to be more than f_s times s_d as expressed in Eq. (3).³

$$f_1 \geq f_s \times s_d. \quad (3)$$

In this experiment, the local clock frequency f_1 is 200 kHz for the nominal system clock frequency f_s of 44 kHz and the number of quantization bit s_d of 4 bits. It is necessary for f_1 to be more than 700 kHz for a 16-bit system and 1 MHz for a 24-bit system by Eq. (3). This is another critical subject for future practical design.

D. Accumulator

The accumulator is a kind of electronic counter in which ± 1 -bit data supplied from the sigma–delta modulator are added to the prior data in the accumulator. This incrementally forms a prescribed magnitude and format for the digital signal output of the direct-converting digital microphone. In this experiment, the output data format is binary-coded 4 bits plus sign. This is synchronized to the local clock frequency. The result is the raw digital output, which is redundant enough due to its excess clock frequency for the nominal system clock rate. The raw digital output is sampled, therefore, by the nominal system frequency, which is 44 kHz, and is the digital output of this direct-converting digital microphone. This digital output is fed to the local DAC driver to generate the electrode driving signal.

E. Local DAC driver

It is necessary to convert the binary-coded digital signal to ternary form again, as mentioned previously, to drive the electrostatic transducer. The local DAC driver converts the signal from binary form to ternary and assembles this as parallel data to drive all electrodes simultaneously. The driver includes parallel gates that switch the driving voltage

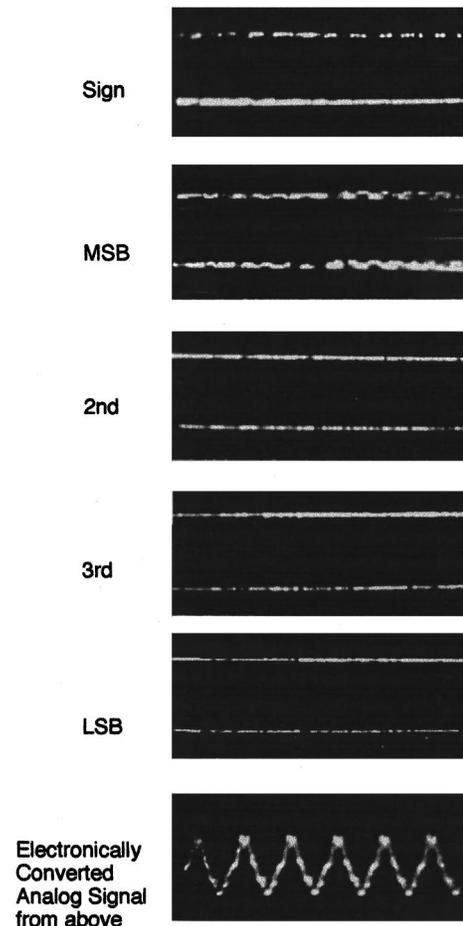


FIG. 6. An example of digital output, sign, MSB, 2nd, 3rd, and LSB. These are converted to analog again by external electronic DAC.

to add to the polarization voltage in the electrostatic transducer controlled by the digital signal. The driving voltage is + and -40 V and the polarization voltage is 300 V given by the facing surface charge of the electret layer of the diaphragm in this experiment. Therefore, the actual driving-voltage between the diaphragm and each electrode is 340 V for the logical true bit, 300 V for 0, and 260 V for -1 bit. The total feedback loop of the ADC of the differential type is completed by this path as the direct-converting digital microphone.

III. ANALOG-TO-DIGITAL CONVERTING TRANSDUCER EXPERIMENT

An example of the operation of a direct-converting digital microphone is the experimental system shown in Fig. 6.

Figure 6 illustrates the digital output waveform of each binary digit, sign bit, most significant bit (MSB), 2nd, 3rd, and least significant bit (LSB). This 5-bit stream was confirmed by reconvert to analog form with an external electronic DAC. In Fig. 6, the reconverted analog signal is illustrated at the bottom, even though the waveform is distorted because of the saturation and error of the surface-area distribution of the driving electrodes for the local DAC. However, it confirms the concept of this direct-converting digital microphone.

IV. CONCLUSION AND APPROACH TO PRACTICAL DESIGN

A. Conclusion of this experiment

It could be concluded from the above experiments that the fundamental concept of the direct-converting digital microphone is confirmed. In this process, we have replaced the local DAC and the subtractor in a conventional electronic ADC by an electroacoustic device which includes an electroacoustic DAC and a subtractor on its diaphragm. This electroacoustic DAC is the same concept as described in a paper by Flanagan in 1980. In this paper, the new electroacoustic DAC is designed as a derived structure from that previous work in order to meet the requirements as the component of an A-to-D Converter. The structure involves a development to detect the displacement of the diaphragm, having a resonance at the highest point of the covering frequency which acts as a low-pass filter for electroacoustic DA conversion. This structure enables the integration on the subtractor in the electroacoustic transducer. The advantages of the direct conversion between analog acoustic and digital electric signals are different in each application and depend on the degree of integration, accuracy, physical dimension, linearity, and cost competitiveness, which are the technical subjects of further studies. Consequently, it is necessary to work in collaboration with system manufacturers as the customer for future development of this technology.

B. Some issues for practical design

1. The resolution or dynamic range for sufficient signal quality

One of the most interesting and critical issues of this microphone design is the possibility of higher resolution or larger number of quantization bits. From the paper by Yanagisawa⁴ and the authors' experience, it would be almost impossible to increase the resolution of the divided electrodes by more than 8 bits. For higher quality by increasing the resolution of the electroacoustic DAC, the summation method in Eq. (1) should be changed to another procedure. The acoustical summation employing the integration of small electroacoustic elements is to be adopted for this purpose.

Figure 7 shows the experimental 4-bit electroacoustic DAC and the subtractor of an acoustical summation type. This is composed of three essential components: small loudspeakers, small microphones, and a cavity in which to embed them. The small loudspeakers are classified into groups in which the number of small loudspeakers are 2^n (n is 0,1,2, ...) corresponding to the place of the significance of each bit and connected in parallel in each group. Sound is radiated into the cavity from these loudspeakers as a pulse sound whose sound pressure is proportional to the number of the activated loudspeakers to satisfy each term of Eq. (1). The summation in Eq. (1) is performed by sound pressure in the cavity.

Small microphones are also implemented in the cavity to detect the result of the acoustical summation. The cavity is small enough not to have any standing wave inside. Figure 7 also shows the picture of the small loudspeakers and microphones used in this device. In this acoustical summation

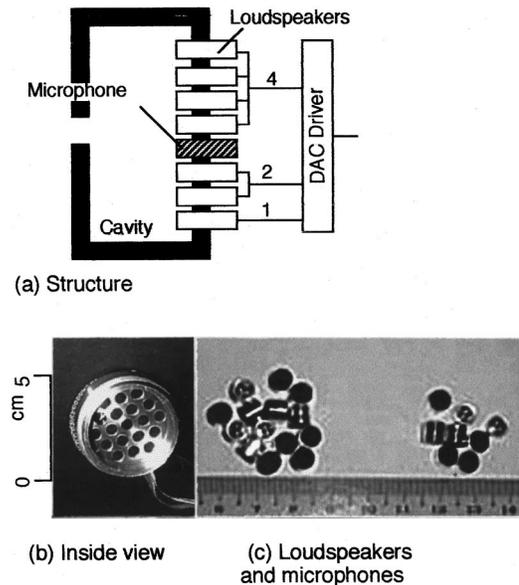


FIG. 7. Configuration of acoustical summation-type DAC.

type, the receiving acoustic signal through the orifice of the cavity is mixed with the synthesized radiated sound from the loudspeakers that is the output of the electroacoustic DAC.

This mixing is regarded as the subtraction when the phase of digital driving signal is set appropriately. Figure 8 shows an example of subtraction of this type. The output of the microphone which detects the sound pressure in the cavity shows a decrease of more than 10 dB by subtracting. This value is acceptable for the present time because the deviation of the efficiency of the loudspeakers is more than 2 dB and this is to be minimized in the future.

Figure 9 shows data on the linearity of this transducer. Small loudspeakers are driven at constant frequency and constant amplitude. The sound-pressure level in the cavity was observed by the coupled sound-level meter through a 2-cc coupler. Since the small loudspeakers are driven by a pulse train of 44-kHz clock frequency in actual operation, these data do not represent the characteristics of the transducer perfectly. However, these data indicate the cause of

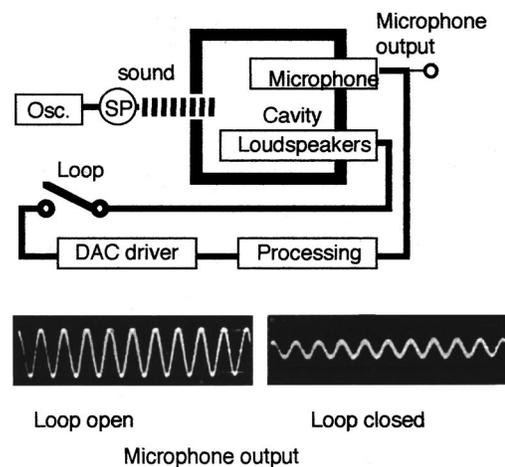


FIG. 8. Subtraction result of the digitally feedback radiated sound and received sound from outside.

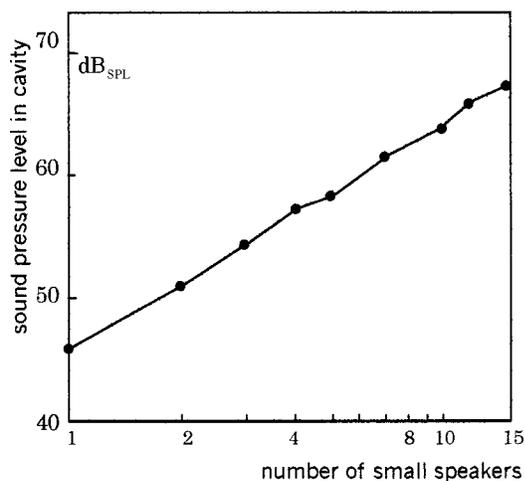


FIG. 9. Sound-pressure level in cavity radiated from small loudspeakers. The frequency of 1 kHz and amplitude of 5 V rms are both held constant.

distortion due to the incomplete uniformity of loudspeakers as mentioned above.

2. Implementation process

The practical fabrication utilizing some kind of integrated mechanism is to be introduced using micro-electro-mechanical technology (MEMS). Compared to the vibrational summation-type DAC, the acoustical summation type seems to be more suitable to MEMS or semiconductor technology because of the uniformity of electroacoustic elements. In order to have effective implementation, collaboration with semiconductor process engineers or MEMS is indispensable.

V. CONCLUSION

This paper should be regarded as a conceptual proposal for the future development of electroacoustic devices by introducing digital technology to the connection between electroacoustics and electronics. Also, when MEMS technology is introduced to the small electroacoustic devices to increase the number of elements, the direct-converting digital microphone and other new integrated systems could be practical.

ACKNOWLEDGMENTS

The authors would like to express their thanks to T. Yokoyama, the managing director of the Audio-Visual Systems Division of Matsushita Communication Industrial Co., Ltd. for his support of this study. They also wish to express their thanks to Dr. J. L. Flanagan, the Vice President of Rutgers University, for his kind comments on his paper. They also wish to thank Professor K. Kido of Chiba Institute of Technology and Professor Dr. T. Kamakura of Electrical Communication University for their useful advice.

¹J. L. Flanagan, "Direct digital-to-analog conversion of acoustic signals," *BSTJ* **59**, 9 (1980).

²Y. Yasuno and Y. Riko, "A conceptual experiment of direct converting digital microphone," *J. Acoust. Soc. Am.* **103**, 2872(A) (1998).

³Y. Yamazaki, M. Noma, H. Iizuka, H. Ohta, and M. Nishikawa, "High speed 1 bit processing of wide range acoustic signal," Technical Report of IECE, EA93-102 (1994-03).

⁴T. Yanagisawa, "On the digital loudspeaker," *JIEICE J72A* 11 (Nov. 1989), pp. 1724-1752.

Magnetic excitation and acoustical detection of torsional and quasi-flexural modes of spherical shells in water

Brian T. Hefner and Philip L. Marston^{a)}

Department of Physics, Washington State University, Pullman, Washington 99164-2814

(Received 7 January 1999; accepted for publication 14 September 1999)

A simple electromagnetic acoustic wave transducer (EMAT) based on a wire coil was used to excite the resonant modes of a stainless steel spherical shell without direct mechanical contact. The coupling produced by this EMAT was examined first for shells in air and then for shells immersed in water to examine the effects of fluid loading on the shell's spectrum. It was found that the torsional modes were excited using this method and these modes radiated sound into the surrounding water contrary to expectations. This excitation is shown to depend on the presence of a permanent magnetization in the shell itself or on the presence of a static external field applied at right angles to the axis of the coil. Possible mechanisms for the excitation and the acoustic radiation of the torsional modes are considered. The excitation of quasi-flexural shell modes is also discussed for shells in air and in water. The shell responds at the oscillation frequency of the applied field and at twice the frequency. Some potential applications of this method of measuring modes are noted. © 1999 Acoustical Society of America. [S0001-4966(99)07012-5]

PACS numbers: 43.40.At [PJR]

INTRODUCTION

Various acoustical methods have been developed for investigating the modes of fluid-loaded shells including measuring the acoustic response to tone bursts at a succession of frequencies¹ and the acoustic response to wide-bandwidth incident waves radiated by an impulsive source.² For some applications, however, certain modes of interest are only weakly excited by sound waves incident on the shell and it is appropriate to consider electromagnetic approaches for the noncontact excitation of shell modes while monitoring the shell's response from the resulting acoustic radiation. The research described here gives an example of a hybrid electromagnetic-acoustic approach for an empty spherical shell in air or in water.

We demonstrate the excitation of the resonant modes of a stainless steel spherical shell with a simple electromagnetic acoustic wave transducer (EMAT). With the shell in air, the resulting acoustic radiation was detected with a microphone placed on the axis of the coil near the surface of the shell as shown in Fig. 1(a). When the shell was in the water, a hydrophone was used. The sinusoidal currents through the coils shown in Fig. 1 result in oscillating electromagnetic stresses on the shell referred to as Maxwell stresses.³ From the spatial distribution of Maxwell stresses discussed in Sec. I, this transduction technique is shown to excite the quasi-flexural modes of the shell. It was observed, however, that it is also possible to drive the torsional modes of the shell. The excitation and detection of these modes is important since for a perfectly spherical shell with a wall of homogeneous composition and thickness, the torsional modes are acoustically inactive. For example, there are no contributions from torsional modes to the computed total scattering cross section for such an ideal shell.⁴ The excitation of these modes in our

experiments requires either that the shell be given a permanent magnetization or the addition of a bias magnetic field as discussed in Secs. IV–VI. The observed torsional modes have extremely high quality factors when compared to the lowest quasi-flexural mode when the shell is in water, indicating that very little energy is being lost to the surrounding medium. This is expected since the torsional modes of the shell should exhibit little or no motion of the surface in the radial direction. The presence of these oscillations was, nevertheless, detected via a significant acoustic response.

The particular type of coil transducer employed here was first developed by Johnson *et al.* for the ultrasonic study of resonant modes of solid metallic spheres.⁵ Like most conventional EMATs, the transducer used by Johnson makes use of a Lorentz force interaction between eddy currents induced on the sample and an external bias magnetic field.^{6,7} One form of the transduction mechanism demonstrated here in our observation of the excitation of quasi-flexural modes is similar to Johnson's, however, we also investigate another contribution to the Lorentz force on the sample when using a coil EMAT. We demonstrate that the eddy currents interact with the induction field itself to produce a weak stress on the shell that oscillates at twice the frequency of the coil's field. The external bias field can then be removed and the transducer becomes simply a coil and an ac current source. An approximate description of these oscillating stresses is given in Sec. I.

This type of EMAT was summarized previously by us for thin, nearly spherical aluminum shells in air;⁸ however, the present study reveals more spectral features. The present research utilizes a shell with a more uniform wall thickness for the purpose of investigating this transduction method. In Secs. II and III we summarize mode calculations appropriate when the shell is surrounded by air and the corresponding observations. This serves to illustrate the spectrum in the limit of weak fluid loading.

^{a)}Electronic mail: marston@wsu.edu

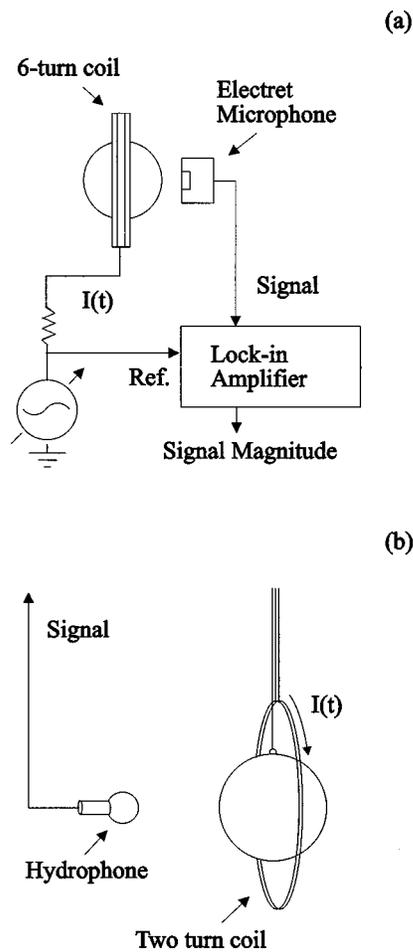


FIG. 1. Configurations used for measuring the response of the shell in air (a) and in water (b). In (a) the electret microphone was uncalibrated and the shell sat on cotton-covered mounts. The measurements in Fig. 2(a) were taken with the shell, coil, and microphone between the poles of a large electromagnet which is not shown. The hydrophone in (b) was 22 cm from the center of the shell for the measurements in Figs. 3–7.

I. APPROXIMATION FOR THE MAXWELL STRESS ON A CONDUCTING SPHERE

The creation of oscillating stress distributions appropriate for exciting the quasi-flexural modes of the shell can be explained by considering the simplified case where the shell is treated as a nonmagnetic perfect electrical conductor (PEC). This approximation should give quantitatively useful approximations for the applied surface stress when the electromagnetic skin depth δ is significantly less than the shell thickness h and it is possible to neglect any ferromagnetic properties of the shell. The similarity of the magnetic B -field distribution around a highly conducting sphere with the PEC result is well known from analytical results for homogeneous spheres of finite conductivity in an applied oscillating field that is spatially uniform in the absence of the sphere.^{9,10} It is found that the magnitude and phase of the B field generated by the induced current is such that the total B field goes over to the PEC result when the electromagnetic skin depth $\delta \ll a$ where a is the radius of the sphere. The skin depth is given by^{3,9,10} $\delta = (\pi f \mu \sigma)^{-1/2}$ in SI units where f is the frequency, μ is the permeability, and σ is the electrical conductivity. For the stainless steel shell used in our experiments,

$1/\sigma \approx 0.71 \times 10^{-6} \Omega\text{m}$ and for the purpose of estimating an upper bound on δ , μ is approximated by the free space value $\mu_0 = 4\pi \times 10^{-7}$. The resulting values of δ range from 2.3 mm at 35 kHz to 1.4 mm at 95 kHz. The shells studied had a thickness $h = 3.086$ mm so that the stress distribution predicted for the lowest frequencies used are only qualitatively useful. An additional simplification in the analysis given below is the assumption that the applied oscillating B field is spatially uniform in the absence of the sphere. As shown in Fig. 1, the oscillating B field is generated by circular coils which are not very much larger than the sphere. One consequence is that the actual stress distribution is more favorable for the excitation of higher-order quasi-flexural modes than the stress distribution approximated below.

For the transduction technique discussed here, a metallic spherical shell is placed in a magnetic field with a uniform static component, $\mathbf{B}_1 = B_1 \hat{\mathbf{z}}$, and an oscillating component such that the total magnetic flux density is $\mathbf{B}(\mathbf{r}, t) = \mathbf{B}_1 + \mathbf{B}_2(\mathbf{r}, t)$. When the shell is treated as a perfect electrical conductor (PEC), $\mathbf{B}_2(\mathbf{r}, t)$ is completely excluded from the shell. (See, e.g., p. 460 of Ref. 3.) The spatial dependence of the oscillating field \mathbf{B}_2 becomes similar to what a static field has around a perfectly diamagnetic sphere ($\mu = 0$). (See, e.g., p. 200 of Ref. 3.) For a uniform applied oscillating field, $\mathbf{B}_0 = B_0 \sin \omega t \hat{\mathbf{z}}$, the magnetic field at the sphere's surface becomes

$$\mathbf{B}_2(r = a, t) = -\frac{3}{2} B_0 (\sin \omega t) (\sin \theta) \hat{\theta}. \quad (1)$$

Neglecting any ferromagnetic response, the static field is unaffected by the presence of the shell and hence remains uniform. The total field at the surface is then

$$\mathbf{B} = B_1 \hat{\mathbf{z}} - \frac{3}{2} B_0 (\sin \omega t) (\sin \theta) \hat{\theta}. \quad (2)$$

Using the Maxwell stress tensor,³

$$\vec{T} = \frac{1}{\mu_0} \left[\mathbf{B}\mathbf{B} - \frac{1}{2} |\mathbf{B}|^2 \vec{T} \right],$$

and the field distribution given by Eq. (1), the radial force-per-area on the shell's surface is given by the oscillating part of $\hat{\mathbf{r}} \cdot \vec{T}$ since there is no oscillating B field within a PEC sphere. The normal force becomes

$$F_r = -\frac{3}{2\mu_0} B_0 B_1 (\sin \theta)^2 \sin \omega t - \frac{9}{8\mu_0} B_0^2 (\sin \theta)^2 (\sin \omega t)^2. \quad (3)$$

The first term describes the Lorentz force of the static field on the eddy current and corresponds to the coupling depicted in Fig. 1 of Ref. 5. This force oscillates at the frequency of the oscillating field. The second term describes the Maxwell stress of the oscillating field alone. From the identity $(\sin \omega t)^2 = (1 - \cos 2\omega t)/2$, this force varies at a frequency that is twice that of the oscillating field. Both terms have the same distribution and couple well into the quadrupole quasi-flexural mode of the shell, denoted by A_2 . This is evident by expressing $\sin^2 \theta$ as $-\left(\frac{2}{3}\right)[P_2(\cos \theta) - 1]$, where P_2 is a Legendre polynomial with an argument of $\cos \theta$. Also there are

TABLE I. Material properties for the stainless steel shell and shell dimensions.

Density (g/cm ³)	Transverse velocity, c_s (mm/ μ s)	Longitudinal velocity, c_l (mm/ μ s)	Radius, a (mm)	Ratio of thickness to outer radius, h/a
7.46	3.36	6.11	19.05	0.164

no forces which act in the θ or ϕ directions. The sign of the second term in (3) corresponds to a radially inward force while the first term has that direction only when $\sin \omega t$ is positive. The direction of the equatorial stress can be explained by considering the magnetic pressure associated with local time-dependent flux density. (See, e.g., p. 320 of Ref. 3.) For the purposes of the present application, the magnitudes of the strain oscillations need not be determined since prior experience with an EMAT indicates that signals radiated into water may be easily detected with hydrophones.⁷

II. CALCULATED MODAL PROPERTIES NEGLECTING FLUID LOADING

The stainless steel shell examined with this transduction method had a radius of $a = 19.05$ mm and the ratio of the thickness to the outer radius of $h/a = 0.164$. The elastic properties of the 440 C stainless steel used were found for a rectangular parallelepiped cut from a shell fabricated by the same vendor as the shell used in experiments. These were determined using resonant ultrasound of a rectangular parallelepiped¹¹ and the results are presented in Table I.¹² The properties found were similar to those assumed in the analysis of previous burst scattering for the same type of shell in water.^{4,13} The resonant modes of the shell in the absence of significant fluid loading were calculated using these values by solving the elastic continuum equations for a spherical shell of finite thickness subject to the stress-free boundary conditions.¹⁴ For the spherical shell in the frequency range of interest, there are three types of modes: quasi-compressional, quasi-flexural, and torsional. The quasi-compressional and quasi-flexural modes have similarities to standing symmetric and antisymmetric Lamb waves, respectively. Torsional modes, however, are characterized by motion that involves no radial displacement of the shell surface.^{14,15} Details about the calculation of torsional modes are given in the Appendix.

Due to equipment constraints, the frequency range of interest for this experiment was 20 to 100 kHz. With the exception of the breathing mode of the shell, the quasi-compressional modes lie outside this range, so attention here is given primarily to the quasi-flexural and torsional modes. The notation used and the calculated resonant frequencies are given in Table II. See also the Appendix.

III. MEASURED MODE SPECTRUM IN AIR

Figure 1(a) shows a simplified diagram of the experiment performed to measure the response of the shell in air. The shell was placed on a three-point support inside a six-turn copper wire coil with a diameter of 7 cm. Cotton was placed between the shell and the supports in order to reduce

TABLE II. Mode frequencies.

Mode type	Notation	Calculated ^a (kHz)	Measured ^b (kHz)
Quasi-flexural mode	A_2	37.0	36.6
	A_3	48.1	47.7
	A_4	60.6	59.7
	A_5	77.2	75.6
	A_6	97.4	96.7
	Breathing mode	E_0	82.0
Torsional mode	T_2	60.5	59.8
	T_3	95.8	95.2

^aFor the empty shell in a vacuum.

^bWith the shell in air except for T_2 and T_3 which are for the shell in water. For the T_2 and T_3 modes, ka is 4.9 and 7.7, respectively, where k is the acoustic wave number in water.

damping due to contact. For the configuration where a static field B_0 was superposed on an oscillating field, the static field was provided by an electromagnet having large iron poles that produced field strengths up to 1200 Ga. A continuous sine-wave current of 0.08 A was provided to the coil when the static field was present. Spectra were also obtained without the static field but with the current through the copper coil increased to 3.8 A. Since the coupling was expected to be strongest to the A_2 mode, an electret condenser microphone was placed on the axis with the coil and shell to measure the response of the shell. The signal magnitude was monitored with a two-phase lock-in amplifier with the reference frequency at ω or 2ω , where ω is the frequency of the oscillating magnetic field.

The response of the shell in the presence of the static field is shown in Fig. 2(a). As expected there was strong coupling into the A_2 mode. Other modes of the shell were also excited and in particular the E_0 mode, or ‘‘breathing mode,’’ of the shell produced a very strong signal. The coupling to the E_0 mode is expected from the constant term in the expression for $\sin^2 \theta$ given below Eq. (3). Note that the A_3 mode is absent from the spectrum. This can be explained by noting that the mode is not symmetric about the equator of the shell and hence the coupling should be extremely weak. It was possible to detect the A_3 mode, however, by offsetting the shell so that the oscillating magnetic field was no longer symmetric, thus producing a greater ‘‘squeezing’’ of one-half of the shell over the other.

The spectrum found using this transduction technique compared well to the calculated spectrum as is seen in Table II. There was, however, a great deal of mode splitting in all modes except the E_0 mode. The shell was used previously in experiments that examined the backscattering of short tone bursts and details of the shell’s construction are given there.¹³ In those experiments, the presence of the seam which connects the two halves of the shell and imperfections on the inner surface of the shell would not have produced perturbations strong enough to be easily detected in that experiment. In the present work, however, these imperfections would contribute to the mode splitting. Evidence for this conclusion comes from the observation that the orientation of the shell has an influence on the magnitude of the individual peaks that cluster about each mode.

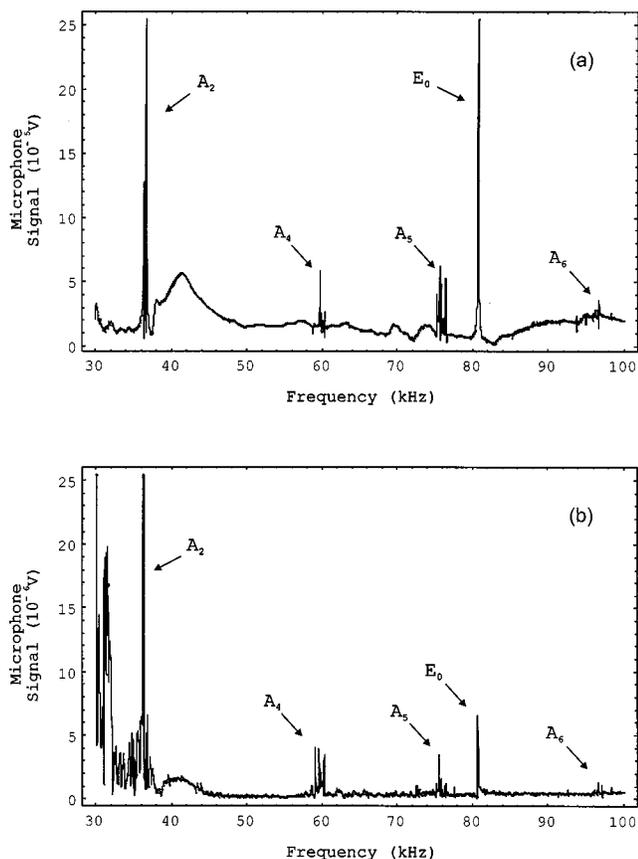


FIG. 2. Mode spectrum with the shell in air. In (a) a static magnetic field is present and the mode frequency plotted is the drive frequency. In (b) the static field is removed and the mode frequency plotted is twice the drive frequency. The modes are classified as tabulated for an ideal shell in Table II. The coupling in (a) is associated with the Lorentz force of the static field on the eddy currents while in (b) the second term in Eq. (3) is relevant.

When the static field was removed, the response of the sphere was measured at twice the frequency of the oscillating field of the transduction coil because of the $\sin^2 \omega t$ term in Eq. (3). The resulting spectrum is given in Fig. 2(b). As expected, even though the amplitude of the current across the transduction coil was increased, the signal of the response dropped significantly. Once again the A_2 mode is the most prominent feature. There is a clear correspondence between the two spectra and comparisons with the calculated frequencies indicate that both methods can be used to examine the resonant structure of the shell.

IV. MEASURED MODE SPECTRUM IN WATER

To examine the effects of fluid loading, a second coil was constructed and the experiment was performed in a 3000-gallon wooden tank of water. A simplified diagram of this experiment is shown in Fig. 1(b). The new coil was 5.75 cm in diameter and consisted of two turns of copper wire. It was attached to the end of a plastic rod and lowered to a depth of 1.3 m to avoid any reflections or interactions with the surface. The tank itself had a depth and diameter of 2.4 m. The shell was attached to a length of fishing line by a small drop of epoxy and suspended so that it was positioned in the center of the coil. A spherical hydrophone (Edo model 6166) was positioned along the axis of the coil to monitor the

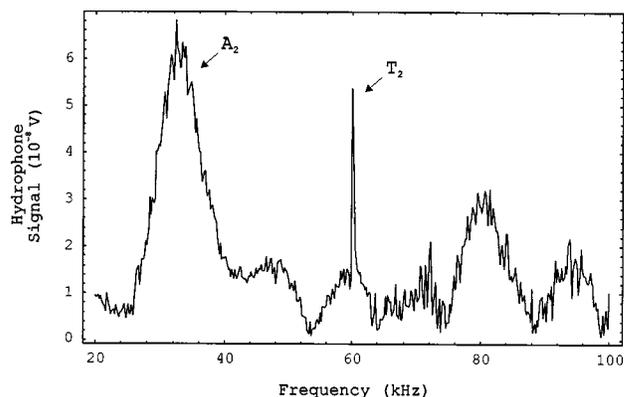


FIG. 3. Mode spectrum with the shell placed in water with the significant modes identified as shown. The frequency plotted is twice the drive frequency and the signal is extracted with a lock-in amplifier and the peak signal is $0.06 \mu\text{V}$. While the hydrophone response varies weakly with frequency over this range, here and in Figs. 4–6 the pressure is approximately $0.016 \text{ Pa}/\mu\text{V}$.

response of the shell. In this experiment there was initially no external static magnetic field present and the response of the shell was measured at twice the frequency of the current applied to the coil.

The response magnitude of the shell is shown in Fig. 3. It was very difficult to eliminate the coupling between the oscillating magnetic field and the hydrophone due to the hydrophone's geometry. To remove a majority of the coupling from the spectrum, the frequency response was also measured with the shell absent from the system and this background was subtracted from the raw data in Fig. 3. As expected, the most prominent feature is the A_2 mode. The mode is shifted from the frequency found in air by $\Delta f = -4.2 \text{ kHz}$ down to 32.4 kHz with a quality factor, $Q \approx 3.8$. This Q was determined from the decay following tone burst excitation as described in Sec. VI. The frequency for this mode subject to fluid loading was estimated by calculating the partial wave $l=2$ contribution to the partial wave series (PWS) representation of the forward scattering amplitude for an evacuated elastic spherical shell.⁴ From this, the $l=2$ contribution to the forward scattering from a rigid sphere was subtracted in an effort to isolate the resonance of the shell. The frequency for this mode was found to be 32.5 kHz which is close to the measured value.

In addition to the A_2 mode, the spectrum has structure at higher frequencies that appears to be unrelated to higher-order quasi-flexural modes. The most prominent of these features is the very narrow resonance at 59.8 kHz. This peak corresponds to the first torsional mode of the shell, denoted as T_2 , which is calculated to occur at 60.5 kHz. Although it is not presented here, there were also indications of a slight peak at 95.2 kHz indicating the next torsional mode, T_3 , may have been excited as well. The high quality factor of the T_2 mode provides additional support for this identification. However, the excitation of these modes is not explained by the analysis presented in Sec. I, indicating the possibility of some other coupling mechanism. It was found that the stainless steel shell had become permanently magnetized. This shell had been used previously in the experiments described in Sec. III and the repeated application and removal of the large static field appears to have induced the magnetization.

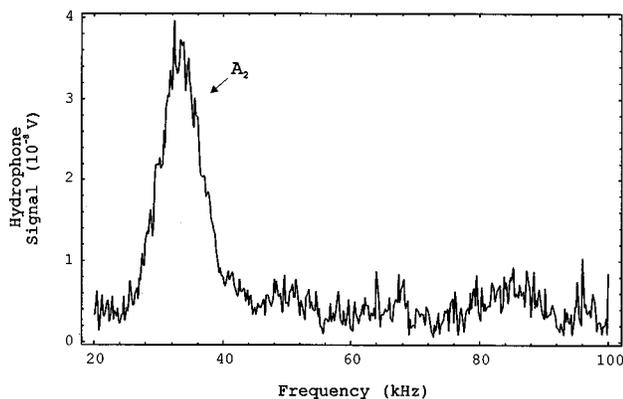


FIG. 4. Spectrum obtained as in Fig. 3 but with an identical shell having no permanent magnetic moment. The coupling to the torsional mode T_2 is suppressed while the coupling to the A_2 mode is expected from the second term in Eq. (3).

Stainless steel 440 C is a martensitic phase¹⁶ of Fe, Cr, and C which is known to allow permanent magnetization. To determine if the magnetization was important to the coupling into these torsional modes, the same experiment was performed with an identical shell with no significant magnetization. As can be seen in Fig. 4, the A_2 mode is once again present and has roughly the same frequency and quality factor as found for the magnetized shell. However, there is no discernible coupling into any higher order quasi-flexural or torsional modes. This indicates that the presence of the magnetization of the shell may be important for the coupling to the torsional mode shown in Fig. 4. Calculations⁴ for an ideal spherical shell in water of the same dimensions reveal no sharp acoustically active resonances in this frequency range. The resonances on the water-loaded shell associated with the modes labeled A_4 and A_6 in Table II are found to be much broader in the total cross section⁴ than the observed peaks which we attribute to the T_2 and T_3 modes.

V. EXPLORATION OF TORSIONAL MODES AND OTHER SPECTRAL FEATURES

The possibility of driving shell oscillations at the frequency ω of the current was explored by first determining the direction of the magnetic moment of the shell so that its orientation could be controlled. This was achieved by repeatedly passing the shell through a coil in air and monitoring the induced current for different orientations of the shell until the orientation of the magnetic moment \mathbf{m} was found. The orientation of \mathbf{m} relative to the seam in the shell could not be determined since the seam was not visible. The driving forces on the shell should then have components that oscillate at both ω and 2ω . Figure 5(a) shows the response of the shell with the lock-in reference signal at ω in the configuration where \mathbf{m} is parallel to the coil's axis. As in the experiment performed in air, the signals can be substantially larger than when the response is measured at 2ω . Also the A_2 mode is very prominent as expected from an excitation mechanism discussed in Sec. VI. To confirm this, Fig. 5(b) shows the response of the shell when \mathbf{m} is aligned perpendicular to the axis of the coil. Analysis shows that the radial component of the driving force that oscillates at ω in this configuration is

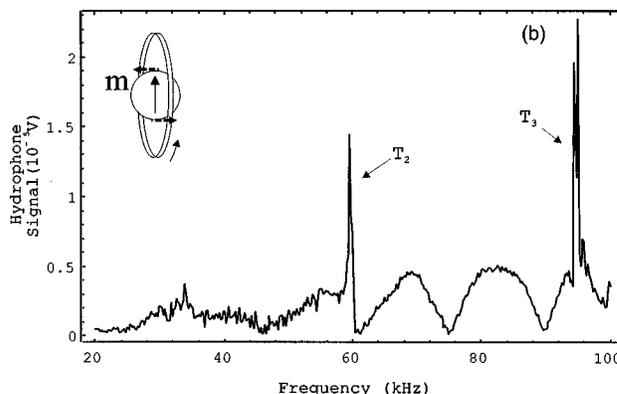
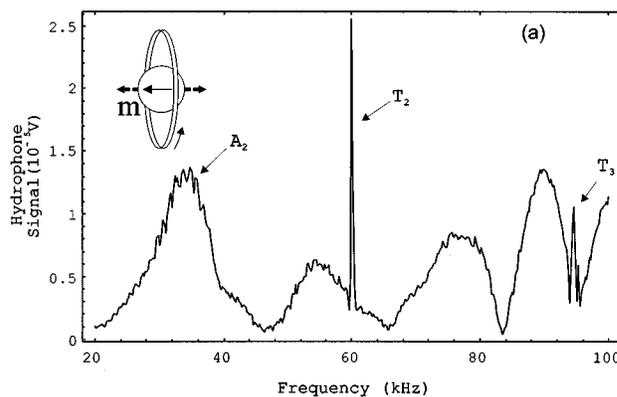


FIG. 5. Mode spectra obtained with the frequency plotted equal to the drive frequency and the permanent magnetic moment \mathbf{m} of the shell oriented according to the thin arrow shown in the inserts. The bold dashed arrows show the relative orientation of the local magnetic force near the poles during one phase of the excitation cycle where the coil current is in the direction indicated. See Sec. VI. Unlike Fig. 2(a), there is no static component to the applied field. Direct coupling to the hydrophone has been suppressed in (a) and (b) as explained in Sec. IV. The peak signals exceed 22 μV .

reduced and, as a result, there should be negligible coupling into the A_2 mode. From Fig. 5(b) it appears that this mode is suppressed.

In addition to the A_2 mode, there is a great deal of structure when the moment is in either orientation. The most prominent of these structures are the torsional modes. The degree of coupling depends on the orientation of the magnetic moment. In Fig. 5(a), the transducer couples well into the T_2 mode but not as strongly into the T_3 mode. When the moment was rotated, the coupling became stronger with T_3 and less so with the T_2 mode. This indicates that the orientation of the moment relative to the oscillating field plays an important role in the coupling mechanism at least for the response at ω . This is considered in more detail in Sec. VI.

The high Q torsional modes are very distinct in Figs. 5(a) and 5(b), but there also appear to be some much broader features in the response which do not correspond to modes of the shell. These may be a result of producing a Lorentz force on the coil due to the shell's magnetic field. When the shell's moment is aligned with the coil axis, there is a radial Lorentz force that couples into the extensional modes of the coil. When the moment is perpendicular, the forces should act out of the plane of the coil and couple into the coil's bending modes. This suggests the cause of the differences in the spec-

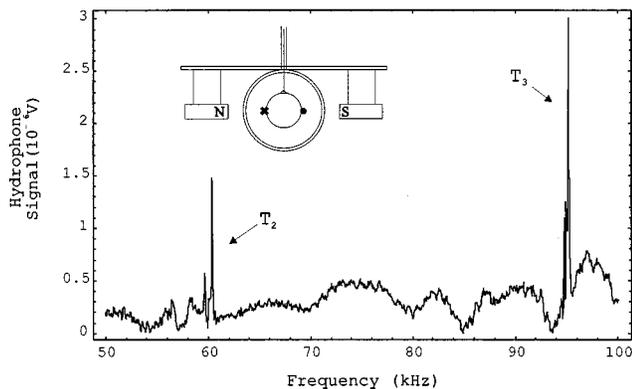


FIG. 6. Mode spectrum which confirms that torsional modes can be excited on the shell having no permanent magnetization provided a static magnetic field is superposed. The torsional features correspond to the T_2 and T_3 in Fig. 5 and the signal is detected at the frequency of the drive. The cross and point on the equator show the tail and head of the local Lorentz force vector on the eddy current during one phase of the excitation cycle and this torque causes the T_3 mode to be excited. The relevant eddy current is in the meridional plane of the sphere which intersects the coil. The T_2 and T_3 modes are not observed for this shell when the magnets are removed.

trum for the two orientations. For the coil used, the calculated mode spacing for coil's extensional modes should be $\Delta f \approx 10$ kHz, which corresponds well to the spacing of the structures in Fig. 5(a) and this suggests that those anomalous features are a result of acoustic radiation by the coil.

Even if the shell lacks significant permanent magnetization, the Lorentz force of the eddy currents produces a torque on the shell when the static and oscillating components of externally applied B fields are perpendicular. See, e.g., Fig. 2 of Ref. 5 where the torques are shown to oscillate at the frequency ω of the eddy currents. To test this idea, two small bar magnets were hung on either side of the coil with their poles aligned such that they provided a magnetic field perpendicular to the axis of the coil. This arrangement is shown in the inset in Fig. 6. The nonmagnetized shell was used in this experiment. As can be seen in Fig. 6, it was indeed possible to drive the torsional modes and the degree of coupling into each was in a manner consistent with the expected symmetry of the Lorentz forces: the coupling with the T_3 mode exceeding that of the T_2 mode.

VI. TORSIONAL MODE EXCITATION AND ACOUSTIC RADIATION MECHANISMS

The detailed description of the mechanics of magnetized solids (with or without eddy currents) in the presence of oscillating applied magnetic fields is a complicated area of continuum mechanics that is beyond the scope of our investigation.^{6,17,18} We summarize below physical mechanisms for exciting some of the modes observed in Figs. 3–5 evident from elementary symmetry considerations. Throughout our discussion, the relative permeability is taken to be unity for the water external to the shell and any gas contained within the shell. The excitation of the A_2 mode in Figs. 3 and 4 is to be expected from the second term of Eq. (3). A coupling mechanism which explains why the A_2 mode is excited in Fig. 5(a) but not in Fig. 5(b) may be summarized as follows. Neglecting the forces resulting from eddy

currents, the magnetic force on the whole magnetized body can be expressed as the sum of the volume integral of a body force per-unit-volume \mathbf{f}_M and a surface integral of a stress \mathbf{T}_M . Let \mathbf{M}_S denote the static component of the shell's magnetization. The applied field causes forces on the internal dipoles within the solid such that there will be contributions to \mathbf{f}_M and \mathbf{T}_M which oscillate at the frequency of the oscillating part of the applied B field denoted by $\mathbf{B}_2(\mathbf{r}, t)$:

$$\mathbf{f}_{M(S2)} = -\mathbf{B}_2(\nabla \cdot \mathbf{M}_S), \quad (4a)$$

$$\mathbf{T}_{M(S2)} = \hat{\mathbf{n}} \mathbf{M}_S \mathbf{B}_2, \quad (4b)$$

where $\hat{\mathbf{n}}$ is an outward directed surface normal, where \mathbf{B}_2 does *not* include contributions to the total field due to the magnetization of the solid. See, e.g., Eq. (3-4.4) of Refs. 18 and 19. The factors $-\nabla \cdot \mathbf{M}_S$ and $\hat{\mathbf{n}} \cdot \mathbf{M}_S$ represent, respectively, volume and surface distributions of effective magnetic "charge." (See, e.g., Sec. 5.9 of Ref. 3.) The corresponding contributions to the volume and surface torque densities are known to be¹⁹ $\mathbf{r} \times \mathbf{f}_{M(S2)}$ and $\mathbf{r} \times \mathbf{T}_{M(S2)}$, respectively, where \mathbf{r} is the radius vector from the origin which is taken to be the center of the shell.

Consider the situation where \mathbf{M}_S is taken to be uniform within the shell so that the volume force and torque densities vanish. The surface stress will in general be present on both the inner and outer surfaces of the shell where, within a given solid angle $d\Omega$, the signs of the surface charge densities $\hat{\mathbf{n}} \cdot \mathbf{M}_S$ are opposite for the inner and outer surfaces. The stress on the inner surface element is reduced because the eddy currents result in a significant reduction of the field B_2 at the inner surface of the shell. Consequently, there is a net radial force on each pole of the shell in Fig. 5(a) (shown by the dashed vectors), but the corresponding forces are tangential in Fig. 5(b). The resulting stresses are expected to couple strongly with the A_2 mode in Fig. 5(a) but weakly (if at all) in Fig. 5(b) in agreement with the observations. Comparison of the stresses indicates that this mechanism should couple strongly into the T_3 mode in Fig. 5(b) but weakly if at all into the T_2 mode in Fig. 5(a). The contribution to the $\mathbf{r} \times \mathbf{T}_{M(S2)}$ torque on the interior surface is further reduced relative to the outer contribution by the smaller value of \mathbf{r} on the inner surface. In the measurements shown in Fig. 5, if the shell and coils are assumed to be concentric and the shell is uniformly magnetized, it is not possible to explain the observed coupling to the T_2 mode from Eq. (4). Notice, however, that in agreement with these considerations, the excitation of the T_3 mode is stronger in Fig. 5(b) in comparison with the T_2 mode, if differences in the coupling to the radiated sound to hydrophone position can be neglected.

The coupling to the T_2 mode evident in Figs. 3, 5(a), 5(b), and 6 is not explained by a symmetric stress distribution described by Eq. (4b). The high quality factor Q of the torsional modes may cause weak coupling processes to be significant. While no single coupling mechanism may explain all of these examples, a partial list of other coupling mechanisms is summarized as follows.

- (i) The static magnetization may be nonuniform giving rise to body forces from Eq. (4a) in Figs. 3, 5, and 6.

- (ii) Imperfect alignment of the centers of the coil and the sphere could alter the spatial distribution of \mathbf{T}_M and Lorentz stresses in Figs. 5 and 6, respectively.
- (iii) In the shell having a permanent magnetization, the magnetic field from that magnetization causes a Lorentz force on the eddy currents which oscillates at the drive frequency ω which may contribute to the coupling evident in Fig. 5.
- (iv) The *Einstein–de-Haas effect* requires that a change in magnetization gives rise to a torque because of the associated change in microscopic angular momentum.^{20–22} Depending on the spatial distribution of the *time-dependent part* of the magnetization for the measurements in Figs. 5 and 6, the resulting distribution of torques could couple into the modes T_2 and T_3 at the modulation frequency of the applied B field.
- (v) The time-dependent part of the magnetization may yield stresses and body forces oscillating at ω and 2ω from a coupling analogous to that considered in Eq. (4).
- (vi) Magnetostrictive mechanisms may contribute as reviewed in Ref. 17 and by Thompson⁶ and pp. 37–40 of Moon.¹⁸ Note, however, that the stainless steel used has a negligible nickel content unlike most highly magnetostrictive materials. The coupling in (iii) gives a second mechanism for exciting the A_2 mode in Fig. 5(a).

Due to the nature of the torsional modes of spheres, it is surprising to find that their presence could be detected acoustically. The quasi-flexural modes of the shell were able to radiate sound because the surface velocity has a radial component. Since this is not the case for torsional modes of an ideal sphere, one would not expect to either excite or detect these modes via acoustic radiation. The high quality factors do indicate that very little acoustic energy is lost to the surrounding medium. When compared to the degree of damping which the A_2 mode of the shell experiences, these high quality factors support the notion that there is little or no radial displacement of the shell's surface; nevertheless, some component of the shell must be radiating sound. One possibility could be the structure supporting the shell. The drop of epoxy attaching the shell to the fishing line could be oscillating transverse to the surface and thereby radiating sound. It is also possible that inhomogeneities within the shell could produce some radial motion of the surface. The shell consists of two hemispheres bonded together and this seam could create a region where the material properties are different than the rest of the shell. This could provide a means of coupling the rotational and radial motions. Neither of these possibilities could be quantitatively tested, however, and the radiation mechanism was not determined.

The high quality factor of the torsional modes is best illustrated by viewing the response to tone burst excitation of the coil current where the frequency of the burst is adjusted to the T_2 or T_3 mode frequency. Figure 7 shows the response of the T_2 mode excited as shown in Fig. 5(a). After the 300-cycle 59.98-kHz burst is complete, the acoustic signal

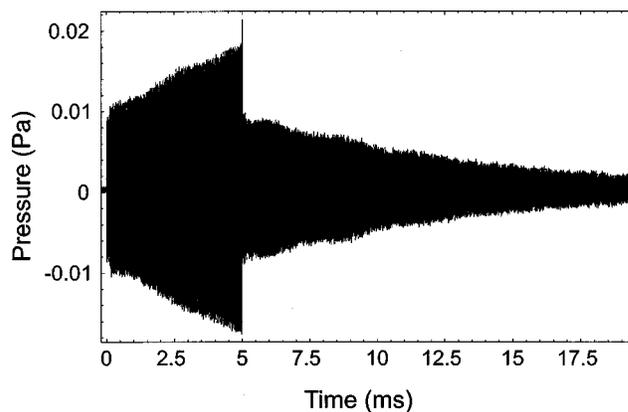


FIG. 7. Response to a 300-cycle tone burst illustrating the remarkably slow free decay of the T_2 mode excited on the shell in water as shown in Fig. 5(a).

decays exponentially with a time constant $\tau=9.046$ ms. The quality factor is $Q=\omega\tau/2=1704$. Two of the dissipation mechanisms contributing to the decay of the mode were the viscosity of water and acoustic radiation. The influence of acoustic radiation on the Q was confirmed by altering the radiation efficiency. When a drop of epoxy was placed at one of the “poles” of the shell, i.e., on the same axis as that of the coil, the damping of the mode increased slightly and $Q=1571$. When the epoxy drop was placed 45 degrees off axis, between the support and the coil, the damping increased significantly giving $Q=890$ and indicating that the corresponding portion of the shell's surface was in significant torsional motion. These results suggest that the axis of rotation for this mode is aligned with the axis of the coil, supporting the possibility of a mechanism such as the Einstein–de-Haas effect where the torsion is around the axis of the coil. Note that in Fig. 7, during the current tone burst, the received signal is a superposition of the acoustic signature and spurious inductive coupling which vanishes at the end of the burst at 5 ms.

VII. CONCLUSIONS

Experimental evidence has been presented for the excitation of the torsional modes of empty spherical shells using an EMAT. Although the simple analysis for a conducting sphere in Sec. I explains the excitation of the quasi-flexural modes and breathing mode, other mechanisms were considered briefly in Sec. VI to account for the coupling into the torsional modes. Among the original motivations for this work was the possibility of investigating the modified spectrum for shells subjected to loading by point masses. While this application has not been developed here, it is noteworthy that the high Q torsional modes are expected to be especially sensitive to such perturbations. It was confirmed that adding a small bead of epoxy to the shell can reduce the Q of a torsional mode. The coupling mechanisms demonstrated here provide an alternative to more traditional acoustical methods described in Refs. 1 and 2 which are not well suited for investigating the torsional modes of spherical shells. They may also facilitate hybrid magnetic-acoustic classification methods for metallic objects.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research.

APPENDIX A: SPHERICAL SHELL, TORSIONAL MODES, AND MODE CLASSIFICATION

The frequencies for the torsional modes of the empty stainless steel shell were calculated by solving the boundary value problem with vanishing shear stress at both the inner and outer boundary. This has been discussed by Lapwood and Usami,¹⁵ Shah *et al.*,¹⁴ and, in the thin shell limit, by Lamb.²³ The frequencies for the free torsional vibrations of an evacuated shell of finite thickness satisfy the characteristic equation,

$$\begin{vmatrix} J(\eta) & Y(\eta) \\ J(\zeta) & Y(\zeta) \end{vmatrix} = 0, \quad (\text{A1})$$

where

$$J(\eta) = \frac{d}{d\eta} j_l(\eta) - \frac{1}{\eta} j_l(\eta), \quad (\text{A2a})$$

$$Y(\eta) = \frac{d}{d\eta} y_l(\eta) - \frac{1}{\eta} y_l(\eta). \quad (\text{A2b})$$

The arguments for the spherical Bessel functions of the first kind, j_l , and of the second kind, y_l , are given by $\eta = k_s a$ and $\zeta = k_s b$, where k_s is the shear wave number, ω/c_s for the solid, and a and b are respectively the outer and inner radii. The solutions for this equation are denoted by T_l^n , where l is the angular index and n is the radial index. For the experiments performed here, modes corresponding to $n > 0$ have frequencies higher than the range of interest so we only consider the T_l^0 modes and the index n is not shown in Table II. The lowest mode is the $l=2$ mode which consists of the opposing rotations of the two hemispheres of the shell. The areas around the poles rotate in the same direction while the midsection of the shell moves in the opposite direction in the $l=3$ mode. These two modes were found to be the most important torsional modes for the present work. The A_n modes in Table I correspond to those on the lower branch of the analogous thin-shell plot in Fig. 7.9 of Junger and Feit,²⁴ with the E_0 mode on the upper branch. In thin-shell theory, the frequencies of the torsional modes become $\omega = [(l-1) \times (l+2)G/\rho R^2]^{1/2}$, where G and ρ are the shear and modulus and density of the shell material and R is the mean radius.²⁵ Using the properties of the shell given in Table I, this approximation gives the frequencies for the T_2 and T_3 modes as 61.1 and 96.7 kHz, respectively, which compare to the values given in Table II.

¹J. Ripoche and G. Maze, "A new acoustic spectroscopy: the resonance scattering spectroscopy by the Method of Isolation and Identification of Resonances (MIIR)," in *Acoustic Resonance Scattering*, edited by H. Überall (Gordon and Breach Science, Philadelphia, PA, 1992), pp. 69–103; S. K. Numrich and H. Überall, "Scattering of Sound Pulses and the

Ring of Target Resonances," in *Physical Acoustics*, edited by A. D. Pierce and R. N. Thurston (Academic, San Diego, CA, 1992), Vol. 21, pp. 235–318.

²G. Kaduchak, C. S. Kwiatkowski, and P. L. Marston, "Measurement and interpretation of the impulse response for backscattering by a thin spherical shell using a broad-bandwidth source that is nearly acoustically transparent," *J. Acoust. Soc. Am.* **97**, 2699–2708 (1995).

³J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1998), pp. 261–262.

⁴S. G. Kargl and P. L. Marston, "Ray synthesis of lamb wave contributions to the total scattering cross section for an elastic spherical shell," *J. Acoust. Soc. Am.* **88**, 1103–1113 (1990).

⁵W. L. Johnson, S. J. Norton, F. Bendec, and R. Pless, "Ultrasonic spectroscopy of metallic spheres using electromagnetic-acoustic transduction," *J. Acoust. Soc. Am.* **91**, 2637–2642 (1992).

⁶R. B. Thompson, "Physical Principles of Measurements with EMAT Transducers," in *Physical Acoustics*, edited by A. D. Pierce and R. N. Thurston (Academic, San Diego, CA, 1988), Vol. 19, pp. 157–200.

⁷T. J. Matula and P. L. Marston, "Energy branching of a subsonic flexural wave on a plate at an air–water interface. I. Observation of the wave field near the interface and near the plate." *J. Acoust. Soc. Am.* **97**, 1389–1398 (1995).

⁸B. T. Hefner and P. L. Marston, "Electromagnetic excitation and acoustic spectroscopy of the elastic mode of an empty nearly spherical shell," *J. Acoust. Soc. Am.* **99**, 2594(A) (1996).

⁹W. R. Smythe, *Static and Dynamic Electricity*, 2nd ed. (McGraw–Hill, New York, 1950), pp. 397–399.

¹⁰J. R. Wait, "A conducting sphere in a time varying magnetic field," *Geophysics* **16**, 666–672 (1951).

¹¹A. Migliori and J. L. Sarrao, *Resonant Ultrasound Spectroscopy* (Wiley, New York, 1997).

¹²T. J. Asaki, Los Alamos National Laboratory, unpublished communication.

¹³S. G. Kargl and P. L. Marston, "Observations and modeling of the backscattering of short tone bursts from a spherical shell: Lamb wave echoes, glory, and axial reverberations," *J. Acoust. Soc. Am.* **85**, 1014–1028 (1989).

¹⁴A. H. Shah, C. V. Ramkrishnan, and S. K. Datta, "Three-dimensional and shell-theory analysis of elastic waves in a hollow sphere," *J. Appl. Mech.* **36**, 431–439 (1969).

¹⁵E. R. Lapwood and T. Usami, *Free Oscillations of the Earth* (Cambridge U. P., Cambridge, 1981), pp. 52–60.

¹⁶R. A. Flinn and P. K. Trojan, *Engineering Materials and Their Applications*, 4th ed. (Houghton Mifflin, Boston, MA, 1990), pp. 423–425.

¹⁷B. A. Kudryavtsev, V. Z. Parton, and N. A. Senik, "Magnetoelastocity," in *Applied Mechanics: Soviet Reviews*, edited by G. K. Mikhailov and V. Z. Parton (Hemisphere, New York, 1990), Vol. 2, pp. 147–224; G. Reyne, J. C. Sabonnadiere, J. L. Coulomb, and P. Brissonneau, "A survey of the main aspects of magnetic forces and mechanical behaviour of ferromagnetic materials under magnetization," *IEEE Trans. Magn.* **23**, 3765–3767 (1987).

¹⁸F. C. Moon, *Magneto-solid Mechanics* (Wiley, New York, 1984), Chap. 3, pp. 53–86.

¹⁹W. F. Brown, Jr., "Electric and Magnetic Forces: A Direct Calculation. I," *Am. J. Phys.* **19**, 290–304 (1951). See Eqs. (1.3-1), (1.3-4), and (1.3-6). Our Eq. (4) follows from integrating $\mathbf{M}_S \cdot \nabla \mathbf{B}_2$ by parts where $\nabla \mathbf{B}_2$ denotes a dyad.

²⁰A. Einstein and W. I. de Haas, "Experimenteller Nachweis Ampèrescher Molekularströme," *Verh. Dtsch. Phys. Ges.* **17**, 152–170 (1915).

²¹I. C. Heck, *Magnetic Materials and Their Applications* (Crane, Russak, New York, 1974), pp. 11–13.

²²S. J. Barnett, "Gyromagnetic and electron inertia effects," *Rev. Mod. Phys.* **7**, 129–166 (1935).

²³H. Lamb, "On the Vibrations of a Spherical Shell," *Proc. London Math. Soc.* **14**, 50–56 (1883).

²⁴M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction*, 2nd ed. (MIT, Cambridge, MA, 1986).

²⁵F. I. Niordson, "Free Vibrations of Thin Elastic Spherical Shells," *Int. J. Solids Struct.* **20**, 667–687 (1984).

Measurement and prediction of diffuse fields in structures

M. J. Evans and P. Cawley^{a)}

*Department of Mechanical Engineering, Imperial College, Exhibition Road,
London SW7 2BX, United Kingdom*

(Received 21 December 1998; accepted for publication 13 September 1999)

Measurements and predictions of the response of a variety of plain and coupled plates when excited by a point source have been reported. Tests on plain plates showed that a diffuse field was generated within the first 30 wave transits across the plate. In plates coupled by a thin ligament, around 50 wave transits across one of the substructures was required for a diffuse field to be established in the whole system. This is equivalent to about 25 transits across the whole plate so the time required to set up a diffuse field was not significantly affected, even when the ligament width was only 2.5% of the total plate width. Tests on plates bolted or adhesively bonded together showed that in both cases, although a diffuse field was established in each of the substructures, the damping was too great for a diffuse field to be set up in the whole structure. Tests on stepped plates have shown that the measured surface amplitudes in the thinner sections tend to be much larger than those in the thick sections, but that the energy in the thinner sections is only slightly larger than that in a thick section of similar plan area. The field is not diffuse in the sense that the amplitude or energy density is the same throughout the coupled structure, but the field is diffuse in each substructure. © 1999 Acoustical Society of America. [S0001-4966(99)06712-0]

PACS numbers: 43.40.At, 43.40.Hb [CBB]

INTRODUCTION

Acoustic emission (AE) techniques were first used commercially in the late 1960s for the testing of pressurized systems for the chemical and aerospace industries and have since become widely used. Applications for which AE is routinely used include quality assurance of new vessels,¹ regular short-term field evaluations to assess vessel degradation, and continuous long-term monitoring.² Hydrostatic proof tests are normally carried out on new pressure vessels to 150% of the working pressure to ensure system integrity, the acoustic emission during the test being monitored.

Acoustic emission has also been successfully used for the nondestructive testing of steel and prestressed concrete structures such as bridges, dams, and skyscrapers.^{3,4} Transducers are often mounted permanently to important points on the structure such as around joints and can be continuously monitored. The number and placement of transducers depends on the specific structure, but single transducers can be used if one particular joint in a structure is under examination.

The experimental layout can affect some or all of the measured waveform parameters in an unpredictable fashion; for example, the relative locations of the source (defect) and receiver can alter the maximum amplitude of the resulting waveform. Consequently, measurements taken on identical structures may not be directly comparable if the source location is different in each case. Therefore, current methods only give a qualitative indication of the change of state of the component rather than a quantitative indication of the absolute level of damage. Ideally, the source severity (energy released) could be measured at any location on the surface, regardless of the source location. For this to be possible the

initial source energy must be evenly spread throughout the structure; in other words, the wave field within the structure must become diffuse.

A diffuse wave field is an enclosed area in which the wave energy is evenly distributed, i.e., the amplitude and directional distribution of the waves throughout the enclosure is random and the waves are uncorrelated with respect to phase (i.e., no standing waves exist). If these requirements are fulfilled, a greatly simplified description of the wave field can be used. The concept of a wave becomes of minor importance and the wave field can be treated as a whole in a stochastic manner. A diffuse field is guaranteed to occur if all boundaries within the enclosure are diffuse reflectors; totally diffuse reflectors obey Lambert's cosine law whereby the distribution of the reflected wave field is independent of the original angle of the incident wave.⁵ In practice no boundary exhibits perfect diffuse reflection but rough and irregular surfaces may be a reasonable approximation if the wavelength is of similar dimensions to the surface features.

A diffuse field cannot strictly be generated in an enclosure if all the boundaries are specular reflectors, but in certain cases the fields generated are approximately diffuse. If the source is nondirectional (small with respect to the wavelength) and excites waves over a finite bandwidth (rather than a single frequency), all normal modes (or structural modes) of the enclosure which have natural frequencies within the excitation bandwidth can be excited simultaneously. Each normal mode, if excited individually, is clearly not diffuse; the amplitude distribution is not random throughout the enclosure due to the nodes and anti-nodes of the mode. However, in the case where many modes are excited over a finite frequency band, the motion at any point within the enclosure is a superposition of all the modes present; thus, as the number of modes increases, the overall

^{a)}Electronic mail: p.cawley@ic.ac.uk

motion pattern becomes more uniform and the diffuse field approximation becomes valid. The modes referred to here are the normal modes of the structure which can be thought of as standing wave patterns produced at particular frequencies due to the reverberation of traveling waves in the structure. There are many different types of traveling wave which can produce structural resonances. In thin plate structures, most structural resonances are due to the reverberation of Lamb or shear horizontal (SH) waves, while in thicker structures, bulk waves (longitudinal or shear) and surface waves dominate the behavior. The different branches of Lamb and SH waves are commonly referred to as modes; these modes of traveling wave should not be confused with the normal modes of the structure. Diffuse field approximations have been used for many years for room acoustics and geometrical optics.⁵

Research has previously been carried out into the use of diffuse field approximations for acoustic emission. The first recorded work in this area was by Egle⁶ whose aim was to calculate the initial power, spectral content, and decay rate of an AE event from ring-down counting data. To do this he assumed that the signal was a product of a slowly decaying function and a stationary Gaussian random process (this requires the field to be diffuse). He found that it is possible to extract the mean square stationary component which represents the initial power and the zero-crossing rate which gives some idea of the frequency content of the signal. He then extended the diffuse field theory of room acoustics to solid structures having two propagating modes to calculate the proportion of energy in these modes.⁷ He assumed that the field generated in a solid is an isotropic random superposition of plane waves and only longitudinal and transverse waves exist (ignoring Rayleigh and Lamb waves). He numerically predicted the flow of energy between the transverse and longitudinal modes caused by mode conversion at the boundaries and predicted that after a calculable time, t^* , the energy present in these modes reaches an equilibrium state. He also predicted that, at equilibrium, the energy present in the transverse mode in an aluminum structure accounts for 97% of the total energy in the system regardless of the initial conditions, the ratio of energies between the longitudinal and transverse modes being dependent only on the Poisson's ratio of the material.

Weaver⁸ analytically reinterpreted and extended the theories of Egle. He assumed that if a structure is excited over a finite frequency band, all structural modes (with natural frequencies within the excitation band) are excited with equal energy. He termed this "The equipartition assumption" which has strong parallels with the statistical energy analysis assumptions (see below). Weaver then subdivided the structural modes into two categories, one for modes which consist of predominantly transverse particle displacements and the other for longitudinal particle displacements. Thus the ratio of energy densities between the transverse and longitudinal motion present is equal to the ratio of the number of transverse to longitudinal modes within the frequency band of interest. He concluded, as Egle, that the equilibrium state is predominately transverse for a steady-state diffuse field in a solid. He also predicted, however, that transducers

mounted on the surface of the structure will be more sensitive to the surface wave field than to either the bulk transverse or bulk longitudinal fields. This is extremely important since virtually all AE transducers measure the surface response.

The relationship between this measurable "surface energy density" and the total energy density was explored⁹ for finite plates. The main assumptions made were that the time domain was long compared to the acoustic transit time across the structure (this allows many randomizing reflections), that the normal modes of the plate can be approximated by standing waves of Rayleigh, shear horizontal, and Lamb types (referred to by Weaver as "Pseudo Modes"), and that the frequency range is large compared to the modal density of the plate. He concluded that a diffuse field in a finite plate will partition its time-averaged energy among the several propagating branches with a weighting independent of source nature and the total spectral energy density can be deduced from a measure of the surface energy density. This theory was later expanded to a half-space¹⁰ where the participation of the surface in the general disturbance was predicted. A term \bar{G} was introduced describing the linear proportionality between the total energy in the body and the mean square displacement at position x_0 on the surface. This value was, however, very difficult to calculate for all but the simplest geometries.

An experimental study¹¹ was carried out to verify the existence of a diffuse field in an aluminum plate. This plate had many randomly oriented saw cuts at the edges which were intended to aid the generation of a diffuse field. Weaver concluded that a diffuse field can be generated long before any appreciable absorption occurs in the system used and the results agreed well with predictions from Ref. 9.

In Ref. 12 an overview of the diffuse field method is given, together with a description of the necessary conditions concerning bandwidth and modal density for a diffuse field to be sustained:

$$\bar{l}\alpha/c \ll 1, \quad (1)$$

$$D_f \Delta f \gg 1, \quad (2)$$

where \bar{l} is the mean free path length (m), α is the material absorption (Np/s), c is the acoustic wave speed of interest (m/s), Δf is the frequency bandwidth (Hz), and D_f is the modal density. For plate structures the modal density is given by

$$D_f \approx A_p/hc, \quad (3)$$

where A_p is the plan area (m²) and h is the thickness (m).¹³ These guidelines can be used to assess whether a given structure is likely to behave in a diffuse manner. It should be emphasized that the relationships (1)–(3) are severe approximations since they assume a constant wave speed, so neglecting dispersion which is strong for the fundamental anti-symmetric (A_0) mode at low frequency-thickness products. They also implicitly assume that only one propagating mode is present. Nevertheless, the relationships give a good indication of whether a field is likely to be diffuse, particularly

when the bandwidth is modest so dispersion is not very strong.

A comparison of the diffuse field and Green's function approaches to source characterization is given by Clough.^{14,15} He argues that a diffuse field should be produced after only three reflections, although no proof is given. A demonstration that high-pass filtering is required to remove the dominating effect of the lower-order structural modes is also furnished.

It has therefore been demonstrated that diffuse fields exist in a randomized ideal structure under laboratory conditions. It has been predicted that in a well developed, steady-state diffuse field, energy partitioning between modes is controlled by the material properties rather than the specific geometry of the structure (given sufficient structural asymmetry). A structure capable of sustaining a diffuse field must be lightly damped, allowing many reflections of the initial wave energy. The number of reflections required will depend on the geometry of the structure and the directivity of the initial source of energy. Previous authors have assumed a value of 20 reflections to be sufficient,⁷ but others have suggested that as few as 3 reflections are necessary.¹⁵

The number of resonant modes (modal density) within the excitation bandwidth is extremely important for diffuse field approximations to be valid. Small structures will have a low modal density, increasing the likelihood of individual modes dominating the structural response. This problem can be reduced if the bandwidth of the signal is broad and the initial source is nondirectional. Large structures will have high modal density but the initial energy will undergo fewer reflections before being damped away; thus the field may not have a chance to become diffuse.

It is not clear from the literature, however, whether a diffuse field can be sustained in a real structure, for example, a structure consisting of several acoustically connected substructures (as will be the case in the vast majority of applications in the field). This paper explores these issues both theoretically and experimentally.

I. PREDICTION OF DIFFUSE FIELDS IN SIMPLE COUPLED SYSTEMS

A. Background

Statistical energy analysis (SEA) is a method by which the acoustic response of a structure is considered as the balance of energy flow between groups of resonant modes. Energy terms are regarded as statistical averages of resonant mode response over finite frequency bands. Statistical methods are most suitable in cases where the frequency is high and many structural modes exist within the bandwidth of interest (high modal density). In these situations traditional modal analysis methods are cumbersome and often impossible to use due to the many hundreds of modes present.

Statistical energy analysis assumes that the detailed modal behavior averages out as the modes become less well defined, thus allowing the structure to be analyzed in a more pragmatic way. Much has been written concerning the use of SEA since its original formulation.¹⁶ An excellent introduction to the subject is given by Woodhouse,^{17,18} in which the

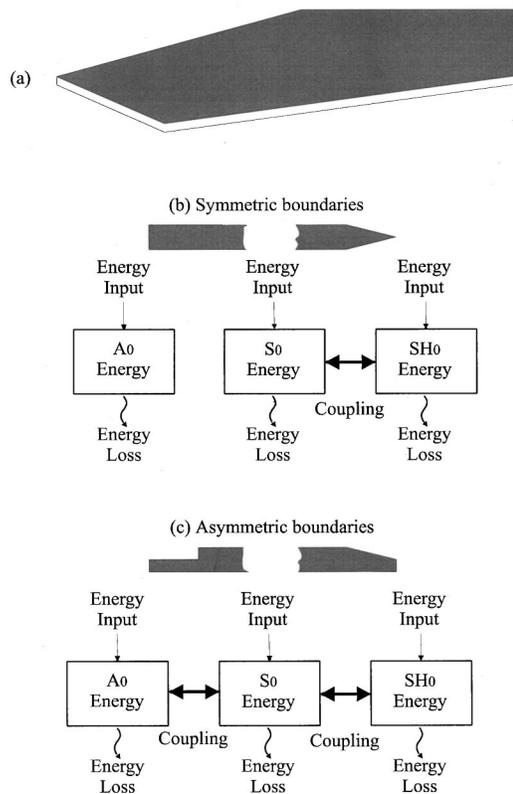


FIG. 1. (a) Simple plate structure; (b) symmetric boundaries produce no coupling between the A_0 and the S_0 and SH_0 fields; (c) asymmetric boundaries couple all the fields.

method, its major assumptions, and limitations are introduced. The early work on SEA was concerned with steady-state predictions where the source of energy was continuous. More recently, transient sources have been examined.^{19,20} It is this transient analysis which will be used to predict the response of structures to simulated acoustic emission sources.

Before discussing the use of SEA to predict diffuse fields, it is necessary to define some commonly used terms which have specific meanings in this field:

Structure: physical assemblage of components which are under analysis.

Substructure: individual component in the structure.

Field: subdivision used for the SEA model which refers to the energy in a single propagating mode type in a single substructure.

Coupling loss factor: coupling between connected fields (not including losses due to internal and external damping).

Damping loss factor: losses in an individual field due to internal and external damping (not including losses to connected fields).

Complex structures are divided into suitable fields; these need not be restricted to physical substructures but can include mode types. For example, consider a plate of uniform thickness and random shape [see Fig. 1(a)]. We make the initial assumption that only two Lamb wave modes can propagate, the fundamental symmetric and antisymmetric modes (S_0 and A_0 , respectively), together with the fundamental shear horizontal (SH_0) mode. This will be the case at

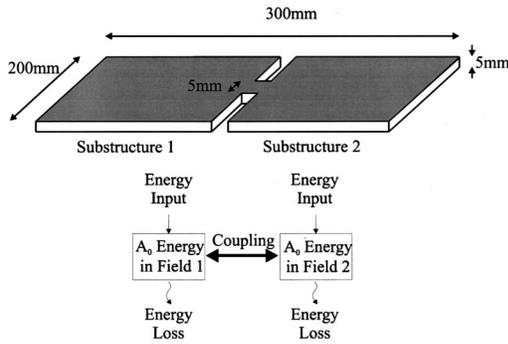


FIG. 2. Geometry and SEA model for simple plate system with no mode conversion.

frequencies below the cutoff frequency of the first higher-order modes (A_1 and SH_1). A SEA model of this structure would consist of three fields, one for each mode type in the structure. Two possible models are schematically represented in Fig. 1(b) and (c), representing the two possible edge conditions (symmetric and asymmetric). The coupling and damping loss factors are at this stage unknowns and will not, in general, be the same for each field, the damping being governed primarily by the mode shape and the material properties of the plate and surrounding media, and the coupling factors being governed by mode conversion between the modes. In the symmetric case, there is no mode conversion between the symmetric S_0 and SH_0 modes. (When waves are normally incident on a symmetric boundary, there is no mode conversion between S_0 and SH_0 ; however, at other incident angles, this mode conversion does occur. Since in a diffuse field waves are incident at all angles, there is coupling between the S_0 and SH_0 fields.) In the asymmetric case, there is mode conversion between all the modes. Hence, in the symmetric case [Fig. 1(b)] there is no coupling between the A_0 and the other fields, while all three fields are coupled in the asymmetric case [Fig. 1(c)]. This is discussed in more detail below.

The analysis of coupled structures is essential if diffuse field methods are to be used on real structures in which several types of coupling are routinely present, for example, joints, abrupt bends, and changes in section. Although these coupling types may differ physically, for the purposes of SEA they may all be treated in the same manner. The energy flow between structures is governed by the coupling loss factors and the difference in energy level between them.

B. Exact and iterative SEA models of two-field systems

A simple example is that of a plate shown in Fig. 2. A single plate has side notches cut in it which reduce the cross-sectional area in the middle portion. The remaining ligament is acting as the coupling element between the two plane regions. A SEA model of such a system would consist of two fields pertaining to the two plane sections with a coupling loss factor between them.

The coupling factors for this system are easily approximated. The through-thickness symmetry of the system means

that there is no mode conversion between the symmetric and antisymmetric modes. Coupling between the plates is assumed to be proportional to the change in width of the plate; therefore, if the remaining ligament is 10% of the width of the entire plate, the coupling will be 10% and so on. The analysis below is based on that presented in Refs. 17–20.

Taking the solution for a single mode, say A_0 , the fundamental antisymmetric mode, the energy balance equations for fields 1 and 2 are

$$P_1 = \frac{dE_1}{dt} + \eta_{12}\bar{n}E_1 + \eta_1E_1 - \eta_{21}\bar{n}E_2 \quad (4)$$

$$P_2 = \frac{dE_2}{dt} + \eta_{21}\bar{n}E_2 + \eta_2E_2 - \eta_{12}\bar{n}E_1,$$

where \bar{n} is the mean reflection frequency (Hz), E is the mean energy in each field, P_i is the steady-state power input to field i , η_{ij} is the coupling loss factor from i to j per reflection, and η_i is the damping loss in field i per second. In this case, the fields are the same size and so the reflection frequency \bar{n} is identical for both systems and is given by

$$\bar{n} = \frac{c}{\bar{l}} \quad (5)$$

where \bar{l} is the mean-free path which, for plates, is given by

$$\bar{l} = \sqrt{A_p}, \quad (6)$$

where A_p is the plan area of the plate. Equations for calculating the mean-free path for other geometries can be found in Ref. 5.

For impulsive excitation P_1 and $P_2 = 0$. Introducing a differential operator D and writing Eq. (4) in matrix format we obtain

$$\begin{bmatrix} D + \eta_a\bar{n} & -\eta_{21}\bar{n} \\ -\eta_{12}\bar{n} & D + \eta_b\bar{n} \end{bmatrix} \begin{Bmatrix} E_1 \\ E_2 \end{Bmatrix} = 0, \quad (7)$$

where $D = d/dt$, $\eta_a = \eta_1 + \eta_{12}$, and $\eta_b = \eta_2 + \eta_{21}$.

The determinant of this equation is

$$D^2 + (\eta_a + \eta_b)\bar{n}D + (\eta_a\eta_b - \eta_{12}\eta_{21})\bar{n}^2 = 0. \quad (8)$$

A transient energy, E_0 , is applied to substructure 1 by setting the initial conditions

$$E_1(0) = E_0, \quad E_2(0) = 0, \quad (9)$$

which when substituted in (4) gives

$$\begin{aligned} dE_1(0)/dt &= -\eta_a\bar{n}E_0, \\ dE_2(0)/dt &= \eta_{12}\bar{n}E_0. \end{aligned} \quad (10)$$

Solving Eq. (8) and applying the initial conditions yields

$$\begin{aligned} E_1(t) &= \frac{E_0}{2b} e^{-a\bar{n}t} \left[\left(\frac{-D_2}{\bar{n}} - \eta_a \right) e^{b\bar{n}t} + \left(\frac{D_1}{\bar{n}} + \eta_a \right) e^{-b\bar{n}t} \right], \\ E_2(t) &= \frac{E_0}{2b} \eta_{12} e^{-a\bar{n}t} [e^{b\bar{n}t} - e^{-b\bar{n}t}], \end{aligned} \quad (11)$$

where $a = -(\eta_a + \eta_b)/2$, $b = \frac{1}{2}\sqrt{(\eta_a - \eta_b)^2 + 4\eta_{12}\eta_{21}}$, and $D_1, D_2 = -\bar{n}(a \pm b)$.

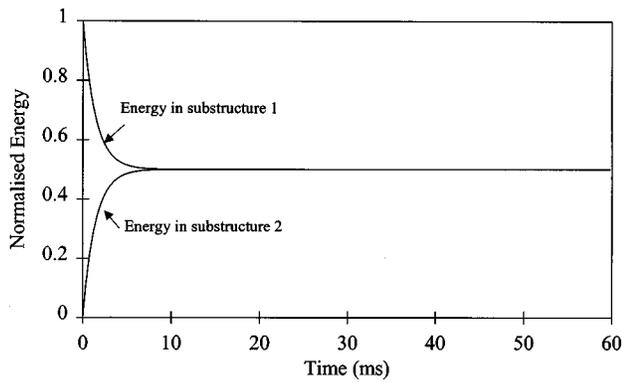


FIG. 3. Predicted response of system of Fig. 2 with no damping.

When dealing with structures having multiple sub-systems it is often easier to use an iterative solution.²¹ A simple scheme has been employed, the equilibrium equations for systems 1 and 2 being written as

$$\begin{aligned} E_1(t + \Delta t) &= E_1(t) - \eta_{12}\bar{n}\Delta t E_1 + \eta_{21}\bar{n}\Delta t E_2 - \eta_1\Delta t E_1, \\ E_2(t + \Delta t) &= E_2(t) - \eta_{21}\bar{n}\Delta t E_2 + \eta_{12}\bar{n}\Delta t E_1 - \eta_2\Delta t E_2. \end{aligned} \quad (12)$$

The energy at $t + \Delta t$ is simply equal to the energy at time t plus the sum of incoming and outgoing energy in time Δt . Predictions obtained using the exact solution of Eq. (10) were compared with the iterative solution of Eq. (11) for the same boundary and initial conditions and showed excellent agreement. It may be more accurate to use a higher-order iterative scheme, but the comparisons with the exact solution indicated that the simple two term Taylor scheme of Eq. (11) was adequate.

Figure 3 shows the predicted response of the structure of Fig. 2 when the coupling loss factors were 2.5% (3.5 dB/ms) and the damping loss factors were zero. This figure clearly shows a steady-state partitioning of energy between the fields being reached after approximately 6 ms. (In this paper, the steady-state condition is defined as being when the decay rates in all fields are equal to within 5%.) It is interesting to note that in this simple case the steady-state ratio of energy is 1:1 as the fields are identical. In the absence of damping the steady-state partitioning of energy will always occur, although, given very low coupling, this may take an extremely long time. We can therefore conclude that in such a system coupling can always be considered to be strong.

Zero damping will never exist in practice and therefore the amount of time available for the partitioning of energy to occur is limited by the reverberation time of the structure. The strength of coupling is therefore affected by the damping loss factors.

C. Coupling strength

Coupling strength describes the ratio of the energy exchanged by coupling between fields and the energy lost due to damping. The coupling between fields encompasses both the coupling between substructures, as demonstrated in the previous example, and the coupling between mode types, for example the S_0 and A_0 mode fields in a plane plate. For

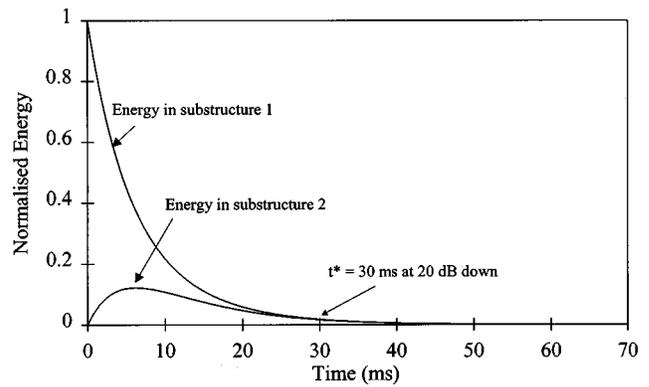


FIG. 4. Predicted response of weakly coupled system.

simplicity, consider first a single structure in which only two propagating modes are present, giving three possible scenarios:

No coupling: In the example of Fig. 1(b), there is no coupling between the A_0 field and the S_0 and SH_0 fields. The A_0 field may be diffuse, but there is no energy flow between it and the other fields. The ratio of energy between uncoupled wave fields is governed by the initial source partitioning and the attenuation of each field. The decay rates of uncoupled fields will, in general, be different due to the differences in attenuation between the propagating modes.

Weak coupling: Weak coupling in the case of Fig. 2 could occur if there was significant damping in the plates. In this work, the coupling is defined as being weak when the coupling loss afforded by the mode conversion at the edges is lower than the damping losses in the structure. Figure 4 gives an example of the response of a weakly coupled system with a damping loss factor of 1 dB/ms and a coupling loss factor of 0.5 dB/ms. The steady-state partitioning of energy is reached (decay rates identical to within 5%) eventually, but not before the initial signal is attenuated by more than 20 dB. Decreasing the coupling loss factor would further delay the time at which the steady state occurs.

Strong coupling: If mode conversion is strong, for example, in a highly asymmetric system, energy transfer will occur very quickly. Thus the coupling loss factors will be larger than the damping losses in the system and a steady state will occur even in the presence of damping. Figure 5 demonstrates the response of a strongly coupled system with a damping loss factor of 0.5 dB/ms and a coupling loss factor

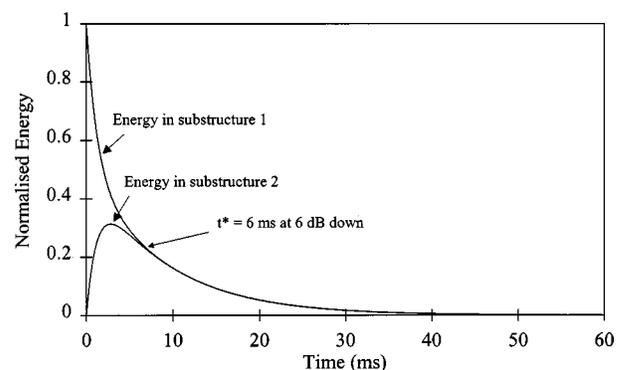


FIG. 5. Predicted response of strongly coupled system.

of 1 dB/ms. This shows the steady state occurring before severe damping losses are seen (after an energy drop of 6 dB). Further increase of the coupling loss factor would make this steady state occur sooner.

II. MEASUREMENT TECHNIQUE

A. Wave generation

The acoustic source used throughout the experimental work was a conical piezoelectric transducer similar to that described by Proctor.²² Using a piezoelectric device allows the experimenter a great deal of control over the amplitude, frequency content, and position of the source. The shot-to-shot repeatability of these devices is excellent, which permits signal averaging and direct comparison of signals measured from different locations on the structure.

The source transducer was driven by applying a carefully controlled electrical signal in the form of a windowed tone burst. The tone burst signals were generated using a wave-packet generator which allowed control over the amplitude, center frequency, and bandwidth of the signals, the bandwidth being controlled by varying the number of cycles in the tone burst and the window employed. The conical transducers have been shown to be sensitive to out-of-plane motion, thus due to reciprocity, they will generate only out-of-plane force on the material surface. The contact area of the conical transducers is 1 mm², which is small with respect to the smallest wavelength used (around 6 mm).

The structures used throughout the experiments were plates which, in the frequency range used, restricts the propagating modes present to Lamb and SH modes. The bandwidth of the excitation signal is therefore extremely important as it determines the number of modes which can exist and the amount of velocity dispersion. All of the experiments were carried out within the frequency range in which only the fundamental modes S_0 , SH_0 , and A_0 exist. However, as the conical transducers are only sensitive to out-of-plane motion (and only generate out-of-plane force), they are not well suited to the generation or reception of S_0 or SH_0 since the motion in the S_0 mode is predominantly in-plane in the frequency-thickness range used and that in SH_0 is entirely in-plane. To quantify this, a finite-element model was used to predict the propagating waves generated by the source on an aluminum plate with a five-cycle tone burst at a center frequency thickness of 0.75 MHz-mm. As expected, only the fundamental modes S_0 and A_0 were generated and the ratio of A_0 to S_0 energy generated by the source was predicted to be 30:1. No SH_0 was generated. Assuming reciprocity, the transducers will be 30 times more sensitive to the reception of A_0 than S_0 and the combined transmit/receive response is the square of this ratio, giving a relative sensitivity of 900:1.

B. Reception, amplification, and sampling

A second conical transducer was used to detect the ultrasonic field generated in the structure. The signal from this transducer was amplified using a 40-dB, wide-band preamplifier and filtered using a 50-kHz, two-pole, high-pass filter. The signal was then rectified and envelope detected using an analog circuit with a 1-ms time constant before sampling, as

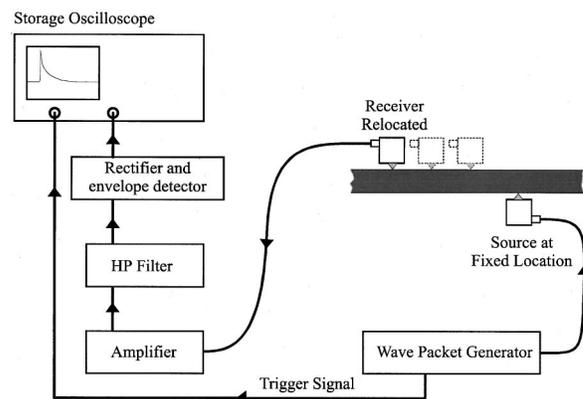


FIG. 6. Experimental setup.

shown in Fig. 6. The envelope detection greatly reduces the bandwidth of the signal, allowing it to be sampled at a lower frequency while preserving the signal envelope. A LeCroy digital oscilloscope was used to sample the demodulated signal, the sampling period required to capture the entire signal envelope being up to 100 ms for some structures.

Given the relative sensitivity of the transducer to the A_0 mode, it is reasonable to assume that the signals measured by the conical transducer are governed entirely by the behavior of the A_0 field within the structure.

C. Signal processing

Three parameters have been measured from the demodulated signal envelopes in order to allow a quantitative comparison of the characteristics of signals measured at different locations on the structure, as shown in Fig. 7.

Maximum amplitude: The maximum amplitude was measured directly from the signal envelopes. The typical shape of the envelope shown in Fig. 7 means that the maximum amplitude is always near the beginning of the signal.

Envelope area: The envelope area was calculated by integrating the signal envelope between fixed starting and ending times. The beginning of the signal was taken to be time zero (the time at which the source was activated) and the end of the signal was defined as the time at which the signal decayed by 60-dB from its maximum amplitude. The time interval between these two points is known as the reverberation time.⁵ Initially the envelope area was measured for the entire reverberation time; however, in certain cases it was found to be advantageous to delay the envelope area mea-

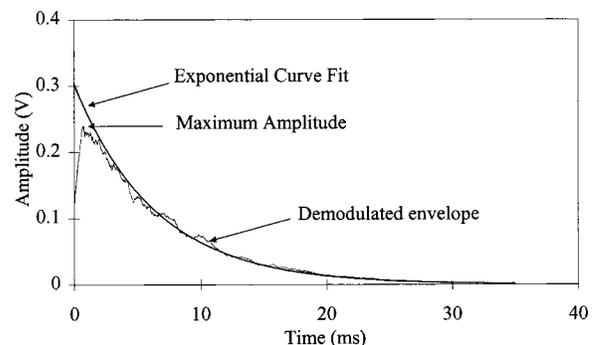


FIG. 7. Signal processing applied to demodulated envelope.

surement until the field has had time to become diffuse. The improvements to the diffuse field measurements gained by using a delay are demonstrated during the discussion of the large plate tests (see below).

Decay rate: The decay rate was measured by fitting an exponential function to the decaying envelopes; a similar technique has been used for previous diffuse field measurements¹¹ and for the measurement of structural damping.²³ As recommended in this work, the beginning and end of the signal envelopes were ignored and the exponential fit was carried out using the midsection of the signal; the first few milliseconds of the signal show the randomizing of the field and consequently the decay in this region does not follow an exponential curve.

D. Measurement of energy density

During all diffuse field measurements the important field parameter is the energy, since diffuse fields are always defined in terms of energies. It is often assumed that the energy of an AE signal is proportional to the area under the squared and demodulated signal.^{15,24} The envelope signals measured can be closely represented as an exponential decay function of the form

$$x(t) = Xe^{-\alpha t}, \quad (13)$$

where X is the initial amplitude, α is the decay rate, and t is the time. Squaring and integrating with respect to time yields

$$\int x(t)^2 dt = \frac{X^2}{2\alpha} e^{-2\alpha t}, \quad (14)$$

which is assumed to be proportional to energy. By performing the operations in reverse order we can see a similar result,

$$\left(\int x(t) dt \right)^2 = \left(\frac{X}{\alpha} \right)^2 e^{-2\alpha t}. \quad (15)$$

This result shows that both methods yield results which are of the same form, differing only in the constant of proportionality. The energy in a signal is, therefore, simply proportional to the square of the envelope area and this value will be termed the envelope energy. There are several advantages to be gained by squaring the signal after the envelope area measurement has been taken: the signal processing is simpler and the original amplitude and decay rate of the measured envelopes are preserved.

The proportionality between envelope energy and the actual energy in the structure has been demonstrated experimentally. This was done by mounting the source and receiver on a 5-mm-thick aluminum plate of plan dimensions $0.2 \times 0.3 \text{ m}^2$. The source amplitude was varied by known amounts by altering the peak voltage supplied by the wave packet generator. (The source energy is proportional to the square of the source amplitude.) An envelope energy measurement was then taken at each source amplitude level and the results recorded. Figure 8 shows the envelope energy plotted against the energy input showing a linear relationship

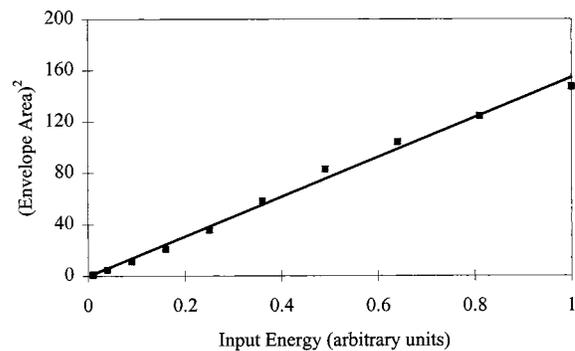


FIG. 8. Envelope area squared as a function of input energy (square of excitation voltage).

(the solid line is a least squares fit to the measured data). This demonstrates that the envelope energy is indeed proportional to the energy in the structure.

A second experiment was carried out to measure the effect of the size of the structure on the measured envelope energy. Five 5-mm-thick rectangular aluminum plates were used; each plate had a different area and aspect ratio. The source transducer was mounted on each structure and driven with a five-cycle, Hanning windowed tone burst at 150-kHz center frequency. The amplitude of the excitation signal remained constant for each experiment (200 V peak-peak), ensuring that the source energy for all the experiments was the same.

Ten measurements were taken at random locations on each of the five structures and the envelope energies measured and averaged for each structure. The envelope energy was found to vary inversely with the plan area. This indicates that the envelope energy is a measure of the energy density in the structure (energy per square meter). If this is the case, it should be possible to predict the source energy for each of the experiments on the five plates. This was done by multiplying the envelope energy by the plan area for each plate geometry, which should result in the same value for each experiment. Figure 9 shows the predicted source energy for each plate plotted against plan area, the solid line being the mean source energy calculated from all the predictions and the error bars representing one standard deviation. The results show reasonable correlation, indicating that envelope energy is a reasonable indication of the energy density in the structures tested.

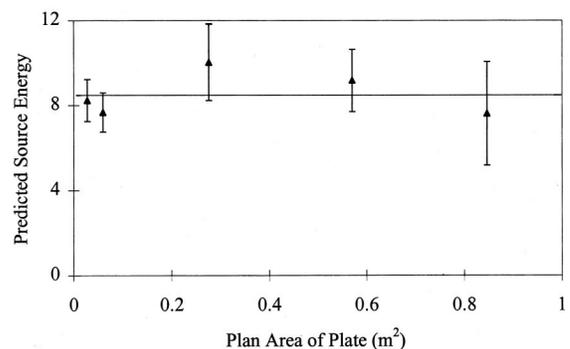


FIG. 9. Predicted source energy (arbitrary units) versus plan area of structure. Error bars indicate one standard deviation.

III. EXPERIMENTAL CONFIRMATION OF DIFFUSE FIELD GENERATION IN PLANE PLATES

A. Background

The diffuse field approximation asserts that the initial source energy will become evenly distributed throughout the structure, i.e., the energy field will become homogeneous. The measured envelope area extracted from the signal envelopes has been shown to be indicative of the signal amplitude. Determining whether a field is diffuse is then simply a case of mapping the variation of the envelope area across the surface of the structure.

Initial experiments were carried out on very simple structures, the intention being to excite these structures using a transient force and to measure the subsequent field across the entire surface of the structure to assess the diffuse field approximation. The excitation signal was a five-cycle, Hanning windowed tone burst with a center frequency of 150 kHz having a bandwidth (20 dB down) of 100 kHz.

It is possible to assess whether structures are likely to sustain a diffuse field by using the conditions of Eqs. (1)–(3). Equation (1) simply asserts that for a diffuse field to be generated, the attenuation in signal caused by one plate transit must be much less than 1 Np (8.7 dB) and Eq. (2) states that the number of structural modes of vibration within the excitation bandwidth must be large. These equations are guidelines only and do not stipulate any cutoff values of attenuation or bandwidth, but they do provide a useful method of assessing a structure objectively for its diffuse field properties. Some recommendations are given below regarding the cutoff value of attenuation for diffuse field generation to occur.

The experiments described here to determine the field generated by a source on a structure have been carried out using a single receiver. The receiver was moved to many locations on the surface of the structure and measurements were taken at each location. A major shortcoming of this technique, as opposed to using a large number of receivers sampling simultaneously, is that repositioning the receiver may affect its response and the response of the structure under test. These effects are quantified in the next section.

B. Measurement of experimental errors

Before assessing the variation of signal strengths measured at different locations on a structure it is first important to get an indication of the errors associated with each step of the measurement protocol. The major errors are expected to be due to random variations of receiver coupling and structural damping caused by the transducers themselves being removed and relocated.^{11,25}

In order to investigate the effect of each of these uncertainties on the measured parameters a series of experiments was carried out, each experiment attempting to isolate a single variable. A description of each experiment follows and the results are presented in Table I.

Recoupling: The first experiment was designed to measure the effect of removal and recoupling of the transducer. It is well known that contact transducers suffer from random variations in their sensitivity which is attributed to

TABLE I. Experimental errors expressed as the ratio of the standard deviation of the mean value for each parameter (%).

	Maximum amplitude	Envelope area	Decay rate
Recoupling	6.2%	3.2%	5.2%
Damping	1.7%	2.0%	4.5%

coupling.^{26,27} Coupling variations are thought to be due to changes in the contact area or angle of the transducer (if the surface is not completely flat). In order to examine this effect the receiving transducer was systematically removed and recoupled at exactly the same location on the structure. This procedure was repeated ten times and the received signal was measured for each coupling; the averages and standard deviations of the measured parameters were then calculated, see Table I.

Damping: The effect of transducer relocation on the damping present in the system was then measured. This was done by placing an additional “dummy” transducer onto the structure which was removed and recoupled at ten randomly chosen locations. Each time the dummy transducer is recoupled it is subject to the same coupling errors previously discussed and these affect the amount of damping caused by the transducer. This damping has been shown to be governed by the contact area of the transducer.²⁵ Since the transducer introduces asymmetry into the system, there will be some mode conversion as waves pass the transducer location. This would contribute to the apparent damping, though it is likely that the major effect is mechanical damping, rather than mode conversion.

The results from the first experiment show that the measured parameters are fairly unaffected by transducer removal and recoupling (for the transducer setup used). It is interesting to note that the area measurements show better repeatability than maximum amplitude. The maximum amplitude measurement can be taken to be a direct indication of coupling efficiency (since decay rate changes will have minimal effect at the beginning of the signal envelope). Area measurements on the other hand will be affected by both the coupling efficiency and the decay rate; however, these effects are contradictory (poor coupling efficiency will reduce overall amplitude and decrease the decay rate). The combination of these effects results in an improvement in repeatability. Another possible reason for the increased repeatability of the area measurements compared to the maximum amplitude is that the maximum amplitude results from interference between waves traveling in different directions and small changes in transducer position could cause large changes in the peak amplitude. The area measurement is effectively an average over time which would tend to reduce the statistical effect of small changes in position.

The decay rate variations measured in the dummy transducer experiment are very similar to those measured in the previous recoupling experiment. It therefore seems reasonable to assume that the variation in damping caused by the dummy transducer is directly caused by the coupling variations mentioned and is not affected by the dummy transducer location.

TABLE II. Results for plates with plan dimensions (1) $0.2 \times 0.3 \text{ m}^2$ and (2) $1 \times 1 \text{ m}^2$. Standard deviations of maximum amplitude and envelope area are expressed as a percentage of the mean value. Values in brackets show the expected standard deviations due to experimental errors.

Plate	Model density (modes/kHz)	Modes in bandwidth	Attenuation per Transit (Np)	Standard deviation maximum amplitude	Standard deviation Envelope Area
1	4	400	0.0051	16.6% (7.9%)	7.1% (5.2%)
2	66	6600	0.0206	14.8% (7.9%)	18.1% (5.2%)

It can be concluded that maximum amplitude and area measurements give consistent, repeatable results which are relatively unaffected by removal and recoupling of the receiver. The combined effect of damping and recoupling will be approximately the sum of the individual effects giving standard deviations of 7.9% for the maximum amplitude and 5.2% for the area measurements. The area measurement is favored over maximum amplitude as it is an indication of the amplitude of the entire signal envelope and will be used in the experiments that follow to assess structures for their diffuse field properties.

C. Diffuse field verification

Two 5-mm-thick aluminum plates with plan dimensions $0.2 \times 0.3 \text{ m}^2$ and $1 \times 1 \text{ m}^2$ were obtained which will be referred to as plates 1 and 2, respectively. Table II shows the modal density, number of modes in the bandwidth, and attenuation per transit for the two plates of interest when the function used to excite the source was a five-cycle, 150 kHz Hanning windowed tone burst having a bandwidth of 100 kHz (40 dB down points). The attenuation per transit was computed from the measured decay rate of the field, knowing the group velocity and the plate dimensions. The velocity of interest in these experiments is the A_0 group velocity which is 3000 m/s at 0.75 MHz-mm for an aluminum plate in vacuum.

The results shown in Table II demonstrate that both plates satisfy the criteria for diffuse field generation stated in Eqs. (4) and (5). Plate 1 has an attenuation of 0.0051 Np per transit and 400 structural modes within the excitation bandwidth while plate 2, being larger, has a higher modal density and more attenuation per transit. The attenuation is, however, still much less than 1 Np per transit.

Experiments were carried out on these plates in which the source transducer was attached to the central region of the plate on the underside. Measurements were taken at 63 locations on the top surface of the plates in a regular grid. The results from plate 1 were extremely consistent over the entire sample of 63 measurements. Standard deviations of the maximum amplitude and envelope area can be seen in Table II, the combined experimental errors discussed in the previous section being given in brackets. The standard deviation of the maximum amplitude measurements on plate 2 is slightly lower than that on plate 1. However, this change is within 2% and is not regarded as significant. More importantly, the standard deviation of the envelope area measurements is more than a factor of 2 higher on plate 2. This clearly demonstrates that the field generated in plate 2 is less uniform than in plate 1 and is, by definition, less diffuse.

One possible explanation for the large standard deviation in area measurements seen on the larger plate (plate 2) is the time taken for the diffuse field to generate. Previous work has suggested that it is necessary to delay measurement of area for a short period to allow the energy to be evenly distributed,¹¹ the initial part of the signal being assumed to account for most of the variation. This hypothesis was tested by repeating the experiment with delays of 10, 20, and 30 ms; to put this in perspective, 10 ms is equivalent to 30 wave transits across the plate and the total reverberation time for the plates was around 90 ms. The variation of the envelope area measurements is shown in Table III; the standard deviation of the maximum amplitude measurements was found not to vary appreciably with delay. The standard deviation of the area measurement is seen to reduce to 14.3% after 10 ms and then stay fairly constant; further increases in delay time were not found to affect this value significantly. This demonstrates that in this case only the first 10 ms of the signal envelope is affected by the first arrival amplitude and the gradual randomizing of the field. It can therefore be concluded that the steady-state condition is reached within the first 10 ms (30 wave transits). The results of Table III demonstrate that the larger standard deviation in the area measurements on the larger plate is not solely due to the longer time required for the field to become diffuse. Further investigation is required to discover whether the standard deviation is generally a function of plate size.

The results therefore show that the field generated in plate 1 has a high degree of uniformity, area measurements having a standard deviation of 7.1% and perhaps less if the experimental errors are taken into account. Area measurements taken on plate 2 were less uniform than for plate 1, the standard deviation being around 14.3%. The question of whether this field can still be regarded as diffuse depends on what signal variation is deemed acceptable (this will ultimately be governed by the accuracy to which the source amplitude is required to be measured). A value of 15% seems reasonable at this stage, bearing in mind that the current measurement system has an inherent error of around 5%. Using this criterion both plates behave diffusely given sufficient time for the field to develop.

TABLE III. Variation with delay of standard deviations of the envelope area measurements, expressed as a percentage of the mean value.

Delay (ms)	0	10	20	30
Standard deviation	18.1%	14.3%	14.3%	13.9%

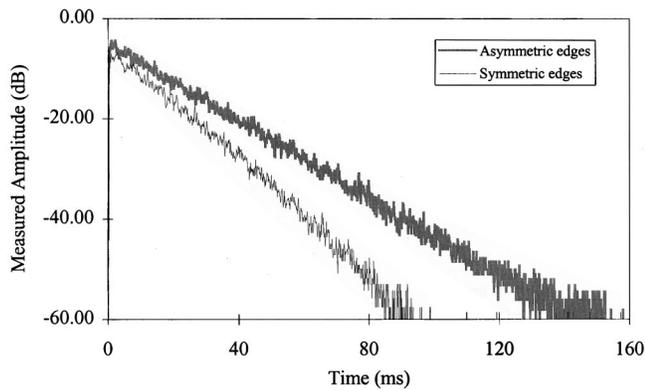


FIG. 10. Signal envelopes measured on plates with symmetric and asymmetric edges.

D. The effect of mode conversion on the decay rate

The discussion above has assumed a single propagating mode, A_0 , to be present. This seems reasonable given the relative efficiencies of generation and reception of A_0 and S_0 by the experimental equipment used. However, mode conversion between the A_0 , S_0 , and SH_0 modes will occur upon interaction with asymmetric features through the plate thickness (such as nonsquare edges), the rate at which this mode conversion takes place being governed primarily by the degree of asymmetry. In a general case, where some structural asymmetry exists, there will be A_0 , S_0 , and SH_0 fields present in the structure. The signals measured using the current experimental setup can only give information regarding the A_0 field, but the cross-coupling between the S_0 , SH_0 , and A_0 fields may significantly alter the A_0 behavior since the decay rate of the signal envelope will be governed by the attenuation of the modes present, and the steady-state ratio of these modes. The particle motion in both the S_0 and SH_0 modes in the frequency range considered here is predominantly in-plane, whereas that of A_0 is predominantly out-of-plane. Therefore the damping caused by radiation into air will be considerably higher for the A_0 mode. The material damping of the S_0 mode will be lower than that of the other two modes as the stresses in this mode are primarily direct, whereas the other modes have significant shear stresses and the shear material attenuation is generally larger than that of the compression wave. This is discussed further in Ref. 28.

During the experiments on plain plates discussed above it was initially assumed that the plate edges were square and no significant mode conversion was present. On closer inspection it was found that due to inaccurate machining the edges were visibly asymmetric, causing rapid mode conversion from the antisymmetric to symmetric modes and vice versa. The subsequent decay rate of the envelope must then be caused by the combined losses due to the S_0 , SH_0 , and A_0 fields. To demonstrate this, the plate was carefully remachined ensuring that the edges were as symmetric as possible. Typical signal envelopes measured from the original and modified plates are shown in Fig. 10. The averaged decay rate measured for ten tests on the original, slightly asymmetric, plate was 0.40 dB/ms with a standard deviation of 5%. The averaged decay rate measured for ten tests on the symmetric edged plate was 0.54 dB/ms with a standard de-

viation of 3%; this represents an increase of 26% in the overall damping of the structure. This result supports the hypothesis that if all the energy is retained in the A_0 mode, the decay rate is more rapid than if mode conversion can occur. There may also be an increase in damping caused by machining the plate edges, but this is unlikely to be very significant since, although the machining operation may have produced a region of deformed material with higher attenuation close to the new plate edge, it removed material that was deformed in the initial cutting process that produced the original plate and so may have had higher material damping. Also, the region affected by the machining covered a small fraction of the overall plate area so the net effect of the machining on the material damping is likely to be relatively small.

IV. EXPERIMENTAL MEASUREMENTS OF ENERGY SHARING IN COUPLED STRUCTURES

A. Simple coupled plates

The experiments reported so far have concentrated on plane structures without discontinuities which have exhibited broadly diffuse characteristics; however, most practical structures contain discontinuities. Any discontinuity in a structure, whether a boundary between assembled components or a sudden change in cross section, will act as an obstruction to the propagation of energy throughout the structure. Initial experiments have been carried out on the simple plate structure shown in Fig. 2. A 5-mm-thick aluminum plate with plan dimensions $0.2 \times 0.3 \text{ m}^2$ was machined to produce two plates with plan dimensions $0.2 \times 0.145 \text{ m}^2$ connected by a 5-mm-long ligament. Using the statistical energy analysis terminology discussed above, the two plate sections may be regarded as substructures, the level of coupling between them being governed by the width of the coupling element. The coupling has been assumed to be proportional to the percentage of the total plate width remaining at the ligament. In this case the coupling ligament was 2.5% of the width of the plate so the coupling factor has been assumed to be 2.5%. The edges of the plate were machined carefully so as not to allow significant mode conversion between A_0 and S_0/SH_0 and it is therefore assumed that A_0 is the only propagating mode.

Experiments were carried out to ascertain whether this structure behaves in a diffuse manner. The source transducer was permanently attached to the central region of one substructure and measurements were taken at ten locations on each substructure. The ten measured envelopes on each substructure have been compared and averaged, the averaged envelopes being shown in Fig. 11(a). The diffuse field establishment time, i.e., the time at which the two envelopes coincide and then decay together, can clearly be seen. The initial amplitude of the envelopes shows the effect of the large first arrival signals present on the near (source) side and the gradual leakage of this energy to the far side. The time taken for the diffuse field to generate in this case is around 3 ms (50 wave transits). The standard deviation of the area measurements from all 20 locations was found to be 11%, but much of this variation exists in the first 3 ms where the field

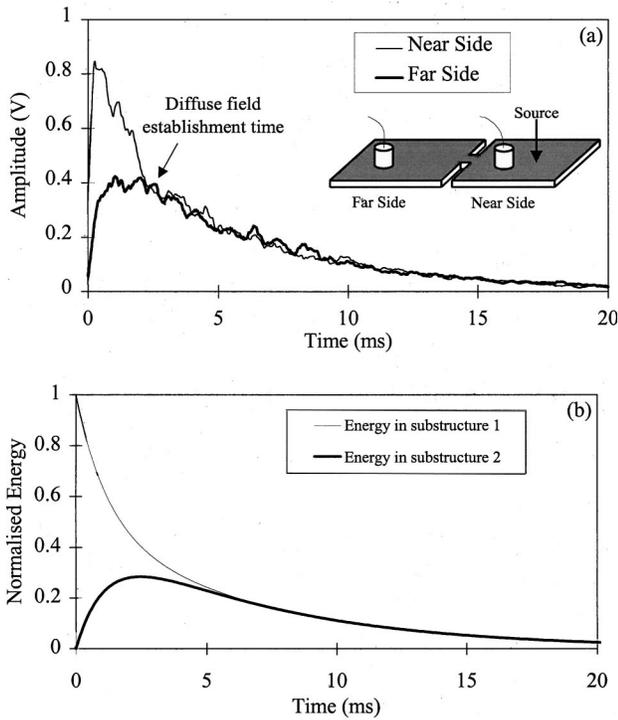


FIG. 11. Signal envelopes on plates coupled by narrow ligament: (a) averaged measurements and (b) corresponding predictions.

is not diffuse. By adding a delay of 3 ms before area measurements were taken the standard deviation dropped to 7.5%, which is similar to that measured on a plane plate of the same dimensions. The decay rate for both substructures was found to be 0.5 dB/ms with a standard deviation of 5%; this value also agrees well with the measurements previously made on the plane structures where only A_0 was present.

This simple plate case was modeled using the SEA techniques described above. The coupling factor between the substructures was assumed to be 2.5% and the measured damping factor of 0.5 dB/ms was used. The SEA model consists of two fields, one A_0 field in each substructure. The predicted response of the system using these parameters is shown in Fig. 11(b), which compares favorably with the measured response shown in Fig. 11(a). The predicted diffuse field establishment time is 6 ms rather than the measured 3 ms; this discrepancy could be caused by an underestimation of the coupling factor between the two substructures.

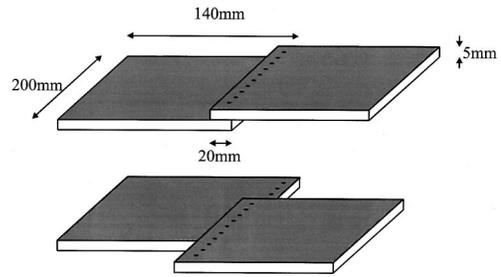


FIG. 12. Geometry of bolted plates showing two configurations with different overlapping areas.

B. Plates with bolted joints

Bolting is a common method used for joining components in assembled structures. These joints are formed by adjacent faces of the structures being forced together by the action of one or more bolts, the intimately contacting faces providing the acoustic coupling between the structures. Friction between the contacting surfaces of the joint also causes losses which contribute to the overall damping of the structure. It is this damping effect which is of most concern for the purposes of diffuse field measurements. As previously discussed, diffuse field generation relies on many tens or hundreds of reflections to spread the initial source energy evenly around the structure. An increase in damping will reduce the total number of reflections and may render the structure incapable of sustaining a diffuse field. Some previous work has concentrated on the damping effect caused by surfaces in contact, specifically to predict the energy loss caused by directly coupled transducers.²⁵ This work has shown that the damping effect caused by transducers is directly proportional to their area of contact.

An experimental structure consisting of two 5-mm-thick aluminum plates with plan dimensions $0.2 \times 0.14 \text{ m}^2$ was constructed. A row of ten equally spaced holes was drilled along one long edge on each plate and the substructures were then bolted together using M5 bolts, with a range of areas of contact, as shown in Fig. 12. The coupling between the substructures will depend on the amount of contact but, for all the possible configurations, it will be much higher than the 2.5% coupling case of Fig. 2.

Measurements were made at ten locations on each substructure for five different structural configurations. The measured parameters of interest for each configuration were the averaged decay rate, the average standard deviation of

TABLE IV. Variation of decay rate, standard deviation of envelope area (EA) measurements, envelope area difference between the near and far sides, and attenuation per transit as a function of the overlap area of bolted and adhesively bonded plates.

Joint overlap area (m^2)	Decay rate (dB/ms)		EA standard deviation (% mean)		EA difference (% mean)		Attenuation per transit (N_p)	
	Bolted	Bonded	Bolted	Bonded	Bolted	Bonded	Bolted	Bonded
0	0.53	0.53	6.3	5.8	0.005	0.005
0.0012	1.19	0.85	11.8	7.6	+18	+16	0.011	0.008
0.0020	2.44	1.24	14.1	12.3	+29	+18	0.023	0.012
0.0028	2.88	1.35	16.0	6.4	+39	+19	0.027	0.013
0.0040	3.40	1.28	23.9	22.8	+60	+59	0.032	0.012

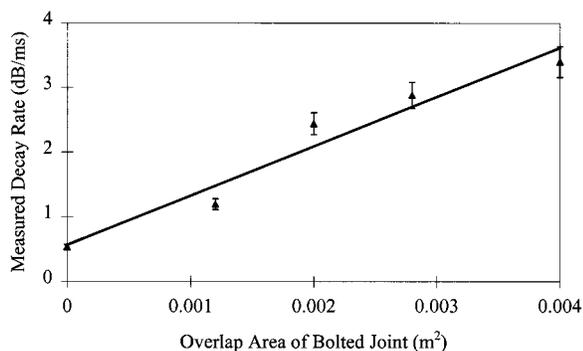


FIG. 13. Decay rate versus joint overlap area for bolted plates; error bars indicate one standard deviation.

the envelope area measurements taken across both substructures, and the difference between the averaged near (source) and far side envelope area measurements. Table IV shows this difference as a percentage of the mean value, a positive sign indicating that the measurements taken on the near side were greater. The decay rate is plotted against overlap area in Fig. 13, the error bars indicating one standard deviation and the solid line a least squares fit applied to the data. The results indicate a linear relationship between overlap area and damping.

The standard deviation of the area measurements taken on each individual substructure shows a steadily increasing trend with joint overlap area. This indicates that the field generated in the substructures is becoming less diffuse as the damping increases, as would be expected. Table IV shows the measured attenuation per wave transit across a substructure for each overlap area, taking the mean transit length as $\bar{l} = \sqrt{A_p}$, where A_p is the plan area of the substructure. Previously the criterion used to determine whether a field is diffuse was that the standard deviation of envelope area measurements taken at random locations across the structure was less than 15%. Table IV shows that this condition is only satisfied by the first three bolted configurations, i.e., overlap areas up to 0.002 m². Table IV shows that the equivalent attenuation per transit for the field to be diffuse must have an upper limit of between 0.023 and 0.027 Np/transit. This agrees well with the results obtained from the large plane plate studied in Sec. II in which the attenuation was measured to be 0.021 Np/transit, the standard deviation of envelope area measurements being 14.3%. The individual substructures are behaving diffusely (up to a joint overlap of 0.002 m²), but the question remains as to whether the structure is behaving diffusely as a whole.

The difference between the average envelope area measurements taken on either side of the joint shown in Table IV demonstrates that the energy is not equally divided between the substructures; even for the smallest of the overlap areas the difference is still 18%. A delay of 3 ms was used before envelope area measurements were taken, as described in Sec. II; increasing this delay was not found to reduce the difference in envelope area measurements. The energy difference between the substructures has an increasing trend with overlap area in all cases. It can be concluded that the energy field

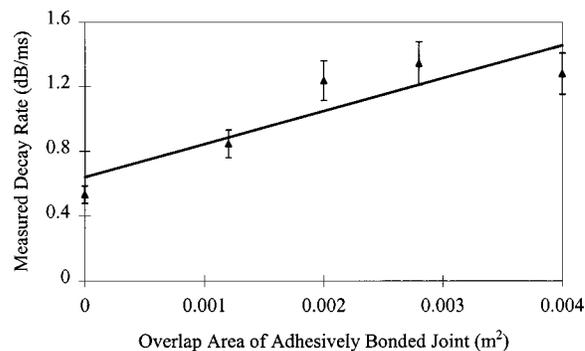


FIG. 14. Decay rate versus joint overlap area for adhesively bonded plates; error bars indicate one standard deviation.

within the bolted structure as a whole has not become truly diffuse and the two substructures are weakly coupled.

C. Plates with adhesively bonded joints

Adhesive joints are commonly used in aerospace structures, particularly the fuselage and wings of modern military and civil aircraft. It is therefore important to understand the effect of adhesive joints on the generation of diffuse fields. An experimental structure was made consisting of two plates of 5-mm-thick aluminum plates with the same size and shape as the plates used for the bolted joint experiments. The joint overlap area was varied in an identical fashion as for the bolted joints and the same parameters were measured. The bonds were made using a hot curing film adhesive (REDUX 322-300GSM) which is commonly used in the aerospace industry for bonding panels and stiffeners. The adhesive was cured for one hour at 180 °C and the resulting bonds were measured to be 0.1 mm thick. The results are shown in Table IV. The measured decay rate is plotted against overlap area in Fig. 14; the error bars indicate one standard deviation and the solid line represents a linear least squares fit. The decay rate exhibits a broadly increasing trend with overlap area, but this increase is much less than that previously measured for bolted joints. Unlike bolted joints, the mating surfaces of an adhesive joint are rigidly bonded, so if there is no relative motion between the surface, there can be no frictional losses. The damping increase is believed to be caused by the energy loss in the adhesive.

The standard deviation of the area measurements taken on each individual structure shows a generally increasing trend with joint overlap area. This indicates that the field generated in the substructures is becoming less diffuse as the damping increases, as was previously found for the bolted joints. The attenuation per wave transit across a bonded substructure as a function of overlap area is also shown in Table IV. The results show that the damping is considerably less than that seen for bolted structures of the same geometry, the highest attenuation being 0.013 Np/transit. This value is lower than the cutoff value of around 0.025 Np/transit which was found to be the highest allowable attenuation for a diffuse field to be generated.

The difference between the average envelope area measurements taken on either side of the joint is shown in Table

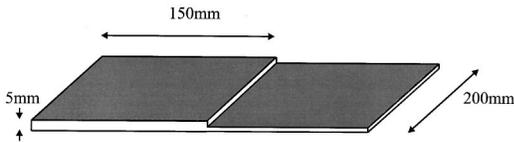


FIG. 15. Geometry of the stepped plates.

IV; as in the bolted joint case, a delay of 3 ms was used for the envelope area measurements. The difference between these energies increases with overlap area as was seen for the bolted joints; however, this increase is not proportional to overlap area. The energy difference increases slowly from a minimum of 16% (at an overlap of 0.0012 m^2) to 19% (at an overlap of 0.0012 m^2), then jumps suddenly to 59%. The reason for this large jump is not entirely understood, although it is thought to be due to inconsistent joint properties. Again this demonstrates that the substructures are weakly coupled and it can be concluded that the energy field within the adhesively structure as a whole has not become truly diffuse.

D. Plates with step changes of thickness

The experiments carried out in the previous sections have concentrated on plates with a uniform thickness of 5 mm. Many practical structures contain thickness changes due to joints, stiffening members, or tapering members, so it is important to understand how these thickness changes are likely to affect the diffuse field properties of the structure. Changes of thickness are expected to affect the partitioning of energy between areas of a structure, which may have important implications for diffuse field acoustic emission measurements.

The characteristics of each mode, such as mode shape and group velocity, vary with the frequency-thickness product, as shown in the dispersion curves.^{29,30} Thus, a change of thickness will shift the operating point to another part of the dispersion curve, assuming constant frequency. In order to study this effect, two plates were machined with step changes in thickness. Both plates had plan dimensions of $0.2 \times 0.3 \text{ m}^2$ and were machined from 5-mm-thick aluminum as shown in Fig. 15. The thickness was reduced over half the area of the plates to 2.7 mm (plate A) and 1.3 mm (plate B) while the remaining half of each plate remained 5 mm thick. A five-cycle tone burst at 150 kHz was used as the source function, the excitation transducer being on the 5-mm-thick section of the plate. Ten measurements were taken on each side of the step, the source being undisturbed between measurements in order to ensure consistent coupling. The envelope area measurements were averaged and the standard deviations over measurements taken on each substructure were

TABLE V. Measured envelope area squared on sections of the stepped plates.

	(EA) ² in 5 mm section (mVs) ²	(EA) ² in reduced section (mVs) ²	Corrected (EA) ² in reduced section (mVs) ²
Plate A	56.0	144.1	58.9
Plate B	42.3	278.6	44.9

all found to be less than 8%. Table V shows the square of the measured envelope area for each substructure; this is assumed to be indicative of only the A_0 field due to the poor sensitivity of the receiver to S_0/SH_0 . The square of the envelope areas cannot be compared directly between the substructures of different thicknesses as the thickness change strongly affects the relationship between surface displacement and energy. If the mode is nondispersive and the mode shape does not vary with the frequency-thickness product, the energy in the plate for a given surface displacement at a particular frequency is simply proportional to the plate thickness. If the mode is dispersive, the frequency-thickness change causes a change of mode shape which further alters the ratio of surface displacement to energy. Applying energy correction factors derived by the procedure described in Ref. 28 to the raw values yields the corrected figures shown in the fourth column of Table V. These figures can now be compared directly with the values in the 5-mm-thick sections.

The results clearly show that the surface amplitude in the thin sections is much larger than in the thicker sections, as might be expected. However, when the energy correction factor is applied, the difference is much smaller. Since the source amplitude was the same on both plates, it might be expected that the sum of the squares of the (corrected) envelope areas in the original thickness and reduced sections would be the same on each plate. However, this sum is 114.9 (mVs)^2 on plate A and only 87.2 (mVs)^2 on plate B. This difference is probably due to the decay rate on plate B being larger than on plate A since air damping will be more significant on the thin (1.3 mm thick) section of plate B. There may also be differences in the damping change produced by the machining process. Since the envelope area measurement is an integral over the time history of the decay, it cannot be used to compare the source energy in two structures if they have different decay rates.

The corrected energies in the two halves of each plate are very similar, but this may be due to coincidence. Simple predictions neglecting the presence of the SH_0 mode²⁸ suggest that if the damping in the two halves of the plate is the same, the fraction of the overall energy in the thinner part of the plate will increase as the thickness of this section is reduced. This effect is due to the group velocity of the A_0 mode at 150 kHz reducing as the plate thickness decreases. This reduces the reflection frequency (number of edge reflections per second) in the thinner half of the plate and so decreases the opportunity for mode conversion in this section. It is probable that in the experiments, this effect was cancelled out by the increase in air damping in the thinner section.

V. CONCLUSIONS

Measurements and predictions of the response of a variety of plain and coupled plates when excited by a point source have been reported. Tests on plain plates showed that a diffuse field was generated within the first 30 wave transits across the plate. In plates coupled by a thin ligament, around 50 wave transits across one of the substructures was required for a diffuse field to be established in the whole system. This is equivalent to about 25 transits across the whole plate so

the time required to set up a diffuse field was not significantly affected by reducing the coupling between parts of the structure in this way, even when the ligament width was only 2.5% of the total plate width. Tests on plates bolted together showed that both the coupling between the plates and the damping increased with the overlap area. While a diffuse field was established in each of the substructures, the damping was too great for a diffuse field to be set up in the whole structure. The damping in adhesively bonded plates was less than that in the bolted systems, but was still too high for the establishment of a diffuse field. Tests on stepped plates have shown that the measured surface amplitudes in the thinner sections tend to be much larger than those in the thick sections, but that the energy in the thinner sections is only slightly larger than that in a thick section of similar plan area. The field is not diffuse in the sense that the amplitude or energy density is the same throughout the coupled structure, but the field is diffuse in each substructure.

ACKNOWLEDGMENTS

This work was funded by an EPSRC CASE studentship for Dr. Evans in conjunction with Rolls Royce PLC. The authors are grateful for helpful discussions with Dr. John Webster of Rolls Royce and Dr. Brian Pavlakovic of Imperial College.

- ¹N. O. Cross, *Acoustic Emission Testing of Pressure Vessels for Petroleum Refineries and Chemical Plants* (ASTM, Philadelphia, 1982).
- ²P. Tscheliesnig and H. Theiretzbacher, "Leakage test by acoustic emission testing (AET) on flat bottomed tanks," in Proceedings Second International Conference on Acoustic Emission, Lake Tahoe (1985), pp. 24–27.
- ³S. E. Dunne, *Discussion of the application of acoustic emission methods for field monitoring of corrosion induced structural deterioration of reinforced and pre-stressed concrete structures* (Florida Atlantic University, Boca Raton, FL, 1985).
- ⁴S. Lovass, "Acoustic emission of offshore structures: attenuation; noise; crack monitoring," *J. Acoust. Emiss.* **26**(8), A161–A164 (1985).
- ⁵H. Kuttruff, *Room Acoustics* (Applied Science, London, 1973).
- ⁶D. M. Egle, "A stochastic model for transient acoustic emission signals," *J. Acoust. Soc. Am.* **65**, 1198–1203 (1979).
- ⁷D. M. Egle, "Diffuse wave fields in solid media," *J. Acoust. Soc. Am.* **70**, 476–480 (1981).
- ⁸R. L. Weaver, "On diffuse waves in solid media," *J. Acoust. Soc. Am.* **71**, 1608 (1982).

- ⁹R. L. Weaver, "Diffuse waves in finite plates," *J. Acoust. Soc. Am.* **94**, 319–335 (1984).
- ¹⁰R. L. Weaver, "Diffuse elastic waves at a free surface," *J. Acoust. Soc. Am.* **78**, 131–136 (1985).
- ¹¹R. L. Weaver, "Laboratory studies of diffuse waves in plates," *J. Acoust. Soc. Am.* **79**, 919–923 (1986).
- ¹²R. L. Weaver, "Diffuse waves for materials NDE," in *Acousto-Ultrasonics* (Plenum, New York, 1988).
- ¹³R. H. Lyon, *Statistical Energy Analysis of Dynamical Systems: Theory and Applications* (Harvard U. P., Cambridge, MA, 1976).
- ¹⁴R. B. Clough, "The energetics of acoustic emission source characterisation," *Mater. Eval.* **45**, 556–563 (1987).
- ¹⁵R. B. Clough, "A scalar approach to acoustic emission," in Proc. International Symposium on Vibroacoustic Characterisation of Materials and Structures (ASME, New York, 1992), pp. 1–6.
- ¹⁶R. H. Lyon and E. Eichler, "Random vibration of connected structures," *J. Acoust. Soc. Am.* **36**, 1344–1354 (1963).
- ¹⁷J. Woodhouse, "An approach to the theoretical background of statistical energy analysis applied to structural vibration," *J. Acoust. Soc. Am.* **69**, 1695–1709 (1981).
- ¹⁸J. Woodhouse, "An introduction to statistical energy analysis of structural vibration," *Appl. Acoust.* **14**, 455–469 (1981).
- ¹⁹M. L. Lai and A. Soom, "Prediction of transient vibration envelopes using statistical energy analysis techniques," *Trans. ASME, J. Vib. Acoust.* **112**, 127–137 (1990).
- ²⁰R. J. Pinnington and D. Lednik, "Transient statistical energy analysis of an impulsively excited two oscillator system," *J. Sound Vib.* **189**(2), 249–264 (1996).
- ²¹R. E. Powell and L. R. Quartararo, "Statistical energy analysis of transient vibration," in Proc. ASME Winter Meeting, Boston, 1987, edited by K. H. Hsu, D. J. Nefske, and A. Akay, Vol. 3, pp. 3–8.
- ²²T. M. Proctor, "An improved piezoelectric acoustic emission transducer," *J. Acoust. Soc. Am.* **71**, 1163–1168 (1982).
- ²³T. J. Holroyd, J. R. Webster, P. E. Cox, and R. B. Price, "Uses of AE in Rolls-Royce plc," in Proc. 8th International Acoustic Emission Symposium, Tokyo (1986), pp. 1–7.
- ²⁴R. V. Williams, *Acoustic Emission* (Hilger, Bristol, 1980).
- ²⁵H. A. L. Dempsey and D. M. Egle, "The effects of transducers on the decay of diffuse energy fields," *J. Acoust. Emiss.* **4**(1), S46–S49 (1985).
- ²⁶M. G. Silk, *Ultrasonic Transducers for Non-destructive Testing* (Hilger, Bristol, 1984).
- ²⁷K. F. Bainton and M. Silk, "Some Factors Which Affect The Performance of Ultrasonic Transducers," *Br. J. Non-Destr. Test.* **22**, 15–20 (1980).
- ²⁸M. J. Evans, "The use of diffuse field measurements for acoustic emission," Ph.D. thesis, Imperial College, University of London, 1997.
- ²⁹M. J. S. Lowe, "Matrix techniques for modelling ultrasonic waves in multilayered media," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **42**(4), 525–542 (1995).
- ³⁰B. Pavlakovic, M. Lowe, D. Alleyne, and P. Cawley, "Disperse: a general purpose program for creating dispersion curves," in *Proc. Review of Progress in Quantitative NDE*, edited by D. Thomson and D. Chimenti (Plenum, New York, 1997), Vol. 16A, pp. 185–192.

Reduced models for structures in the medium-frequency range coupled with internal acoustic cavities

Christian Soize

Structural Dynamics and Coupled Systems Department, ONERA, BP 72, F-92322 Chatillon Cedex, France

(Received 16 February 1999; revised 3 September 1999; accepted 4 September 1999)

This paper concerns the development of a method adapted for constructing reduced models in the medium-frequency range to a general three-dimensional dissipative structure consisting of an anisotropic, inhomogeneous, viscoelastic bounded medium coupled with an internal acoustic cavity. The reduced models are obtained using the Ritz–Galerkin method for which the projection subspace corresponds to the dominant eigensubspace of the energy operator of the structure in the medium-frequency band of analysis. Two fundamental cases are considered: (1) both the structure and the internal acoustic cavity have a medium-frequency behavior in the frequency band of analysis; (2) the structure has a medium-frequency behavior in the frequency band of analysis while the internal acoustic cavity has a low-frequency behavior. © 1999 Acoustical Society of America. [S0001-4966(99)05912-3]

PACS numbers: 43.40.At, 43.40.Dx, 43.40.Rj, 43.20.Tb [CBB]

INTRODUCTION

This paper is the continuation of two previous papers published by the author, concerning the development of a method for constructing reduced models in the medium-frequency (MF) range for general structural dynamics systems¹ (structures in a vacuum) and external structural-acoustic systems² [structure coupled with an external acoustic fluid (gas or liquid)]. For a mechanical system such as a structure in a vacuum or a structural-acoustic system, its response in the frequency domain (its “behavior”) depends on the frequency range defined as follows (see Ref. 3).

The low-frequency range (LF) can be defined as the modal domain for which the associated conservative system has a small number of modes; there are no modal overlaps due to the dissipation effects. In this LF range, the finite element method can be used for spatial discretization and the dynamic analysis can be performed in the frequency domain using reduced models which are very efficient and popular tools in constructing the solution. Such reduced models correspond to a Ritz–Galerkin reduction of the dynamical model using the normal modes corresponding to the lowest eigenfrequencies of the associated conservative system. The efficiency of this kind of reduced model is due to the small number of generalized dynamical degrees of freedom used in the representation and, in addition, is obtained by solving a well-stated generalized symmetric eigenvalue problem for which only the first eigenvalues and the corresponding eigenfunctions have to be calculated. In addition when such a reduced model is obtained, responses to any deterministic or random excitations can be calculated for no significant additional numerical cost.

The high-frequency range (HF) can be defined as the frequency band for which there is a uniform high modal density; there is a uniform modal overlap due to the high modal density and the dissipation effects. For this HF range, the finite element method cannot be used for spatial discretization and the dynamic analysis has to be performed using the wave approach, the global statistical energy approach

(such as the very popular Statistical Energy Analysis) and the local energy approach (such as the power flow analysis based on continuous energy equations).

For complex systems such as general three-dimensional structures, an intermediate frequency range called medium-frequency range (MF) appears. The modal density exhibits large variations over the band. In addition, if there is a structural complexity related to the presence of fuzzy substructures, mechanical models have to be adapted. This MF range cannot be analyzed with the tools used for the LF and HF ranges. A complex structure is then considered as a master structure coupled with all the dynamical subsystems (fuzzy substructures). In the MF range, probabilistic models have to be used to model the effects of fuzzy substructures on the master structure and the finite element method can only be used for spatial discretization of the master structure. In addition, random uncertainties have to be modeled in the master structure in order to increase robustness of response predictions. Concerning dynamical analysis of the master structure, the modal method which is very efficient in the LF range for constructing reduced models cannot be used in the MF range for general three-dimensional structures. It should be noted that it is essential to have intrinsic reduced models in the MF range (as we have in the LF range) for calculating the MF responses to any multiple loads constituted of deterministic and random excitations. This is the reason why a reduced model method in the MF range has been recently proposed for general dissipative structural-dynamics systems.^{1,2} In this method, the reduced model is constructed using the Ritz–Galerkin projection of the variational formulation of the boundary value problem on the dominant eigensubspace of the energy operator of the structure over the medium-frequency band of analysis. Similarly to the LF reduced models, the efficiency of the proposed MF reduced model is due to the small number of generalized dynamical degrees of freedom used in the representation and in addition, is obtained by solving a well-stated generalized symmetric eigenvalue problem for which only the first eigenval-

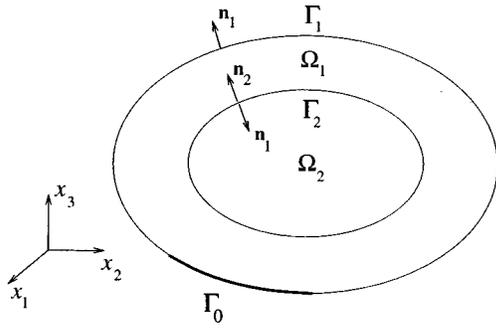


FIG. 1. Geometrical configuration of the structural-acoustic system.

ues and the corresponding eigenfunctions of the energy operator related to the MF band have to be calculated. This means that the first normal modes of the LF range are replaced by the first eigenfunctions of the energy operator in the MF range. Finally, it should be noted that the efficiency of such a reduced model approach in the MF range can be very high with respect to any direct approaches validated in the MF range as soon as the MF responses have to be calculated for a large number of multiple deterministic and random loads, particularly for random excitations.

In this paper, we apply and adapt this method for constructing a reduced model in the MF range to a general three-dimensional dissipative structure consisting of an anisotropic, inhomogeneous, viscoelastic bounded medium coupled with an internal acoustic cavity filled with a gas or a liquid. Two fundamental cases are considered: (1) both the structure and the internal acoustic cavity have a medium-frequency “behavior” in the frequency band of analysis; (2) the structure has a medium-frequency “behavior” in the frequency band of analysis while the internal acoustic cavity has a low-frequency “behavior.”

In Sec. I, we present the boundary value problem to be solved in the frequency domain and we recall its variational formulation. In Sec. II, we introduce a finite dimension approximation. Section III concerns the construction of a vector basis for the reduced model of the structure and the internal acoustic cavity in the MF range. We will see the role played by the low-frequency structural modes on the acoustic cavity response in the MF range. As a consequence, we deduce a structural vector basis adapted to the MF range when the structure is coupled with an internal acoustic cavity. Section IV is devoted to the construction of the reduced model while Sec. V deals with the construction of the dominant eigensubspaces. In Sec. VI, we present the time-stationary random response using the reduced model. Finally, we present a validation for the two fundamental cases introduced above.

I. BOUNDARY VALUE PROBLEM AND ITS VARIATIONAL FORMULATION

A. Definition of the boundary value problem

We consider linear vibrations (formulated in the frequency domain ω) of a three-dimensional structural-acoustic system around a static equilibrium configuration considered as a natural state at rest (see Fig. 1). Let Ω_1 be the three-

dimensional bounded domain occupied by the structure and made of viscoelastic material. Let $\partial\Omega_1 = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2$ be its boundary and $\mathbf{n}_1 = (n_{1,1}, n_{1,2}, n_{1,3})$ be its outward unit normal. Let $\mathbf{u}(\mathbf{x}, \omega) = (u_1(\mathbf{x}, \omega), u_2(\mathbf{x}, \omega), u_3(\mathbf{x}, \omega))$ be the displacement field in each point $\mathbf{x} = (x_1, x_2, x_3)$ in Cartesian coordinates and at frequency ω . On part Γ_0 of the boundary, the structure is fixed ($\mathbf{u} = \mathbf{0}$) whereas on part $\Gamma_1 \cup \Gamma_2$ it is free. The structure is coupled with an internal dissipative acoustic fluid (gas or liquid) occupying the three-dimensional bounded domain Ω_2 whose boundary $\partial\Omega_2$ is the coupling interface Γ_2 . The outward unit normal to $\partial\Omega_2$ is denoted as $\mathbf{n}_2 = (n_{2,1}, n_{2,2}, n_{2,3})$ and we have $\mathbf{n}_2 = -\mathbf{n}_1$. We denote the pressure field in Ω_2 as $p(\mathbf{x}, \omega)$. We introduce a narrow MF band B such that

$$B = [\omega_B - \Delta\omega/2, \omega_B + \Delta\omega/2], \quad (1)$$

in which ω_B is the center frequency and $\Delta\omega$ is the bandwidth such that $\Delta\omega/\omega_B \ll 1$ and $\omega_B > \Delta\omega/2$. With B we associate interval

$$\tilde{B} = [-\omega_B - \Delta\omega/2, -\omega_B + \Delta\omega/2]. \quad (2)$$

The structure is submitted to a square integrable surface force field $\mathbf{x} \rightarrow \eta(\omega)\mathbf{g}(\mathbf{x}, \omega)$ from Γ_1 into \mathbb{C}^3 , in which $\eta(\omega)$ is a function from \mathbb{R} into \mathbb{C} , such that $\eta(\omega) = 0$ if ω is not in $B \cup \tilde{B}$, continuous on B , verifying $|\eta(-\omega)| = |\eta(\omega)|$ and such that $|\eta(\omega)| \neq 0$ for all ω in B .

For this structural-acoustic system, we use the model and the boundary value problem developed for the MF range in Chapter XIV of Ref. 3. Introducing components (g_1, g_2, g_3) of \mathbf{g} , the boundary value problem for the structure is written as follows in terms of \mathbf{u} (the convention for the Fourier transform being $v(\omega) = \int_{\mathbb{R}} e^{-i\omega t} v(t) dt$),

$$-\omega^2 \rho_1 u_i - \sigma_{ij,j} = 0 \quad \text{in } \Omega_1, \quad (3)$$

$$\sigma_{ij} n_{1,j} = \eta g_i \quad \text{on } \Gamma_1, \quad (4)$$

$$\sigma_{ij} n_{1,j} = -p n_{1,j} \quad \text{on } \Gamma_2, \quad (5)$$

$$u_i = 0 \quad \text{on } \Gamma_0, \quad (6)$$

in which $i = 1, 2, 3$, where the summation over index j is used, and where $\rho_1(\mathbf{x}) > 0$ is the mass density of the structure and $\sigma_{ij,j} = \sum_{j=1}^3 \partial \sigma_{ij} / \partial x_j$. For a linear viscoelastic material, stress tensor σ_{ij} is written as

$$\sigma_{ij} = a_{ijkh}(\mathbf{x}, \omega) \epsilon_{kh}(\mathbf{u}) + b_{ijkh}(\mathbf{x}, \omega) \epsilon_{kh}(i\omega \mathbf{u}), \quad (7)$$

in which the summation over indices k and h is used and where $\epsilon_{kh}(\mathbf{u}) = (\partial u_k / \partial x_h + \partial u_h / \partial x_k) / 2$ is the linearized strain tensor. Coefficients $a_{ijkh}(\mathbf{x}, \omega)$ and $b_{ijkh}(\mathbf{x}, \omega)$ are real, depend on \mathbf{x} and ω , verify the usual properties of symmetry and positiveness³⁻⁶ and are such that $a_{ijkh}(\mathbf{x}, -\omega) = a_{ijkh}(\mathbf{x}, \omega)$ and $b_{ijkh}(\mathbf{x}, -\omega) = b_{ijkh}(\mathbf{x}, \omega)$.

Concerning the internal dissipative acoustic fluid, the pressure in the fluid is written (see Chapter XIV of Ref. 3) as

$$p(\mathbf{x}, \omega) = -i\omega \rho_2 \psi(\mathbf{x}, \omega) - \kappa \pi_2(\mathbf{u}) \quad \text{in } \Omega_2 \cup \Gamma_2, \quad (8)$$

in which κ is a positive constant such that

$$\kappa = \frac{\rho_2 c_2^2}{|\Omega_2|}, \quad |\Omega_2| = \int_{\Omega_2} d\mathbf{x}, \quad (9)$$

where $\rho_2 > 0$ and c_2 are the constant mass density and the constant speed of sound of the acoustic fluid at equilibrium and where $\pi_2(\mathbf{u})$ is defined by

$$\pi_2(\mathbf{u}) = \int_{\Gamma_2} \mathbf{u}(\mathbf{x}, \omega) \cdot \mathbf{n}_2(\mathbf{x}) ds(\mathbf{x}). \quad (10)$$

The new unknown field $\psi(\mathbf{x}, \omega)$ is related to the velocity field $\mathbf{v}(\mathbf{x}, \omega)$ of the dissipative acoustic fluid by the equation $\mathbf{v}(\mathbf{x}, \omega) = (1 + i\omega\tau)\nabla\psi(\mathbf{x}, \omega)$ in which τ is a constant coefficient³ related to the viscosity of the acoustic fluid (τ may depend on frequency ω). The boundary value problem for the internal acoustic fluid is written as follows in terms of field ψ

$$-\omega^2 \frac{\rho_2}{c_2^2} \psi(\mathbf{x}, \omega) - i\omega\tau\rho_2 \nabla^2 \psi(\mathbf{x}, \omega) - \rho_2 \nabla^2 \psi(\mathbf{x}, \omega) = -\frac{i\omega\kappa}{c_2^2} \pi_2(\mathbf{u}) \quad \text{in } \Omega_2, \quad (11)$$

$$\rho_2(1 + i\omega\tau) \frac{\partial \psi}{\partial \mathbf{n}_2} = i\omega\rho_2 \mathbf{u} \cdot \mathbf{n}_2 \quad \text{on } \Gamma_2, \quad (12)$$

$$\int_{\Omega_2} \psi(\mathbf{x}, \omega) d\mathbf{x} = 0. \quad (13)$$

The boundary value problem of the structural-acoustic problem is defined by Eqs. (3)–(13).

B. Variational formulation

Let V_1 be the space of admissible displacement fields \mathbf{u} defined on Ω_1 with values in \mathbb{C}^3 such that $\mathbf{u} = \mathbf{0}$ on Γ_0 . Let V_2 be the space of admissible fields ψ defined on Ω_2 with values in \mathbb{C} such that $\int_{\Omega_2} \psi(\mathbf{x}) d\mathbf{x} = 0$. Below, \bar{z} denotes the conjugate of the complex number z . The variational formulation of the boundary value problem in \mathbf{u} and ψ defined by Eqs. (3) to (13) is obtained using the test-function method and is expressed as follows (see Chapters II and XIV of Ref. 3). For all ω in $B \cup \tilde{B}$, find $\{\mathbf{u}(\omega), \psi(\omega)\}$ in $V_1 \times V_2$ such that, for all $\{\mathbf{v}, \phi\}$ in $V_1 \times V_2$,

$$a_1(\mathbf{u}, \mathbf{v}; \omega) + \kappa j(\mathbf{u}, \mathbf{v}) + i\omega c(\psi, \mathbf{v}) = f(\mathbf{v}; \omega), \quad (14)$$

$$i\omega c(\phi, \mathbf{u}) - a_2(\psi, \phi; \omega) = 0, \quad (15)$$

in which $f(\mathbf{v}; \omega)$ is defined by

$$f(\mathbf{v}; \omega) = \eta(\omega) \int_{\Gamma_1} \mathbf{g}(\mathbf{x}, \omega) \cdot \overline{\mathbf{v}(\mathbf{x})} ds(\mathbf{x}), \quad (16)$$

and where $a_1(\mathbf{u}, \mathbf{v}; \omega)$ is written as

$$a_1(\mathbf{u}, \mathbf{v}; \omega) = -\omega^2 m_1(\mathbf{u}, \mathbf{v}) + i\omega d_1(\mathbf{u}, \mathbf{v}; \omega) + k_1(\mathbf{u}, \mathbf{v}; \omega), \quad (17)$$

$$m_1(\mathbf{u}, \mathbf{v}) = \int_{\Omega_1} \rho_1(\mathbf{x}) \mathbf{u}(\mathbf{x}, \omega) \cdot \overline{\mathbf{v}(\mathbf{x})} d\mathbf{x}, \quad (18)$$

$$d_1(\mathbf{u}, \mathbf{v}; \omega) = \int_{\Omega_1} b_{ijkh}(\mathbf{x}, \omega) \varepsilon_{kh}(\mathbf{u}) \varepsilon_{ij}(\overline{\mathbf{v}}) d\mathbf{x}, \quad (19)$$

$$k_1(\mathbf{u}, \mathbf{v}; \omega) = \int_{\Omega_1} a_{ijkh}(\mathbf{x}, \omega) \varepsilon_{kh}(\mathbf{u}) \varepsilon_{ij}(\overline{\mathbf{v}}) d\mathbf{x}, \quad (20)$$

in which the summation over indices i, j, k and h is used. It is assumed that $d_1(\mathbf{u}, \mathbf{v}; \omega)$ and $k_1(\mathbf{u}, \mathbf{v}; \omega)$ are continuous functions on band B with respect to ω . In Eqs. (14) and (15), we have

$$c(\psi, \mathbf{v}) = \rho_2 \int_{\Gamma_2} \psi(\mathbf{x}, \omega) \mathbf{n}_2(\mathbf{x}) \cdot \overline{\mathbf{v}(\mathbf{x})} ds(\mathbf{x}), \quad (21)$$

$$j(\mathbf{u}, \mathbf{v}) = \pi_2(\mathbf{u}) \pi_2(\overline{\mathbf{v}}), \quad (22)$$

$$a_2(\psi, \phi; \omega) = -\omega^2 m_2(\psi, \phi) + i\omega d_2(\psi, \phi; \omega) + k_2(\psi, \phi), \quad (23)$$

$$m_2(\psi, \phi) = \frac{\rho_2}{c_2^2} \int_{\Omega_2} \omega(\mathbf{x}, \omega) \overline{\phi(\mathbf{x})} d\mathbf{x}, \quad (24)$$

$$d_2(\psi, \phi; \omega) = \tau(\omega) k_2(\psi, \phi), \quad (25)$$

$$k_2(\psi, \phi) = \rho_2 \int_{\Omega_2} \nabla \psi(\mathbf{x}, \omega) \cdot \nabla \overline{\phi(\mathbf{x})} d\mathbf{x}. \quad (26)$$

For all ω in $B \cup \tilde{B}$, the problem defined by Eqs. (14) and (15) has a unique solution $\{\mathbf{u}(\omega), \psi(\omega)\}$ in $V_1 \times V_2$.

II. FINITE DIMENSION APPROXIMATION

The finite dimension approximation of the problem defined by Eqs. (14) and (15) is obtained by using the Ritz–Galerkin method. We then consider a complete family of independent \mathbb{R}^3 -valued functions $\{\mathbf{u}_\alpha\}_{\alpha \geq 1}$ in admissible space V_1 and a complete family of independent real-valued functions $\{\psi_\beta\}_{\beta \geq 1}$ in admissible space V_2 . We consider (1) a subspace $V_{1,n_1} \subset V_1$ of finite dimension $n_1 \geq 1$ spanned by the family $\{\mathbf{u}_1, \dots, \mathbf{u}_{n_1}\}$ and (2) a subspace $V_{2,n_2} \subset V_2$ of finite dimension $n_2 \geq 1$ spanned by the family $\{\psi_1, \dots, \psi_{n_2}\}$. In practice, each family can be either a finite element basis^{7,8} associated with a finite element mesh of domain Ω_1 for $\{\mathbf{u}_\alpha\}_\alpha$ and domain Ω_2 for $\{\psi_\beta\}_\beta$, or, for particular cases corresponding to simple shapes of geometry, any functional basis constructed in the context of an analytical approach such as a sequence of structural normal modes^{9,10} of the structure in vacuo and a sequence of acoustic modes^{9,11} of the internal acoustic cavity with a rigid wall. Let $\{\mathbf{u}(\omega), \psi(\omega)\} \in V_1 \times V_2$ be the solution of Eqs. (14) and (15). Its projection on $V_{1,n_1} \times V_{2,n_2}$ is written as

$$\tilde{\mathbf{u}}(\mathbf{x}, \omega) = \sum_{\alpha=1}^{n_1} q_{1,\alpha}(\omega) \mathbf{u}_\alpha(\mathbf{x}). \quad (27)$$

$$\tilde{\psi}(\mathbf{x}, \omega) = \sum_{\beta=1}^{n_2} q_{2,\beta}(\omega) \psi_\beta(\mathbf{x}). \quad (28)$$

From Eqs. (14) and (15), we deduce that $\mathbf{q}_1(\omega) = (q_{1,1}(\omega), \dots, q_{1,n_1}(\omega)) \in \mathbb{C}^{n_1}$ and $\mathbf{q}_2(\omega) = (q_{2,1}(\omega), \dots, q_{2,n_2}(\omega)) \in \mathbb{C}^{n_2}$ is the unique solution of the linear equation³

$$\begin{bmatrix} [A_1(\omega)] + \kappa[J] & i\omega[C] \\ i\omega[C]^T & -[A_2(\omega)] \end{bmatrix} \begin{bmatrix} \mathbf{q}_1(\omega) \\ \mathbf{q}_2(\omega) \end{bmatrix} = \begin{bmatrix} \eta(\omega)\mathbf{F}(\omega) \\ 0 \end{bmatrix}. \quad (29)$$

In Eq. (29), $[A_1(\omega)]$ is an $(n_1 \times n_1)$ symmetric complex matrix (invertible for all ω in $B \cup \bar{B}$) which is written as

$$[A_1(\omega)] = -\omega^2[M_1] + i\omega[D_1(\omega)] + [K_1(\omega)], \quad (30)$$

in which $[M_1]$, $[D_1(\omega)]$ and $[K_1(\omega)]$ are positive-definite symmetric $(n_1 \times n_1)$ real matrices such that $[M_1]_{\alpha\alpha'} = m_1(\mathbf{u}_{\alpha'}, \mathbf{u}_\alpha)$, $[D_1(\omega)]_{\alpha\alpha'} = d_1(\mathbf{u}_{\alpha'}, \mathbf{u}_\alpha; \omega)$ and $[K_1(\omega)]_{\alpha\alpha'} = k_1(\mathbf{u}_{\alpha'}, \mathbf{u}_\alpha; \omega)$. Vector $\mathbf{F}(\omega) = (F_1(\omega), \dots, F_{n_1}(\omega)) \in \mathbb{C}^{n_1}$ is such that

$$\eta(\omega)F_\alpha(\omega) = f(\mathbf{u}_\alpha; \omega). \quad (31)$$

Matrix $[J]$ is an $(n_1 \times n_1)$ symmetric real matrix which can be written as

$$[J] = \mathbf{\Pi}_2 \mathbf{\Pi}_2^T,$$

in which $\mathbf{\Pi}_2 = (\Pi_{2,1}, \dots, \Pi_{2,n_1}) \in \mathbb{R}^{n_1}$ is such that

$$\Pi_{2,\alpha} = \pi_2(\mathbf{u}_\alpha). \quad (32)$$

In Eq. (29), $[A_2(\omega)]$ is an $(n_2 \times n_2)$ symmetric complex matrix (invertible for all real ω in $B \cup \bar{B}$) which is written as

$$[A_2(\omega)] = -\omega^2[M_2] + i\omega[D_2(\omega)] + [K_2], \quad (33)$$

in which $[M_2]$, $[D_2(\omega)]$ and $[K_2]$ are positive-definite symmetric $(n_2 \times n_2)$ real matrices such that $[M_2]_{\beta\beta'} = m_2(\psi_{\beta'}, \psi_\beta)$, $[D_2(\omega)] = \tau(\omega)[K_2]$ and $[K_2]_{\beta\beta'} = k_2(\psi_{\beta'}, \psi_\beta)$. It should be noted that $[K_2]$ is positive definite due to the constraint $\int_{\Omega_2} \psi d\mathbf{x} = 0$ which is included in space V_2 . In the context of use of the finite element method, if this constraint is not included in the construction of $\{\psi_\beta\}_\beta$, then $[K_2]$ is only positive and the constraint $\mathbf{L}^T \mathbf{q}_2 = 0$ corresponding to the finite element discretization of $\int_{\Omega_2} \psi d\mathbf{x} = 0$ has to be added to Eq. (29). Finally, in Eq. (29), $[C]$ is an $(n_1 \times n_2)$ real matrix such that

$$[C]_{\alpha\beta} = c(\psi_\beta, \mathbf{u}_\alpha). \quad (34)$$

The approximation of pressure field p defined by Eq. (8) is written as $\tilde{p}(\mathbf{x}, \omega) = -i\omega\rho_2\tilde{\psi}(\mathbf{x}, \omega) - \kappa\pi_2(\tilde{\mathbf{u}})$, i.e.,

$$\tilde{p}(\mathbf{x}, \omega) = -i\omega\rho_2\mathbf{\Psi}(\mathbf{x})^T \mathbf{q}_2(\omega) - \kappa\mathbf{\Pi}_2^T \mathbf{q}_1(\omega), \quad (35)$$

in which $\mathbf{\Psi}(\mathbf{x}) = (\psi_1(\mathbf{x}), \dots, \psi_{n_2}(\mathbf{x})) \in \mathbb{R}^{n_2}$.

III. CONSTRUCTION OF A VECTOR BASIS FOR THE REDUCED MODEL

A. Vector bases adapted to MF band B for the structure *in vacuo* and the internal acoustic cavity with rigid wall

Two vector bases adapted to MF band B can be constructed for the structure *in vacuo* and the internal acoustic cavity with rigid wall by applying the method presented in Refs. 1 and 2 (for the details, we refer the reader to these references). These two vector bases correspond to the dominant eigensubspaces of the energy operators relative to band B for the structure *in vacuo* and the internal acoustic fluid with rigid wall. In the context of the finite dimension approximation introduced in Sec. II, the procedure can be summarized as follows. Let $a = 1$ or $a = 2$ be the index related to the structure or the internal acoustic fluid. Let $N_a \ll n_a$ be the

order of the reduced model related to the structure ($a = 1$) or the internal acoustic fluid ($a = 2$). Let $[P_a]$ be the $(n_a \times N_a)$ real matrix whose columns are the N_a eigenvectors $\{\mathbf{P}_a^1, \dots, \mathbf{P}_a^{N_a}\}$ corresponding to the N_a highest eigenvalues $\lambda_a^1 \geq \dots \geq \lambda_a^{N_a}$ of the generalized symmetric eigenvalue problem

$$[H_a][P_a] = [G_a][P_a][\Lambda_a], \quad (36)$$

such that

$$[P_a]^T [G_a] [P_a] = [I_a], \quad (37)$$

$$[P_a]^T [H_a] [P_a] = [\Lambda_a], \quad (38)$$

in which $[I_a]$ is the $(N_a \times N_a)$ identity matrix, $[\Lambda_a]$ is the $(N_a \times N_a)$ diagonal matrix of eigenvalues $\lambda_a^1, \dots, \lambda_a^{N_a}$ and where $[G_a]$ and $[H_a]$ are positive-definite symmetric $(n_a \times n_a)$ real matrices such that

$$[G_1]_{\alpha\alpha'} = \int_{\Omega_1} \mathbf{u}_{\alpha'} \cdot \mathbf{u}_\alpha d\mathbf{x}, \quad [G_2]_{\beta\beta'} = \int_{\Omega_2} \psi_{\beta'} \cdot \psi_\beta d\mathbf{x}, \quad (39)$$

$$[H_a] = [G_a][E_a][G_a]. \quad (40)$$

In Eq. (40), positive-definite symmetric $(n_a \times n_a)$ real matrix $[E_a]$ is the projection of the energy operator of the structure ($a = 1$) or the internal acoustic fluid ($a = 2$), such that

$$[E_a] = \int_B [e_a(\omega)] d\omega, \quad (41)$$

$$[e_a(\omega)] = \frac{1}{\pi} \omega^2 |\eta(\omega)|^2 \text{Re}\{[T_a(\omega)]^* [M_a] [T_a(\omega)]\}, \quad (42)$$

$$[T_a(\omega)] = [A_a(\omega)]^{-1}; \quad (43)$$

$$[T_a(\omega)]^* = \overline{[T_a(\omega)]^T} = \overline{[T_a(\omega)]},$$

in which $[A_a(\omega)]$ is defined by Eq. (30) for $a = 1$ and Eq. (33) for $a = 2$.

B. Remark concerning the acoustic vector basis relative to MF band B

It can easily be proved that for the internal acoustic fluid ($a = 2$), eigenvectors $\{\mathbf{P}_2^1, \dots, \mathbf{P}_2^{N_2}\}$, associated with the N_2 highest eigenvalues of generalized symmetric eigenvalue problem $[H_2][P_2] = [G_2][P_2][\Lambda_2]$ relative to MF band B , coincide with the finite dimension approximation of the acoustic modes of the internal acoustic cavity with rigid wall, whose eigenfrequencies are in MF band B . This particular property is due to the fact that two conditions are simultaneously satisfied: mass density ρ_2 of the acoustic fluid is a constant and damping sesquilinear form $d_2(\psi, \phi; \omega) = \tau(\omega)k_2(\psi, \phi)$ is diagonalized by the acoustic modes. If either of these two conditions is not satisfied, then this particular property does not hold; generally, these conditions are not satisfied for structures in the MF range. It should be noted that, in the context of a finite element model of the internal acoustic fluid, if the rank of the acoustic modes belonging to MF band B is high (case of MF behavior of the internal acoustic fluid), then the proposed method is an

efficient tool for computation of these acoustic modes. Conversely, if the rank of the acoustic modes belonging to MF band B is low (case of LF behavior of the internal acoustic fluid), then the usual numerical methods for computation of the acoustic modes (such as subspace iteration or Lanczos methods^{12–15}) are more efficient.

C. Role played by the low-frequency structural modes on the acoustic cavity response in MF band B

In this section, in order to analyze this role, we introduce particular assumptions for simplifying the explanations. Let us assume that $\{\psi_\beta\}_\beta$ corresponds to the acoustic modes verifying

$$k_2(\psi_\beta, \phi) = \omega_{2,\beta}^2 m_2(\psi_\beta, \phi), \quad \forall \phi \in V_2,$$

with the normalization $m_2(\psi_\beta, \psi_\beta) = 1$ and such that $[D_2(\omega)]_{\beta\beta'} = 2\xi_{2,\beta}\omega_{2,\beta}\delta_{\beta\beta'}$ in which $\xi_{2,\beta} > 0$ is an acoustic damping rate depending on β but independent of ω . We choose $\{\mathbf{u}_\alpha\}_\alpha = \{\hat{\mathbf{u}}_\alpha\}_\alpha$ as the structural modes taking into account the additional stiffness term $\kappa j(\mathbf{u}, \mathbf{v})$. Consequently, the spectral problem is written as

$$\hat{k}_1(\hat{\mathbf{u}}_\alpha, \mathbf{v}) = \hat{\omega}_{1,\alpha}^2 m_1(\hat{\mathbf{u}}_\alpha, \mathbf{v}), \quad \forall \mathbf{v} \in V_1,$$

in which $\hat{k}_1(\mathbf{u}, \mathbf{v}) = k_1(\mathbf{u}, \mathbf{v}; 0) + \kappa j(\mathbf{u}, \mathbf{v})$. The normalization is such that $m_1(\hat{\mathbf{u}}_\alpha, \hat{\mathbf{u}}_\alpha) = 1$ and it is assumed that $[D_1(\omega)]_{\alpha\alpha'} = 2\hat{\xi}_{1,\alpha}\hat{\omega}_{1,\alpha}\delta_{\alpha\alpha'}$ where $\hat{\xi}_{1,\alpha} > 0$ is a structural damping rate depending on α but independent of ω . From Eqs. (29) to (34), we deduce that $\mathbf{q}_2(\omega)$ is the solution of the linear matrix equation

$$[A_2(\omega)]\mathbf{q}_2(\omega) = \mathbb{F}_2(\omega),$$

in which $\mathbb{F}_2(\omega)$ is a given vector in C^{n_2} and $[A_2(\omega)]$ is an $(n_2 \times n_2)$ symmetric complex matrix which can be written as $[A_2(\omega)] = [A_2^F(\omega)] + [A_2^S(\omega)]$ where, for all β and β' in $\{1, \dots, n_2\}$,

$$[A_2^F(\omega)]_{\beta\beta'} = (-\omega^2 + 2i\omega\omega_{2,\beta}\xi_{2,\beta} + \omega_{2,\beta}^2)\delta_{\beta\beta'},$$

$$[A_2^S(\omega)]_{\beta\beta'} = -\sum_{\alpha=1}^{n_1} \frac{\omega^2 C_{\alpha\beta} C_{\alpha\beta'}}{-\omega^2 + 2i\omega\hat{\omega}_{1,\alpha}\hat{\xi}_{1,\alpha} + \hat{\omega}_{1,\alpha}^2}.$$

Let us consider the MF response of the acoustic cavity due to an acoustic mode ψ_β whose eigenfrequency is $\omega_{2,\beta}$. Consequently, ω and $\omega_{2,\beta}$ belong to MF band B . Let us investigate the contribution of the first M low-frequency structural modes (whose eigenfrequencies $\hat{\omega}_{1,1}, \dots, \hat{\omega}_{1,M}$ belong to the low-frequency range) to the acoustic mode ψ_β whose eigenfrequency is $\omega_{2,\beta}$ belonging to MF band B . We then have

$$0 < \hat{\omega}_{1,1} \leq \dots \leq \hat{\omega}_{1,M} \ll \omega \in B,$$

and therefore, the contribution in $[A_2(\omega)]_{\beta\beta}$ of structural mode $\hat{\mathbf{u}}_\alpha$ whose eigenfrequency $\hat{\omega}_{1,\alpha}$ is such that $\hat{\omega}_{1,\alpha} \ll \omega \in B$, is equal to

$$\frac{-\omega^2 C_{\alpha\beta}^2}{-\omega^2 + 2i\omega\hat{\omega}_{1,\alpha}\hat{\xi}_{1,\alpha} + \hat{\omega}_{1,\alpha}^2} \simeq C_{\alpha\beta}^2.$$

This positive term $C_{\alpha\beta}^2$ contributes to increase the value of the acoustic eigenfrequency $\omega_{2,\beta}$ because generalized stiff-

ness term $\omega_{2,\beta}^2$ in $[A_2(\omega)]_{\beta\beta}$ is increasing of positive term $C_{\alpha\beta}^2$ for each structural mode $\hat{\mathbf{u}}_\alpha$ such that $\hat{\omega}_{1,\alpha} \ll \omega \in B$. Consequently, the rate of convergence of the internal acoustic MF response can be increased by “adding” the lowest structural modes $\{\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_M\}$ (or equivalently the lowest structural modes $\{\mathbf{u}_1, \dots, \mathbf{u}_M\}$ of the structure *in vacuo*) to vectors $\{\mathbf{P}_1^1, \dots, \mathbf{P}_1^{N_1}\}$ adapted to the prediction of the structural response in MF band B .

D. Structural vector basis adapted to MF band B for prediction of the internal acoustic response in MF band B

Taking into account the conclusion of Sec. III C, MF band B being fixed, we consider eigenvectors $\{\mathbf{P}_1^1, \dots, \mathbf{P}_1^{N_1}\}$ introduced in Sec. III A, such that [see Eq. (37)],

$$\mathbf{P}_1^{kT} [G_1] \mathbf{P}_1^j = \delta_{jk}, \quad j \text{ and } k \text{ in } \{1, \dots, N_1\},$$

and we consider the N_0 structural modes $\{\mathbf{Q}_1^1, \dots, \mathbf{Q}_1^{N_0}\}$ corresponding to the lowest structural eigenfrequencies $0 < \omega_{1,1} \leq \dots \leq \omega_{1,N_0}$ such that, for all α in $\{1, \dots, N_0\}$,

$$[K_1(0)]\mathbf{Q}_1^\alpha = \omega_{1,\alpha}^2 [M_1]\mathbf{Q}_1^\alpha.$$

We assume that

$$N_1^S = N_1 + N_0 < n_1,$$

and that vectors $\{\mathbf{P}_1^1, \dots, \mathbf{P}_1^{N_1}, \mathbf{Q}_1^1, \dots, \mathbf{Q}_1^{N_0}\}$ constitute a set of linearly independent vectors in \mathbb{R}^{n_1} . In practice, M structural modes $\{\mathbf{Q}_1^{\alpha_1}, \dots, \mathbf{Q}_1^{\alpha_M}\}$ are considered, the $(n_1 \times (N_1 + M))$ real matrix $[X] = [\mathbf{P}_1^1 \dots \mathbf{P}_1^{N_1} \mathbf{Q}_1^{\alpha_1} \dots \mathbf{Q}_1^{\alpha_M}]$ is constructed and finally, the $[Q][R]$ factorization of matrix $[X]$ is computed in which $[Q]$ is an $(n_1 \times n_1)$ orthogonal matrix and $[R]$ is an $(n_1 \times (N_1 + M))$ real upper triangular matrix. If none of the diagonal elements of $[R]$ is equal to zero, then $N_0 = M$; if there are m_0 diagonal elements equal to zero, then $N_0 = M - m_0$ and the corresponding structural modes $\mathbf{Q}_1^{\alpha_j}$ are eliminated. We then deduce the set $\{\mathbf{Q}_1^1, \dots, \mathbf{Q}_1^{N_0}\}$. Finally, a Gram–Schmidt algorithm with respect to the inner product defined by matrix $[G_1]$ is applied to the set $\{\mathbf{P}_1^1, \dots, \mathbf{P}_1^{N_1}, \mathbf{Q}_1^1, \dots, \mathbf{Q}_1^{N_0}\}$ for constructing the N_1^S linearly independent vectors $\{\mathbf{P}_1^1, \dots, \mathbf{P}_1^{N_1}, \mathbf{P}_1^{N_1+1}, \dots, \mathbf{P}_1^{N_1+N_0}\}$ in \mathbb{R}^{n_1} , such that for j and k in $\{1, \dots, N_1^S\}$, we have $\mathbf{P}_1^{kT} [G_1] \mathbf{P}_1^j = \delta_{jk}$. Vectors $\{\mathbf{P}_1^{N_1+k}, k=1, \dots, N_0\}$ are constructed by the recurrence $\mathbf{P}_1^{N_1+k} = a_k \mathbf{W}_k$ in which

$$\mathbf{W}_k = \mathbf{Q}_1^k - \sum_{j=1}^{N_1+k-1} (\mathbf{P}_1^{jT} [G_1] \mathbf{Q}_1^k) \mathbf{P}_1^j,$$

$$a_k = (\mathbf{W}_k^T [G_1] \mathbf{W}_k)^{-1}.$$

Finally, we introduce the $(n_1 \times N_1^S)$ real matrix $[P_1^S]$ such that

$$[P_1^S] = [\mathbf{P}_1^1 \dots \mathbf{P}_1^{N_1} \mathbf{Q}_1^1 \dots \mathbf{Q}_1^{N_0}].$$

It should be noted that if $N_0 = 0$, then $N_1^S = N_1$ and $[P_1^S] = [P_1]$.

IV. CONSTRUCTION OF THE REDUCED MODEL ADAPTED TO MF BAND B

The reduced model adapted to MF band B is obtained^{1,2} by introducing the new variable $\boldsymbol{\theta}_1(\omega) = (\theta_{1,1}(\omega), \dots, \theta_{1,N_1^S}(\omega))$ for the structure and the new variable $\boldsymbol{\theta}_2(\omega) = (\theta_{2,1}(\omega), \dots, \theta_{2,N_2}(\omega))$ for the internal acoustic fluid, such that

$$\mathbf{q}_1(\omega) = [P_1^S] \boldsymbol{\theta}_1(\omega); \quad \mathbf{q}_2(\omega) = [P_2] \boldsymbol{\theta}_2(\omega). \quad (44)$$

From Eqs. (29) and (44), we deduce that for all ω in $B \cup \tilde{B}$, $\{\boldsymbol{\theta}_1(\omega), \boldsymbol{\theta}_2(\omega)\} \in \mathbb{C}^{N_1^S} \times \mathbb{C}^{N_2}$ is the unique solution of the linear equation

$$\begin{bmatrix} [\mathcal{A}_1(\omega)] + \kappa[\mathcal{J}] & i\omega[\mathcal{C}] \\ i\omega[\mathcal{C}]^T & -[\mathcal{A}_2(\omega)] \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1(\omega) \\ \boldsymbol{\theta}_2(\omega) \end{bmatrix} = \begin{bmatrix} \eta(\omega)\mathcal{F}(\omega) \\ 0 \end{bmatrix}, \quad (45)$$

in which $\mathcal{F}(\omega) \in \mathbb{C}^{N_1^S}$ is written as

$$\mathcal{F}(\omega) = [P_1^S]^T \mathbf{F}(\omega), \quad (46)$$

and where

$$[\mathcal{A}_1(\omega)] = [P_1^S]^T [A_1(\omega)] [P_1^S], \quad (47)$$

$$[\mathcal{A}_2(\omega)] = [P_2]^T [A_2(\omega)] [P_2], \quad (48)$$

$$[\mathcal{J}] = [P_1^S]^T [J] [P_1^S] = ([P_1^S]^T \mathbf{\Pi}_2) ([P_1^S]^T \mathbf{\Pi}_2)^T, \quad (49)$$

$$[\mathcal{C}] = [P_1^S]^T [C] [P_2]. \quad (49)$$

From Eqs. (27) and (44), we deduce that for all \mathbf{x} in Ω_1 , the displacement field of the structure is written as

$$\tilde{\mathbf{u}}(\mathbf{x}, \omega) = \sum_{k=1}^{N_1^S} \mathbf{u}_k(\mathbf{x}) \theta_{1,k}(\omega), \quad (50)$$

in which $\mathbf{u}_k(\mathbf{x}) \in \mathbb{C}^3$ is written as

$$\mathbf{u}_k(\mathbf{x}) = \sum_{\alpha=1}^{n_1} [P_1^S]_{\alpha k} \mathbf{u}_\alpha(\mathbf{x}). \quad (51)$$

From Eqs. (35), (44) and (45), we deduce that for all \mathbf{x} in Ω_2 , the pressure field in the internal acoustic fluid is written as

$$\tilde{p}(\mathbf{x}, \omega) = \sum_{k=1}^{N_1^S} \mathcal{P}_k(\mathbf{x}, \omega) \theta_{1,k}(\omega), \quad (52)$$

in which $\mathcal{P}(\mathbf{x}, \omega) = (\mathcal{P}_1(\mathbf{x}, \omega), \dots, \mathcal{P}_{N_1^S}(\mathbf{x}, \omega)) \in \mathbb{C}^{N_1^S}$ is such that

$$\mathcal{P}(\mathbf{x}, \omega) = \omega^2 \rho_2 [\mathcal{C}] [\mathcal{A}_2(\omega)]^{-1} [P_2]^T \boldsymbol{\Psi}(\mathbf{x}) - \kappa [P_1^S]^T \mathbf{\Pi}_2. \quad (53)$$

Equations (45), (50) and (52) constitute the reduced model adapted to MF band B .

V. CONSTRUCTION OF DOMINANT EIGENSUBSPACES

Concerning the construction of the dominant eigensubspace of the energy operator relative to MF band B for the structure on the one hand, and for the internal acoustic fluid

on the other hand, we can use the indirect procedure in the frequency domain or the procedure based on the use of the MF solving method in the time domain presented in Refs. 1 and 2. A detailed analysis of these procedures cannot be reproduced here. Nevertheless, in order to facilitate the understanding of Secs. VII and VIII, we summarize below the main results of the MF solving method in the time domain that we use for the examples. For $a=1$ and $a=2$, the problem defined by Eqs. (36) to (38) is solved by calculating the N_a lowest eigenvalues of the following generalized symmetric eigenvalue problem

$$[G_a][S_a] = [H_a][S_a][\Gamma_a], \quad (54)$$

$$[S_a]^T [H_a][S_a] = [I_a], \quad (55)$$

$$[S_a]^T [G_a][S_a] = [\Gamma_a], \quad (56)$$

for which the subspace iteration algorithm^{1,12-15} is used. The dimension m_a of the subspace used for iterations is such that $N_a < m_a \leq n_a$ with $m_a = \{2N_a, N_a + 8\}$. Consequently, $[S_a]$ is an $(n_a \times m_a)$ real matrix and $[\Gamma_a]$ is a diagonal $(m_a \times m_a)$ real matrix. We have

$$[\tilde{\Lambda}_a] = [\Gamma_a]^{-1}, \quad (57)$$

$$[\tilde{P}_a] = [S_a][\Gamma_a]^{-1/2}, \quad (58)$$

where $[\tilde{P}_a]$ is the $(n_a \times m_a)$ real matrix whose first N_a columns are eigenvectors $\mathbf{P}_a^1, \dots, \mathbf{P}_a^{N_a}$ defining matrix $[P_a]$. For each iteration of the subspace iteration algorithm, we only need to calculate an $(n_a \times m_a)$ real matrix $[W_a] = [E_a] \times [X_a]$, in which $[X_a]$ is a given $(n_a \times m_a)$ real matrix. Let $\chi_0(t)$ be the complex-valued function defined on \mathbb{R} by $\chi_0(t) = e^{-i\omega_B t} \chi(t)$ in which $\chi(t) = (1/2\pi) \int_B e^{i\omega t} \hat{\chi}(\omega) d\omega$ with

$$\tilde{\chi}(\omega) = \frac{1}{\pi} \omega^2 |\eta(\omega)|^2 1_B(\omega). \quad (59)$$

Therefore, χ_0 is an LF signal whose band is $[-\Delta\omega/2, \Delta\omega/2]$. Then it is proved¹ that $[W_a]$ can be calculated by

$$[W_a] = 2\pi \operatorname{Re}\{[Z_a(0)]\}, \quad (60)$$

in which $[Z_a(t)]$ is the solution of the following LF equations in the time domain associated with the MF equations,

$$[M_a][\ddot{Y}_a(t)] + [\tilde{D}_a][\dot{Y}_a(t)] + [\tilde{K}_a][Y_a(t)] = \chi_0(t)[X_a], t \in]-\infty, +\infty[, \quad (61)$$

$$[M_a][\ddot{Z}_a(t)] + [\tilde{D}_a][\dot{Z}_a(t)] + [\tilde{K}_a][Z_a(t)] = [M_a][\overline{Y_a(-t)}], t \in]-\infty, 0[, \quad (62)$$

in which symmetric $(n_a \times n_a)$ complex matrices $[\tilde{D}_a]$ and $[\tilde{K}_a]$ are written, for $a=1$ and $a=2$, as

$$[\tilde{D}_a] = [D_a(\omega_B)] + 2i\omega_B [M_a], \quad (63)$$

$$[\tilde{K}_a] = -\omega_B^2 [M_a] + i\omega_B [D_a(\omega_B)] + [K_a(\omega_B)]. \quad (64)$$

It should be noted that for $a=2$, $[K_2(\omega_B)] = [K_2]$. In addition, if the constraint $\int_{\Omega_2} \psi d\mathbf{x} = 0$ is not included in the dis-

cretization, then, for $a=2$, Eqs. (61) and (62) must be solved with the constraint (see Sec. II)

$$\mathbf{L}^T[Y_2(t)] = [0]; \quad \mathbf{L}^T[Z_2(t)] = [0]. \quad (65)$$

The LF Eqs. (61) and (62) are solved using an unconditionally stable implicit step-by-step integration method. Concerning Eq. (61), time interval $]-\infty, \infty[$ is replaced by the finite interval $]t_I, t_F[$ with the initial conditions $[Y_a(t_I)] = [\dot{Y}_a(t_I)] = [0]$. Concerning Eq. (62), time interval $]-\infty, 0[$ is replaced by the finite interval $]-t_F, 0[$ with the initial conditions $[Z_a(-t_F)] = [\dot{Z}_a(-t_F)] = [0]$.

VI. TIME-STATIONARY RANDOM RESPONSE USING THE REDUCED MODEL

The structural-acoustic system is submitted to a time-stationary second-order centered random wall pressure field $\{p(\mathbf{x}, t), \mathbf{x} \in \Gamma_1, t \in \mathbb{R}\}$ with values in \mathbb{R} and we are interested in the stationary response of the structural-acoustic system. The cross-correlation function of random field p is denoted as $R_p(\mathbf{x}, \mathbf{y}, \tau) = \mathbb{E}\{p(\mathbf{x}, t + \tau)p(\mathbf{y}, t)\}$ in which \mathbb{E} is the mathematical expectation and is such that¹⁶

$$R_p(\mathbf{x}, \mathbf{y}, \tau) = \int_{\mathbb{R}} e^{i\omega\tau} S_p(\mathbf{x}, \mathbf{y}, \omega) d\omega, \quad (66)$$

in which $S_p(\mathbf{x}, \mathbf{y}, \omega)$ is the cross-spectral density function which is written as

$$S_p(\mathbf{x}, \mathbf{y}, \omega) = |\eta(\omega)|^2 s_p(\mathbf{x}, \mathbf{y}, \omega). \quad (67)$$

Let $\{\mathbb{F}_\alpha(t), t \in \mathbb{R}\}$ be the stochastic process defined by

$$\mathbb{F}_\alpha(t) = - \int_{\Gamma_1} p(\mathbf{x}, t) \mathbf{n}_1(\mathbf{x}) \cdot \mathbf{u}_\alpha(\mathbf{x}) ds(\mathbf{x}). \quad (68)$$

Therefore the $(n_1 \times n_1)$ matrix-valued spectral density function $[S_{\mathbb{F}}(\omega)]$ of stationary stochastic process $\mathbb{F} = (\mathbb{F}_1, \dots, \mathbb{F}_{n_1})$ indexed by \mathbb{R} with values in \mathbb{R}^{n_1} is such that

$$[S_{\mathbb{F}}(\omega)]_{\alpha\alpha'} = \int_{\Gamma_1} \int_{\Gamma_1} |\eta(\omega)|^2 s_p(\mathbf{x}, \mathbf{y}, \omega) \{\mathbf{n}_1(\mathbf{x}) \cdot \mathbf{u}_\alpha(\mathbf{x})\} \times \{\mathbf{n}_1(\mathbf{y}) \cdot \mathbf{u}_{\alpha'}(\mathbf{y})\} ds(\mathbf{x}) ds(\mathbf{y}). \quad (69)$$

From Eq. (50), we deduce that, for all \mathbf{x} and \mathbf{y} fixed in Ω_1 , the (3×3) matrix-valued cross-spectral density function $[S_{\tilde{\mathbf{u}}}(\mathbf{x}, \mathbf{y}, \omega)]$ of the \mathbb{R}^3 -valued stochastic field $\{\tilde{\mathbf{u}}(\mathbf{x}, t), \mathbf{x} \in \Omega_1, t \in \mathbb{R}\}$, such that

$$[R_{\tilde{\mathbf{u}}}(\mathbf{x}, \mathbf{y}, \tau)] = \mathbb{E}\{\tilde{\mathbf{u}}(\mathbf{x}, t + \tau)\tilde{\mathbf{u}}(\mathbf{y}, t)^T\} = \int_{\mathbb{R}} e^{i\omega\tau} [S_{\tilde{\mathbf{u}}}(\mathbf{x}, \mathbf{y}, \omega)] d\omega, \quad (70)$$

can be written as

$$[S_{\tilde{\mathbf{u}}}(\mathbf{x}, \mathbf{y}, \omega)] = \sum_{j=1}^{N_1^S} \sum_{k=1}^{N_1^S} [S_{\theta_1}(\omega)]_{jk} \mathcal{U}_j(\mathbf{x}) \mathcal{U}_k(\mathbf{y})^T, \quad (71)$$

in which $[S_{\theta_1}(\omega)]$ is the matrix-valued spectral density function of $\mathbb{R}^{N_1^S}$ -valued mean-square stationary stochastic process $\{\theta_1(t), t \in \mathbb{R}\}$. From Eq. (52), we deduce that, for all \mathbf{x} and \mathbf{y} fixed in Ω_2 , the complex-valued cross-spectral density func-

tion $S_{\tilde{p}}(\mathbf{x}, \mathbf{y}, \omega)$ of the real-valued stochastic field $\{\tilde{p}(\mathbf{x}, t), \mathbf{x} \in \Omega_2, t \in \mathbb{R}\}$, such that

$$R_{\tilde{p}}(\mathbf{x}, \mathbf{y}, \tau) = \mathbb{E}\{\tilde{p}(\mathbf{x}, t + \tau)\tilde{p}(\mathbf{y}, t)\} = \int_{\mathbb{R}} e^{i\omega\tau} S_{\tilde{p}}(\mathbf{x}, \mathbf{y}, \omega) d\omega, \quad (72)$$

can be written as

$$S_{\tilde{p}}(\mathbf{x}, \mathbf{y}, \omega) = \sum_{j=1}^{N_1^S} \sum_{k=1}^{N_1^S} [S_{\theta_1}(\omega)]_{jk} \overline{\mathcal{P}_j(\mathbf{x}, \omega)} \mathcal{P}_k(\mathbf{y}, \omega). \quad (73)$$

From Eqs. (45) and (46), and using linear filtering of stationary stochastic processes,¹⁶⁻¹⁸ we deduce that matrix-valued spectral density function $[S_{\theta_1}(\omega)]$ [appearing in Eqs. (71) and (73)] is written as

$$[S_{\theta_1}(\omega)] = [\mathcal{T}(\omega)] [P_1^S]^T [S_{\mathbb{F}}(\omega)] [P_1^S] [\mathcal{T}(\omega)]^*, \quad (74)$$

in which $(N_1^S \times N_1^S)$ symmetric complex matrix $[\mathcal{T}(\omega)]$ is written as

$$[\mathcal{T}(\omega)] = ([\mathcal{A}_1(\omega)] + \kappa[\mathcal{J}] - \omega^2[\mathcal{C}][\mathcal{A}_2(\omega)]^{-1}[\mathcal{C}]^T)^{-1}. \quad (75)$$

VII. VALIDATION FOR A STRUCTURE HAVING AN MF BEHAVIOR COUPLED WITH AN INTERNAL ACOUSTIC FLUID HAVING AN LF BEHAVIOR IN MF BAND B

This first example concerns an inhomogeneous structure having an MF behavior in a given narrow MF band B , coupled with an internal acoustic cavity filled with a gas and having an LF behavior in MF band B . We consider the MF time-stationary random response of this structural-acoustic system in MF band B , the structure being excited by a time-stationary random wall pressure field.

A. Description of the structural-acoustic system

The structural-acoustic system is referenced to an (x_1, x_2, x_3) coordinate system. The structure is a rectangular thin plate in bending mode, located in plane Ox_1x_2 , to which are attached two point masses, three springs and five dashpots, as shown in Fig. 2. Domain $\Omega_1 = \Gamma_1$ of the plate (middle surface) is rectangular and the plate is simply supported, homogeneous and isotropic, with constant thickness, width $L_1 = 0.5$ m, length $L_2 = 1.0$ m, surface-mass density $\rho_1 = 40$ kg/m², total mass $\mu_1 = \rho_1 L_1 L_2 = 20$ kg and constant damping rate $\xi_1 = 0.002$. We assume that the usual thin plate theory can be used. The thickness, Young's modulus, and Poisson's ratio of the plate are such that the lowest eigenfrequency of the associated conservative plate (without point masses and springs) is 5 Hz. To this plate are attached (1) two point masses having a mass of 3 kg and 4 kg located at points (0.2, 0.4, 0) and (0.35, 0.75, 0) respectively, (2) three springs having the same stiffness coefficient $k = \varepsilon_k \mu_1 \omega_{\text{ref}}^2$ with $\omega_{\text{ref}} = 2\pi \times 550$ rad/s, $\varepsilon_k = 0.1$, located at points (0.221, 0.278, 0), (0.332, 0.537, 0) and (0.443, 0.826, 0), and (3) five dashpots having the same damping coefficient $d = 2\varepsilon_d \mu_1 \xi_1 \omega_{\text{ref}}$ with $\varepsilon_d = 0.1$, located at points (0.154, 0.165, 0), (0.145, 0.334, 0), (0.465, 0.373, 0), (0.247, 0.462, 0), and

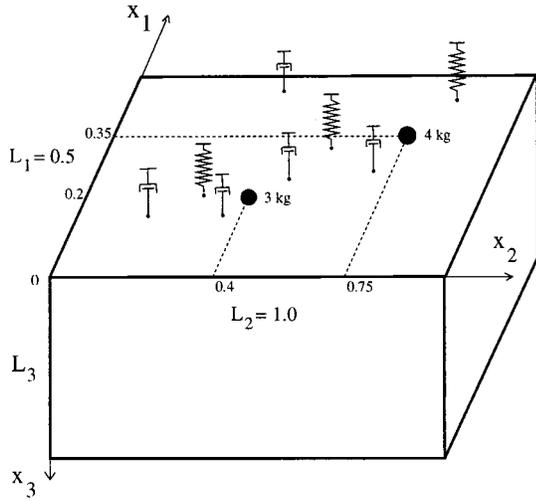


FIG. 2. Rectangular thin plate with point masses, springs, and dashpots, coupled with an acoustic cavity constituted of a bounded rectangular room filled with a gas (air).

(0.268, 0.681, 0). This structure (the plate with point masses, springs, and dashpots) is coupled with an acoustic cavity constituted of a rectangular room Ω_2 (see Fig. 2) bounded by five rigid walls lying along the planes $x_1=0$, $x_1=L_1$, $x_2=0$, $x_2=L_2$, $x_3=L_3$ with $L_3=0.318$ m. The sixth wall lying in plane $x_3=0$ is not rigid and is constituted by the plate. This bounded room is filled with a gas having a constant mass density $\rho_2=1$ kg/m³, speed of sound $c_2=330$ m/s, and dissipative coefficient $\tau=\xi_2/\omega_{\text{ref}}$ with $\xi_2=0.002$. The total mass of the gas is $\mu_2=\rho_2L_1L_2L_3=0.159$ kg. We consider the MF response of this structural-acoustic system in narrow MF band $B=2\pi\times[500,550]$ rad/s, i.e., on the [500, 550] Hz frequency band. The excitation is time-stationary random wall pressure field $\{p(\mathbf{x},t), \mathbf{x}\in\Gamma_1, t\in\mathbb{R}\}$ introduced in Sec. VI, for which the cross-spectral density function defined by Eq. (67) is such that for all ω in B , $|\eta(\omega)|=1$ and $s_p(\mathbf{x},\mathbf{y},\omega)=(L_1L_2)^{-1}\delta_0(x_1-y_1)\delta_0(x_2-y_2)$ where δ_0 is the Dirac delta function on \mathbb{R} at point 0. For the validation, we are interested in calculating the power spectral density function $S_{\text{struc}}(\omega)$ relative to the structure and defined by

$$S_{\text{struc}}(\omega)=\frac{1}{|\eta(\omega)|^2|\Omega_1|}\int_{\Omega_1}\text{tr}[S_{\mathbf{u}}(\mathbf{x},\mathbf{x},\omega)]d\mathbf{x}, \quad (76)$$

and the power spectral density function $S_{\text{acous}}(\omega)$ relative to the acoustic cavity and defined by

$$S_{\text{acous}}(\omega)=\frac{1}{\rho_2^2c_2^4|\eta(\omega)|^2|\Omega_2|}\int_{\Omega_2}[S_{\bar{p}}(\mathbf{x},\mathbf{x},\omega)]d\mathbf{x}. \quad (77)$$

B. Description of the finite dimension approximation

The finite dimension approximation is defined in Sec. II. Concerning the structure *in vacuo* (plate with point masses, springs, and dashpots), family $\mathbf{u}_1, \dots, \mathbf{u}_{n_1}$ is chosen as the first n_1 modes of the associated conservative plate (without point masses and springs) whose corresponding eigenfrequencies are $f_{\text{plate},1}\leq f_{\text{plate},2}\leq \dots \leq f_{\text{plate},n_1}$. For $\alpha=(\alpha_1, \alpha_2)$, $\alpha_1\geq 1$ and $\alpha_2\geq 1$, we have $\mathbf{u}_\alpha(\mathbf{x})=(u_{\alpha,1}(\mathbf{x}), u_{\alpha,2}(\mathbf{x}), u_{\alpha,3}(\mathbf{x}))$ with $u_{\alpha,1}(\mathbf{x})=u_{\alpha,2}(\mathbf{x})=0$ and

$$u_{\alpha,3}(\mathbf{x})=2\sin\left(\frac{\alpha_1\pi x_1}{L_1}\right)\sin\left(\frac{\alpha_2\pi x_2}{L_2}\right).$$

From a convergence study of the structural-acoustic response over the [0, 700] Hz frequency band, we deduced the value of n_1 which is $n_1=407$. We have $f_{\text{plate},1}=5$ Hz and $f_{\text{plate},407}=1097$ Hz. For all α , the normalization of \mathbf{u}_α is such that

$$\int_0^{L_1}\int_0^{L_2}\rho_1|u_{\alpha,3}(x_1,x_2)|^2dx_1dx_2=\mu_1.$$

The $(n_1\times n_1)$ matrices $[M_1]$, $[D_1]$, and $[K_1]$ are independent of the frequency and are dense due to the presence of the point masses, dashpots, and springs. Matrix $[G_1]=L_1L_2[I_1]$ is diagonal, $[I_1]$ being the $(n_1\times n_1)$ unity matrix. The lowest eigenfrequency of the associated conservative structure *in vacuo* (plate with point masses and springs) is $f_{1,1}=7.74$ Hz (to be compared with $f_{\text{plate},1}=5$ Hz). There are 255 modes of the structure *in vacuo* in frequency band [0, 700] Hz, 179 structural modes in frequency band [0, 500] Hz and 20 structural modes in narrow MF band [500, 550] Hz. Then the rank of the first structural mode belonging to frequency band [500, 550] Hz is 180 and consequently, the structure has an MF behavior in frequency band B .

Concerning the acoustic cavity, family $\psi_1, \dots, \psi_{n_2}$ is chosen as the first n_2 acoustic modes of the acoustic cavity with rigid walls whose corresponding eigenfrequencies are $f_{2,1}\leq f_{2,2}\leq \dots \leq f_{2,n_2}$. For $\beta=(\beta_1, \beta_2, \beta_3)$ with $\beta_1+\beta_2+\beta_3\neq 0$ and $\beta_1, \beta_2, \beta_3\geq 0$, we have

$$\psi_\beta(\mathbf{x})=2\sqrt{2}\cos\left(\frac{\beta_1\pi x_1}{L_1}\right)\cos\left(\frac{\beta_2\pi x_2}{L_2}\right)\cos\left(\frac{\beta_3\pi x_3}{L_3}\right).$$

These functions satisfy the constraint $\int_{\Omega_2}\psi_\beta(\mathbf{x},\omega)d\mathbf{x}=0$. From a convergence study of the structural-acoustic response over the [0, 700] Hz frequency band, we deduced the value of n_2 which is $n_2=19$. For all β , the normalization of ψ_β is such that

$$\int_0^{L_1}\int_0^{L_2}\int_0^{L_3}\rho_2|\psi_\beta(\mathbf{x})|^2d\mathbf{x}=\mu_2.$$

The $(n_2\times n_2)$ matrices $[M_2]$, $[D_2]$, and $[K_2]$ are independent of the frequency and are diagonal. Matrix $[G_2]=L_1L_2L_3[I_2]$ is diagonal, $[I_2]$ being the $(n_2\times n_2)$ unity matrix. The lowest acoustic eigenfrequency of the acoustic cavity with rigid walls is $f_{2,1}=165$ Hz. There are 16 acoustic modes in frequency band [0, 700] Hz, 6 acoustic modes in frequency band [0, 500] Hz, and 2 acoustic modes in narrow MF band [500, 550] Hz. Then the rank of the first acoustic mode in MF band [500, 550] Hz is 7 and consequently, the acoustic cavity has an LF behavior in MF band B .

For $\alpha=(\alpha_1, \alpha_2)$, component $\Pi_{2,\alpha}$ of vector $\mathbf{\Pi}_2\in\mathbb{R}^{n_1}$, defined by Eq. (32), is such that

$$\Pi_{2,\alpha}=-\frac{2L_1L_2}{\alpha_1\alpha_2\pi^2}\times\{\cos(\alpha_1\pi)-1\}\times\{\cos(\alpha_2\pi)-1\}. \quad (78)$$

For $\alpha=(\alpha_1, \alpha_2)$ and $\beta=(\beta_1, \beta_2, \beta_3)$, coupling $(n_1\times n_2)$ real matrix $[C]$, defined by Eq. (34), is such that

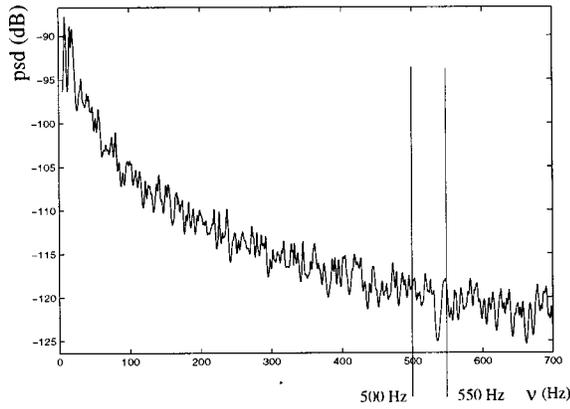


FIG. 3. Graph of function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struct}}^{\text{ref}}(2\pi\nu))$ corresponding to the reference solution for the MF structure coupled with the LF internal acoustic fluid.

$$\begin{aligned}
 [C]_{\alpha\beta} = & -\frac{\sqrt{2}\rho_2 L_1 L_2}{\pi^2} \left\{ \frac{1 - \cos\{\pi(\alpha_1 + \beta_1)\}}{\alpha_1 + \beta_1} \right. \\
 & \left. + \frac{1 - \cos\{\pi(\alpha_1 - \beta_1)\}}{\alpha_1 - \beta_1} \right\} \\
 & \times \left\{ \frac{1 - \cos\{\pi(\alpha_2 + \beta_2)\}}{\alpha_2 + \beta_2} \right. \\
 & \left. + \frac{1 - \cos\{\pi(\alpha_2 - \beta_2)\}}{\alpha_2 - \beta_2} \right\}. \quad (79)
 \end{aligned}$$

Finally, from Eq. (69), we deduce that $[S_F(\omega)]_{\alpha\alpha'} = |\eta(\omega)|^2 \delta_{\alpha\alpha'}$.

C. Constructing the reference solution on a broad frequency band

The reference solution is constructed on the $[5, 700]$ Hz broad frequency band with a sampling frequency step $\delta\nu = 0.165$ Hz. The power spectral density functions $S_{\text{struct}}^{\text{ref}}(\omega)$ and $S_{\text{acous}}^{\text{ref}}(\omega)$ corresponding to the reference solution and defined by Eqs. (76) and (77) are directly calculated in the frequency domain without using the reduced model. From Eqs. (76), (77), (27), (29), and (35), we deduce that

$$S_{\text{struct}}^{\text{ref}}(\omega) = \text{tr}\{[T(\omega)][T(\omega)]^*\}, \quad (80)$$

$$\begin{aligned}
 S_{\text{acous}}^{\text{ref}}(\omega) = & \frac{1}{\rho_2^2 c_2^4} (\text{tr}\{([V(\omega)][T(\omega)])([V(\omega)][T(\omega)]^*)\} \\
 & + \kappa^2([T(\omega)]\mathbf{\Pi}_2)^*([T(\omega)]\mathbf{\Pi}_2)), \quad (81)
 \end{aligned}$$

in which $[T(\omega)]$ is the symmetric $(n_1 \times n_1)$ complex matrix which is written as

$$[T(\omega)] = ([A_1(\omega)] + \kappa[J] - \omega^2[C][A_2(\omega)]^{-1}[C]^T)^{-1},$$

where $[A_1(\omega)]$, $[J]$, $[A_2(\omega)]$ and $[C]$ are the matrices defined by Eqs. (30), (32) with (78), (33), and (79), respectively. In Eq. (81), $(n_2 \times n_1)$ complex matrix $[V(\omega)]$ is defined by

$$[V(\omega)] = \omega^2 \rho_2 [A_2(\omega)]^{-1} [C]^T.$$

Figures 3 and 4 show the mappings $\nu \rightarrow 10$

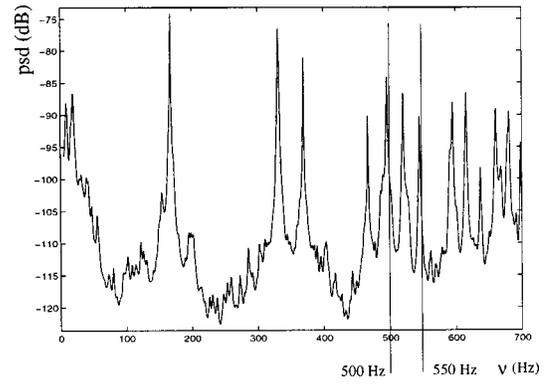


FIG. 4. Graph of function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ corresponding to the reference solution for the LF internal acoustic fluid coupled with the MF structure.

$\times \log_{10}(S_{\text{struct}}^{\text{ref}}(2\pi\nu))$ and $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$, respectively, over the $[5, 700]$ Hz broad frequency band.

D. Reference solution on the narrow MF band

The reference solution on narrow MF band $B = [500, 550]$ Hz is constructed as explained in Sec. VII C using the frequency resolution $\delta\nu = 0.166$ Hz. The graphs of functions $\nu \rightarrow 10 \times \log_{10}(S_{\text{struct}}^{\text{ref}}(2\pi\nu))$ and $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ (corresponding to Figs. 3 and 4) of the reference solution on narrow MF band B are used below (see the solid lines in Figs. 7–10) to evaluate the accuracy of the response constructed using the reduced model.

E. Constructing the dominant eigensubspaces

For the structure, the dominant eigensubspace of the energy operator relative to band B for the structure *in vacuo* [related to matrix $[E_1]$ defined by Eqs. (41) to (43)] is constructed using the method presented in Sec. V with $N_1 = 50$. Figure 5 shows the graph of the function $j \rightarrow 10 \times \log_{10}(\lambda_j^1)$ for $j \in \{1, 2, \dots, 50\}$ in which $\lambda_1^1, \dots, \lambda_{50}^1$ are the highest eigenvalues of the generalized symmetric eigenvalue problem defined by Eq. (36) for $a=1$. There is a strong decrease in the eigenvalues which means there exists the possibility of constructing an efficient reduced model for the structure. Figure 5 shows that the order N_1 of the reduced model is about 25 for band B .

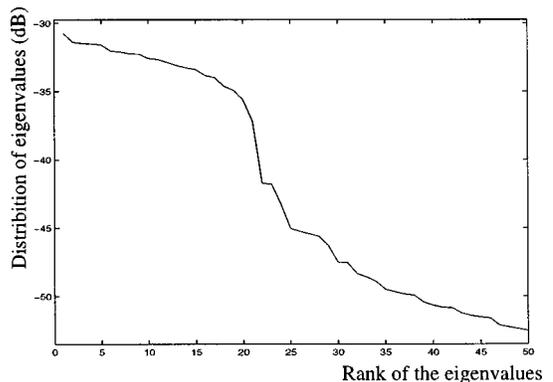


FIG. 5. Graph of function $j \rightarrow 10 \times \log_{10}(\lambda_j^1)$ showing the distribution of eigenvalues λ_j^1 of the energy operator of the MF structure *in vacuo*.

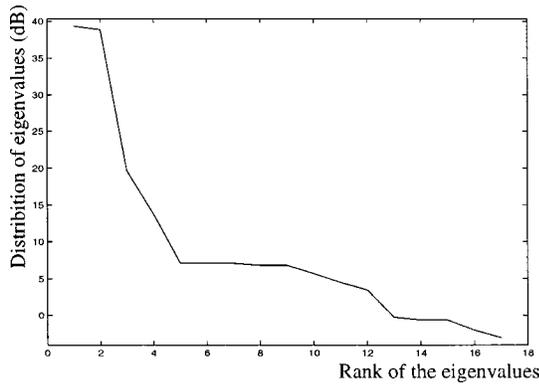


FIG. 6. Graph of function $j \rightarrow 10 \times \log_{10}(\lambda_2^j)$ showing the distribution of eigenvalues λ_2^j of the energy operator of the LF internal acoustic cavity with rigid walls.

For the internal acoustic fluid, the dominant eigensubspace of the energy operator relative to band B for the internal acoustic cavity with rigid walls [related to matrix $[E_2]$ defined by Eqs. (41) to (43)] is constructed using the method presented in Sec. V with $N_2=17$. Figure 6 shows the graph of the function $j \rightarrow 10 \times \log_{10}(\lambda_2^j)$ for $j \in \{1, 2, \dots, 17\}$ in which $\lambda_2^1, \dots, \lambda_2^{17}$ are the highest eigenvalues of the generalized symmetric eigenvalue problem defined by Eq. (36) for $a=2$. Figure 6 shows that the order N_2 of the reduced model is about 4 for band B .

F. Reduced model adapted to the narrow MF band

In this section, we present a comparison of the reference solution constructed in Sec. VII D with the solution obtained by the reduced model constructed using the results of Secs. IV and VI and Eqs. (76) and (77). Two reduced models are considered below. For the first reduced model, the parameters are $N_1^S = N_1 = 25$ (then $N_0 = 0$) and $N_2 = 4$. Figure 7 is related to the structure and shows the comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ (dashed line corresponding to the first reduced model) on narrow MF band [500, 550] Hz. Figure 8 is related to the internal acoustic fluid and shows the com-

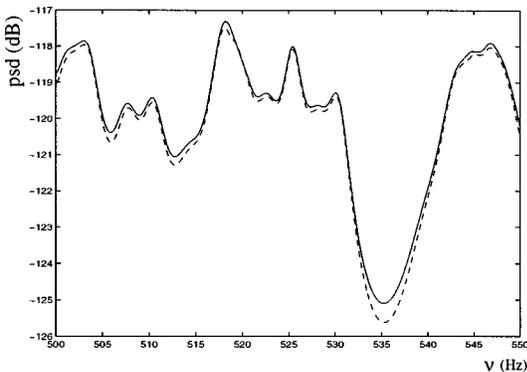


FIG. 7. Reduced model of the dynamical response of the MF structure coupled with the LF internal acoustic fluid: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ [reduced model for $N_1=25$, $N_0=0$, and $N_2=4$ (dashed line)].

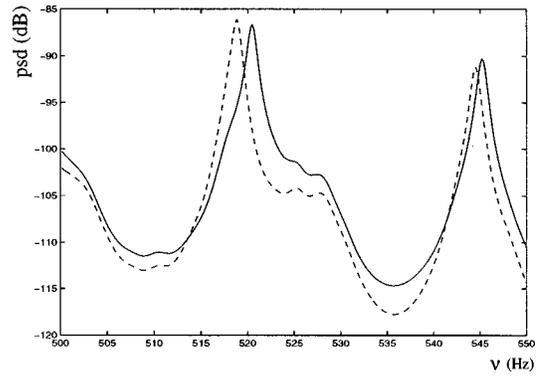


FIG. 8. Reduced model of the dynamical response of the LF internal acoustic fluid coupled with the MF structure: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ [reduced model for $N_1=25$, $N_0=0$, and $N_2=4$ (dashed line)].

parison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ (dashed line corresponding to the first reduced model) on narrow MF band [500, 550] Hz. Figure 7 shows that the comparison is very good for the structure, but Fig. 8 shows that this first reduced model ($N_0=0$) is not so good for the internal acoustic fluid. The problem which appears in Fig. 8 has been explained in Secs. III C and III D; to accelerate convergence of the reduced model for prediction of the acoustic pressure in the internal acoustic fluid, it is necessary to take $N_0 > 0$ (see below).

For the second reduced model, the parameters are $N_1 = 25$, $N_0 = 10$ (then $N_1^S = 35$) and $N_2 = 4$. Figure 9 is related to the structure and shows the comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ (dashed line corresponding to the second reduced model) on narrow MF band [500, 550] Hz. Figure 10 is related to the internal acoustic fluid and shows the comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ (dashed line corresponding to the second reduced model) on narrow MF band [500, 550] Hz. Fig-

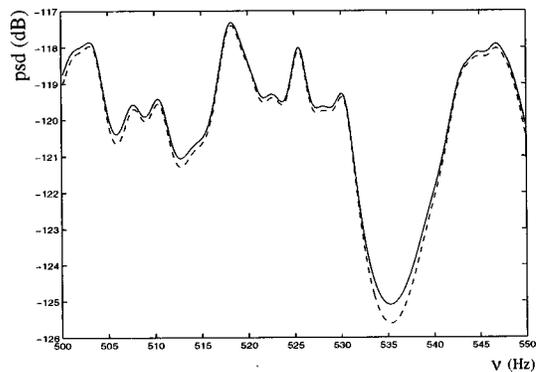


FIG. 9. Reduced model of the dynamical response of the MF structure coupled with the LF internal acoustic fluid: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ [reduced model for $N_1=25$, $N_0=10$, and $N_2=4$ (dashed line)].

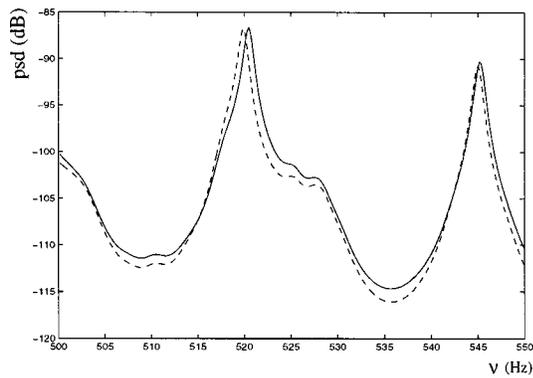


FIG. 10. Reduced model of the dynamical response of the LF internal acoustic fluid coupled with the MF structure: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ [reduced model for $N_1=25$, $N_0=10$, and $N_2=4$ (dashed line)].

ure 9 is quite similar to Fig. 7 and shows that the comparison is very good for the structure and Fig. 10 shows that this second reduced model ($N_0=10$) is good. Consequently, for this structural-acoustic system whose structure has an MF behavior and whose internal acoustic fluid has an LF behavior (MF-LF coupling), the reduced model developed in Sec. IV based on the use of the structural vector basis constructed in Sec. III D is validated.

VIII. VALIDATION FOR A STRUCTURE HAVING AN MF BEHAVIOR COUPLED WITH AN INTERNAL ACOUSTIC FLUID HAVING AN MF BEHAVIOR IN MF BAND B

This second example concerns the inhomogeneous structure defined in Sec. VII, having an MF behavior in given narrow MF band B , coupled with an internal acoustic cavity filled with a gas and having an MF behavior in MF band B . We consider the MF time-stationary random response of this structural-acoustic system in MF band B , the structure being excited by the time-stationary random wall pressure field used in Sec. VII.

A. Description of the structural-acoustic system

We consider the structural-acoustic system defined in Sec. VII A for which acoustic rectangular room Ω_2 is such that $L_3=9.0$ m (instead of $L_3=0.318$ m, see Fig. 2). The total mass of the gas is $\mu_2=\rho_2 L_1 L_3=4.5$ kg. We consider the MF response of this structural-acoustic system in narrow MF band $B=[500,550]$ Hz. The excitation is the time-stationary random wall pressure field defined in Sec. VII A. For the validation, we are interested in calculating the power spectral density functions defined by Eqs. (76) and (77).

B. Description of the finite dimension approximation

The finite dimension approximation of the structure *in vacuo* (plate with point masses, springs, and dashpots) is defined in Sec. VII B. We have $n_1=407$. There are 255 modes of the structure *in vacuo* in frequency band $[0, 700]$ Hz, 179 structural modes in frequency band $[0, 500]$ Hz, and 20 structural modes in narrow MF band $[500, 550]$ Hz. Then

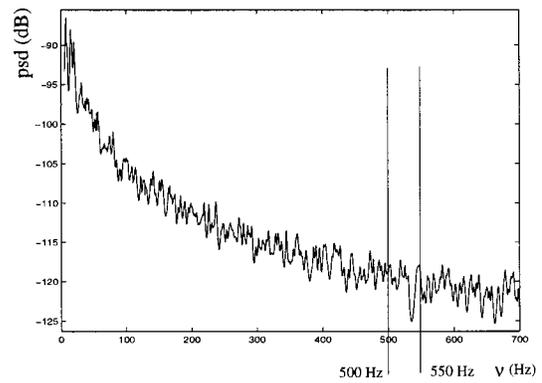


FIG. 11. Graph of function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ corresponding to the reference solution for the MF structure coupled with the MF internal acoustic fluid.

the rank of the first structural mode belonging to frequency band $[500, 550]$ Hz is 180 and consequently, the structure has an MF behavior in frequency band B .

Concerning the acoustic cavity, we use the finite dimension approximation defined in Sec. VII B for which the value of n_2 has to be increased. From a convergence study of the structural-acoustic response over the $[0, 700]$ Hz frequency band, we deduced the value of n_2 which is $n_2=353$ (instead of 19). The lowest acoustic eigenfrequency of the acoustic cavity with rigid walls is $f_{2,1}=18.33$ Hz. There are 289 acoustic modes in frequency band $[0, 700]$ Hz, 128 acoustic modes in frequency band $[0, 500]$ Hz, and 34 acoustic modes in narrow MF band $[500, 550]$ Hz. Then the rank of the first acoustic mode in MF band $[500, 550]$ Hz is 129 and consequently, the acoustic cavity has an MF behavior in MF band B .

C. Constructing the reference solution on a broad frequency band

The reference solution is constructed on the $[5, 700]$ Hz broad frequency band with a sampling frequency step $\delta\nu=0.165$ Hz. The power spectral density functions $S_{\text{struc}}^{\text{ref}}(\omega)$ and $S_{\text{acous}}^{\text{ref}}(\omega)$ corresponding to the reference solution and defined by Eqs. (76) and (77) are calculated directly in the frequency domain, using Eqs. (80) and (81), without using the reduced model. Figures 11 and 12 show the mappings

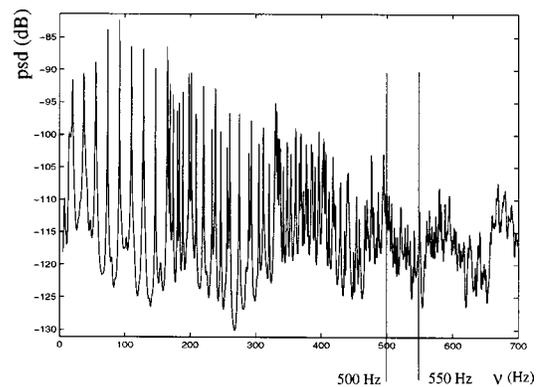


FIG. 12. Graph of function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ corresponding to the reference solution for the MF internal acoustic fluid coupled with the MF structure.

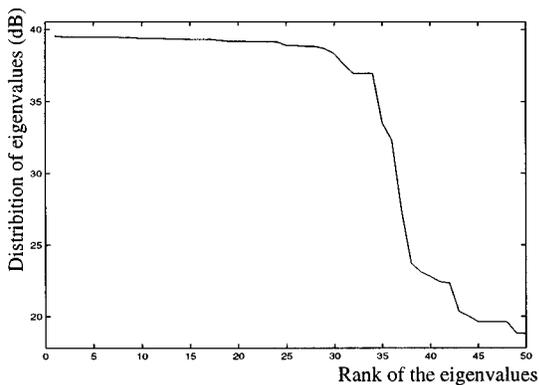


FIG. 13. Graph of function $j \rightarrow 10 \times \log_{10}(\lambda_2^j)$ showing the distribution of eigenvalues λ_2^j of the energy operator of the MF internal acoustic cavity with rigid walls.

$\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ and $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ respectively over the [5, 700] Hz broad frequency band.

D. Reference solution on the narrow MF band

The reference solution on narrow MF band $B = [500, 550]$ Hz is constructed as explained in Sec. VIII C using the frequency resolution $\delta\nu = 0.166$ Hz. The graphs of functions $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ and $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ (corresponding to Figs. 11 and 12) of the reference solution on narrow MF band [500, 550] Hz are used below to evaluate the accuracy of the response constructed using the reduced model (see the solid lines in Figs. 14–17).

E. Constructing the dominant eigensubspaces

For the structure, we use the dominant eigenspace calculated in Sec. VII E. For the internal acoustic fluid, the dominant eigensubspace of the energy operator relative to band B for the internal acoustic cavity with rigid walls [related to matrix $[E_2]$ defined by Eqs. (41) to (43)] is constructed using the method presented in Sec. V with $N_2 = 50$. Figure 13 shows the graph of the function $j \rightarrow 10 \times \log_{10}(\lambda_2^j)$ for $j \in \{1, 2, \dots, 50\}$ in which $\lambda_2^1, \dots, \lambda_2^{50}$ are the

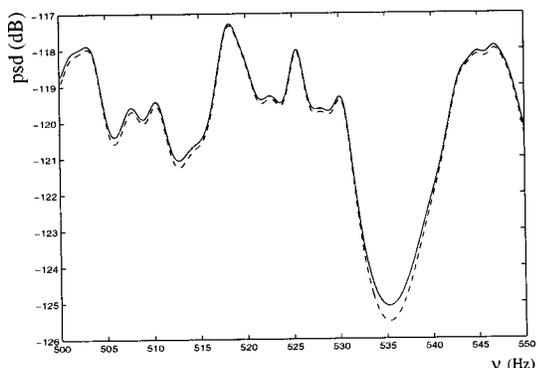


FIG. 14. Reduced model of the dynamical response of the MF structure coupled with the MF internal acoustic fluid: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ [reduced model for $N_1 = 30$, $N_0 = 0$, and $N_2 = 45$ (dashed line)].

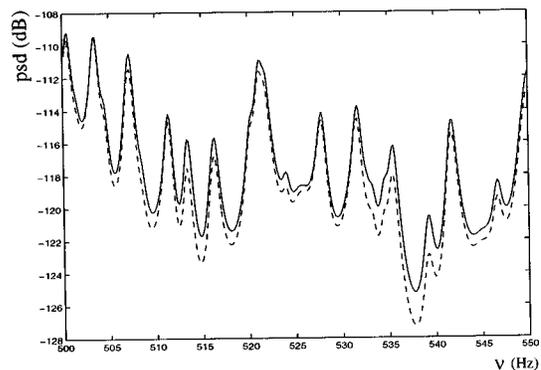


FIG. 15. Reduced model of the dynamical response of the MF internal acoustic fluid coupled with the MF structure: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ [reduced model for $N_1 = 30$, $N_0 = 0$, and $N_2 = 45$ (dashed line)].

highest eigenvalues of the generalized symmetric eigenvalue problem defined by Eq. (36) for $a = 2$. Figure 13 shows that the order N_2 of the reduced model is about 40 for band B .

F. Reduced model adapted to the narrow MF band

In this section, we present a comparison of the reference solution constructed in Sec. VIII D with the solution obtained by the reduced model constructed using the results of Secs. IV and VI and Eqs. (76) and (77). As for the first example presented in Sec. VII, two reduced models are considered below. For the first reduced model, the parameters are $N_1^S = N_1 = 30$ (then $N_0 = 0$) and $N_2 = 45$. Figure 14 is related to the structure and shows the comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ (dashed line corresponding to the first reduced model) on narrow MF band [500, 550] Hz. Figure 15 is related to the internal acoustic fluid and shows the comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ (dashed line corresponding to the first reduced model) on narrow MF band [500, 550] Hz. Figures 14 and 15 show that the comparison is good for both the structure and the internal acoustic fluid. For the second reduced model, the parameters are $N_1 = 30$, $N_0 = 10$ (then $N_1^S = 40$) and $N_2 = 45$. Figure 16 is related to the structure and shows the comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ (dashed line corresponding to the second reduced model) on narrow MF band [500, 550] Hz. Figure 17 is related to the internal acoustic fluid and shows the comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ (solid line corresponding to the reference solution) and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ (dashed line corresponding to the second reduced model) on narrow MF band [500, 550] Hz. Figure 16 is quite similar to Fig. 14 and Fig. 17 to Fig. 15. This means that for an MF–MF coupling, the reduced model of the structure can be constructed with $N_0 = 0$.

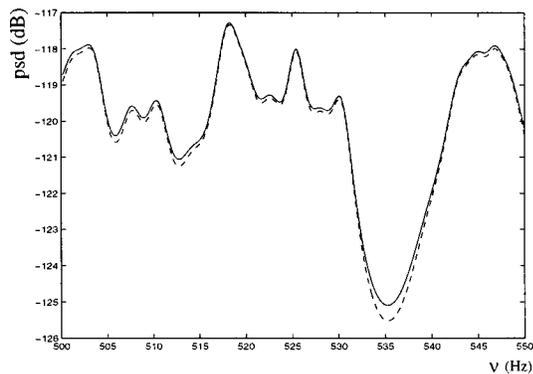


FIG. 16. Reduced model of the dynamical response of the MF structure coupled with the MF internal acoustic fluid: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{struc}}(2\pi\nu))$ [reduced model for $N_1=30$, $N_0=10$, and $N_2=45$ (dashed line)].

IX. CONCLUSIONS

A theoretical approach is presented for constructing a reduced model in the MF range in the area of structural acoustics for a general three-dimensional anisotropic, inhomogeneous, viscoelastic, bounded structure with an arbitrary geometry coupled with an internal acoustic fluid (gas or liquid). The boundary value problem formulated in the frequency domain and its variational formulation are presented. For a given MF band, the energy operator of the structure *in vacuo* and the energy operator of the internal acoustic cavity with rigid walls are positive-definite symmetric operators which have a countable set of decreasing positive eigenvalues. The eigenfunctions corresponding to the highest eigenvalues (dominant eigensubspace) of each energy operator constitute an appropriate functional basis of the corresponding admissible function space for the structure and for the internal acoustic fluid. For an MF structure (i.e., a structure having an MF behavior) coupled with an MF internal acoustic fluid (i.e., an internal acoustic fluid having an MF behavior) in the MF band considered, these two functional bases allow a reduced model of the structural-acoustic system to be constructed using the Ritz–Galerkin method. If the MF structure is coupled with an LF internal acoustic fluid, then it is more efficient to complete the structural vector basis relative to the MF band with a few LF structural modes in order to accelerate convergence of the solution for the internal acoustic cavity response in the MF band. A finite dimension approximation of the continuous case is introduced in a general context (using the finite element method or not). For construction of the dominant eigensubspace of each energy operator, an efficient procedure based on the use of the subspace iteration method is proposed. It does not require explicit calculation of the energy operator. We then obtain an efficient method for constructing a reduced model in the MF range: (1) an intrinsic reduced model is constructed for general structural-acoustic systems and can be considered as a progress independently of any computer time aspects; (2) the efficiency of such a reduced model in the MF range can be very high with respect to any other techniques validated in the MF range as soon as the MF responses have to be calculated for a large number of multiple deterministic and ran-

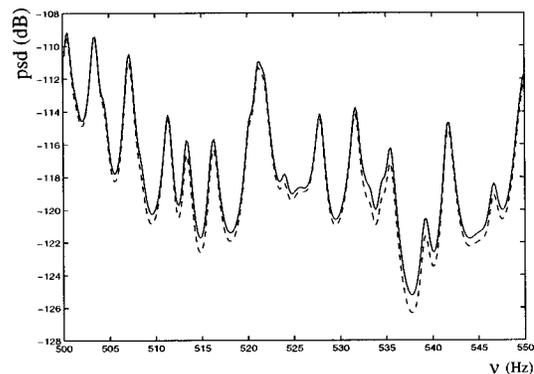


FIG. 17. Reduced model of the dynamical response of the MF internal acoustic fluid coupled with the MF structure: comparison between function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}^{\text{ref}}(2\pi\nu))$ [reference solution (solid line)] and function $\nu \rightarrow 10 \times \log_{10}(S_{\text{acous}}(2\pi\nu))$ [reduced model for $N_1=30$, $N_0=10$, and $N_2=45$ (dashed line)].

dom loads, particularly for random excitations; (3) the implementation of this method is easily to perform in any usual finite element computer code. Finally, concerning the structure, the results presented can be extended straightforwardly to a structure made of beams, plates, and shells. Analysis of the two examples (MF-LF coupling and MF–MF coupling) validates the reduced model methodology proposed for the MF range.

- ¹C. Soize, “Reduced models in the medium-frequency range for general dissipative structural-dynamics systems,” *Eur. J. Mech. A/Solids* **17**, 657–685 (1998).
- ²C. Soize, “Reduced models in the medium-frequency range for general external structural-acoustic systems,” *J. Acoust. Soc. Am.* **103**, 3393–3406 (1998).
- ³R. Ohayon and C. Soize, *Structural Acoustics and Vibration* (Academic, San Diego, 1998).
- ⁴Y. C. Fung, *Foundations of Solid Mechanics* (Prentice-Hall, Englewood Cliffs, NJ, 1968).
- ⁵C. Truesdell, *Mechanics of Solids, Vol. III, Theory of Viscoelasticity, Plasticity, Elastic Waves and Elastic Stability* (Springer-Verlag, Berlin, 1984).
- ⁶P. G. Ciarlet, *Mathematical Elasticity, Vol. I: Three-Dimensional Elasticity* (North-Holland, Amsterdam, 1988).
- ⁷O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method* (McGraw-Hill, New York, 1989).
- ⁸R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology* (Springer-Verlag, Berlin, 1992).
- ⁹P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
- ¹⁰A. Leissa, *Vibration of Plates and Vibration of Shells* (Acoustical Society of America Publications on Acoustics, Woodbury, 1993). Originally published by NASA, 1973.
- ¹¹A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America Publications on Acoustics, Woodbury, 1993). Originally published by McGraw-Hill, New York, 1981.
- ¹²K. J. Bathe and E. L. Wilson, *Numerical Methods in Finite Element Analysis* (Prentice-Hall, New York, 1976).
- ¹³B. N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, NJ, 1980).
- ¹⁴G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. (The Johns Hopkins University Press, Baltimore, 1989).
- ¹⁵F. Chatelin, *Eigenvalues of Matrices* (Wiley, New York, 1993).
- ¹⁶P. Kree and C. Soize, *Mathematics of Random Phenomena* (Reidel, Dordrecht, 1986).
- ¹⁷C. Soize, *The Fokker-Planck Equation for Stochastic Dynamical Systems and its Explicit Steady State Solutions* (World Scientific, Singapore, 1994).
- ¹⁸Y. K. Lin, *Probabilistic Theory of Structural Dynamics* (McGraw-Hill, New York, 1967).

Three-dimensional vibration analysis of a cantilevered parallelepiped: Exact and approximate solutions

C. W. Lim

Department of Mechanical Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

(Received 6 April 1999; accepted for publication 20 August 1999)

The paper investigates the hypothetical assumption of neglecting transverse normal stress in vibration analysis for cantilevered thick plates and rectangular parallelepiped. The analysis solves the three-dimensional elasticity energy functional including, as well as excluding, transverse normal stress and obtains free vibration solutions for a cantilevered parallelepiped. Although it is widely accepted, the omission of transverse normal stress is well justified in Kirchhoff–Love thin-plate theory and higher-order thick-plate models; the transverse normal stress effects and thickness extent to which the thick-plate models apply as the thickness increases are practically unknown. The inconsistency of assuming constant transverse normal displacement through thickness for thick-plate models is also addressed. The paper concludes that for a rectangular parallelepiped with thickness exceeding a certain limit, there is considerable discrepancy if transverse normal stress is neglected. © 1999 Acoustical Society of America. [S0001-4966(99)02312-7]

PACS numbers: 43.40.Dx [CBB]

INTRODUCTION

For decades, the analysis of thick plates has neglected the effects of transverse normal stress. The Reissner–Mindlin first-order plate theory (Reissner, 1945; Mindlin, 1951) extended the classical Kirchhoff–Love thin-plate theory to analyze plates with considerable thickness. Transverse shear deformation was considered by including transverse shear strain effects in the analysis. The first-order theory assumes constant transverse shear strains through the plate thickness and renders a paradoxical implication that the transverse shear strain components do not vanish on the top and bottom surfaces. A shear correction factor ($\kappa = \pi^2/12$) was therefore derived by Reissner (1945) to account for this deficiency. Using this first-order theory, accurate vibration formulation and solutions have been reported for laminated curved beams (Qatu, 1993) and plates (Bert and Chen, 1978).

The inadequacy of the first-order shear deformation theory to overcome the nonvanishing shear strain effects stimulated the development of research in thick plates with various formulations of higher-order theories. One of the earliest attempts was initiated by Soler (1968), who expressed all dependent variables including displacement and stress components in Legendre polynomials. Other developments in the higher-order plate theory include Whitney and Sun (1973) with quadratic and linear distributions for in-plane and transverse displacements; Whitney and Sun (1974) with linear and quadratic distributions for in-plane and transverse displacements; and Iyengar *et al.* (1974) and Lo *et al.* (1977a,b) with cubic and quadratic distributions for in-plane and transverse displacements. Washizu (1980) expressed the in-plane displacement in a power series of transverse coordinate (z) and simplified the function to a first-order expression. By imposing zero transverse shear-stress conditions at the free surfaces, Levinson (1980) developed a third-order plate theory with cubic in-plane displacement and constant transverse displacement without the requirement of a shear

correction factor. Perhaps the most remarkable work on the third-order shear deformation theory for thick plates was attributed to Reddy and his associates (1984a, b, 1985, 1989) based on a similar approach as Levinson (1980), resulting in a parabolic transverse shear strain distribution in thickness. This approach had been extended to investigate the numerical aspects and effects of various boundary conditions on free vibration of thick plates (Lim *et al.*, 1998a,b) and singly and doubly curved shallow shells (Lim and Liew, 1995; Liew and Lim, 1996).

Similar to the hypothesis of Kirchhoff–Love plate theory, the Reissner–Mindlin first-order and Levinson–Reddy higher-order plate theories do not consider transverse normal stress. Some authors have expressed concern over the validity of this hypothesis. Gould (1988) suggested that the concept of transverse inextensibility ($w = w_0$) must be reviewed in analysis of thick plates and shells with transverse shear flexibility. He expressed that the inclusion of transverse shearing strains to extend the bounds to include somewhat thicker plates and shells is difficult to quantify since a true thick-plate or shell theory should account for transverse normal stress as well.

The transverse normal stress is considered in three-dimensional elastic analysis of solids. However, such three-dimensional elastic solutions are particularly scarce. Some of the investigations are concerned with rods and beams (Hutchinson, 1971; Hutchinson, 1981; Leissa and So, 1995a), parallelepiped (Fromme and Leissa, 1970; Hutchinson and Zillmer, 1983; Leissa and Zhang, 1983; Liew *et al.*, 1995a), solid and hollow cylinders (Hutchinson, 1967; Hutchinson, 1980; Leissa and So, 1995b; Liew *et al.*, 1995b; So and Leissa, 1997), truncated hollow cones (Leissa and So, 1995c), and open shells (Lim *et al.*, 1998c).

To the author's knowledge, direct comparison of three-dimensional elasticity solutions including and excluding transverse normal stress is only available in Hutchinson (1979, 1984). In these two papers, Hutchinson analyzed the

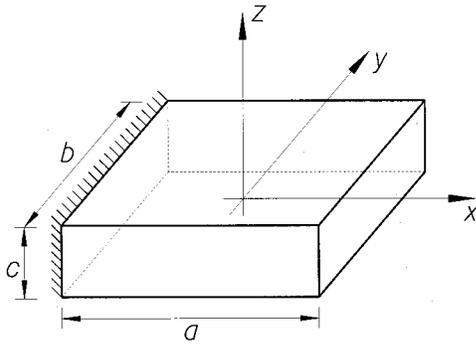


FIG. 1. Geometry of a rectangular parallelepiped.

vibration of thick, free circular plates using the Mathieu series solution and modified Pickett method to obtain exact solutions. The approximate solutions were obtained by the Mindlin thick-plate model including shear deformation and rotary inertia. These analyses concluded that, for thick, free circular plates, the approximate solution yields frequencies of sufficient accuracy for most engineering application within the range of applicability of the approximate theory. However, results are not available for thick rectangular plates. It is the key objective of this paper to address the validity of the hypothetical neglect of transverse normal stress and transverse inextensibility in the specific case of free vibration of a cantilevered rectangular parallelepiped. The assumption of transversely inextensible displacement through thickness will also be examined. This paper also intends to determine to what extent the omission of transverse normal stress in the Kirchhoff–Love thin-plate theory is applicable to cantilevered thick-plate vibrations.

I. FORMULATION

A. Basic definition

An isotropic parallelepiped of length a , width b , and thickness c is illustrated in Fig. 1. With respect to a Cartesian coordinate system with origin located at the center of the body, a point within the body is designated by (x, y, z) . The parallelepiped is clamped at a surface $x = -a/2$ with all other surfaces free.

B. Three-dimensional strain and kinetic energy expressions

For linear, elastic free vibration, the strain energy of a three-dimensional solid is

$$U = \frac{1}{2} \int \int \int_V [(\Delta + 2G)(\epsilon_{xx}^2 + \epsilon_{yy}^2 + \epsilon_{zz}^2) + 2\Delta(\epsilon_{xx}\epsilon_{yy} + \epsilon_{yy}\epsilon_{zz} + \epsilon_{zz}\epsilon_{xx}) + G(\gamma_{yz}^2 + \gamma_{zx}^2 + \gamma_{xy}^2)] dx dy dz, \quad (1)$$

where V is the volume, G is the shear modulus and

$$\Delta = \frac{\nu E}{(1 + \nu)(1 - 2\nu)}, \quad (2a)$$

$$G = \frac{E}{2(1 + \nu)}, \quad (2b)$$

in which E is the Young's modulus.

Let the displacements in the x, y, z directions be u, v, w , respectively. The normal and shear strain-displacement relations are

$$\epsilon_{xx} = \frac{\partial u}{\partial x}, \quad (3a)$$

$$\epsilon_{yy} = \frac{\partial v}{\partial y}, \quad (3b)$$

$$\epsilon_{zz} = \frac{\partial w}{\partial z}, \quad (3c)$$

$$\gamma_{yz} = \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}, \quad (3d)$$

$$\gamma_{zx} = \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}, \quad (3e)$$

$$\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}. \quad (3f)$$

The kinetic energy is

$$T = \frac{\rho}{2} \int \int \int_V \left[\left(\frac{\partial u}{\partial t} \right)^2 + \left(\frac{\partial v}{\partial t} \right)^2 + \left(\frac{\partial w}{\partial t} \right)^2 \right] dV, \quad (4)$$

where ρ is the mass density per unit volume.

For small deformation vibration, the displacement components assume temporal simple harmonic functions in the forms of

$$u(\bar{x}, \bar{y}, \bar{z}, t) = U(\bar{x}, \bar{y}, \bar{z}) \sin \omega t, \quad (5a)$$

$$v(\bar{x}, \bar{y}, \bar{z}, t) = V(\bar{x}, \bar{y}, \bar{z}) \sin \omega t, \quad (5b)$$

$$w(\bar{x}, \bar{y}, \bar{z}, t) = W(\bar{x}, \bar{y}, \bar{z}) \sin \omega t, \quad (5c)$$

where U, V, W are the displacement amplitude functions, ω is the angular frequency of vibration, and

$$\bar{x} = \frac{x}{a}, \quad (6a)$$

$$\bar{y} = \frac{y}{b}, \quad (6b)$$

$$\bar{z} = \frac{z}{c}, \quad (6c)$$

are the nondimensional coordinates.

For a nondissipative system, the total energy in a vibration cycle is conserved. The maximum strain and kinetic energy integral expressions U_{\max} and T_{\max} can be derived easily by substituting Eqs. (5a)–(5c) into Eqs. (1) and (4) and determining the extremum with respect to time t , as

$$\begin{aligned}
U_{\max} = & \frac{bcE}{a(1+\nu)} \int \int \int_V \left\{ \frac{1-\nu}{2(1-2\nu)} \left[\left(\frac{\partial U}{\partial \bar{x}} \right)^2 + \left(\frac{a}{b} \right)^2 \left(\frac{\partial V}{\partial \bar{y}} \right)^2 + \left(\frac{a}{c} \right)^2 \left(\frac{\partial W}{\partial \bar{z}} \right)^2 \right] + \frac{\nu}{1-2\nu} \left[\frac{a}{b} \frac{\partial U}{\partial \bar{x}} \frac{\partial V}{\partial \bar{y}} + \frac{a^2}{bc} \frac{\partial V}{\partial \bar{y}} \frac{\partial W}{\partial \bar{z}} + \frac{a}{c} \frac{\partial U}{\partial \bar{x}} \frac{\partial W}{\partial \bar{z}} \right] \right. \\
& + \frac{1}{4} \left[\left(\frac{a}{c} \right)^2 \left(\frac{\partial V}{\partial \bar{z}} \right)^2 + \frac{2a^2}{bc} \frac{\partial V}{\partial \bar{z}} \frac{\partial W}{\partial \bar{y}} + \left(\frac{a}{b} \right)^2 \left(\frac{\partial W}{\partial \bar{y}} \right)^2 + \left(\frac{a}{c} \right)^2 \left(\frac{\partial U}{\partial \bar{z}} \right)^2 + \frac{2a}{c} \frac{\partial U}{\partial \bar{z}} \frac{\partial W}{\partial \bar{x}} + \left(\frac{\partial W}{\partial \bar{x}} \right)^2 + \left(\frac{a}{b} \right)^2 \left(\frac{\partial U}{\partial \bar{y}} \right)^2 \right. \\
& \left. \left. + \frac{2a}{b} \frac{\partial U}{\partial \bar{y}} \frac{\partial V}{\partial \bar{x}} + \left(\frac{\partial V}{\partial \bar{x}} \right)^2 \right] \right\} d\bar{x} d\bar{y} d\bar{z}, \tag{7}
\end{aligned}$$

and

$$T_{\max} = \frac{\rho \omega^2 abc}{2} \int \int \int_V (U^2 + V^2 + W^2) d\bar{x} d\bar{y} d\bar{z}. \tag{8}$$

C. Three-dimensional energy functional and eigenvalue equation

The displacement amplitude functions for a vibrating parallelepiped can be expressed by a set of three-dimensional (3D) p -Ritz functions. These functions are the products of 2D p -Ritz functions $\phi_u(\bar{x}, \bar{y})$, $\phi_v(\bar{x}, \bar{y})$, $\phi_w(\bar{x}, \bar{y})$ for the midsurface deformation, and 1D p -Ritz functions $\psi_u(\bar{z})$, $\psi_v(\bar{z})$, $\psi_w(\bar{z})$ for the thickness deformation. The displacement amplitude functions are

$$U(\bar{x}, \bar{y}, \bar{z}) = \sum_{i=1}^m \sum_{j=1}^n C_u^{ij} \phi_u^i(\bar{x}, \bar{y}) \psi_u^j(\bar{z}), \tag{9a}$$

$$V(\bar{x}, \bar{y}, \bar{z}) = \sum_{i=1}^m \sum_{j=1}^n C_v^{ij} \phi_v^i(\bar{x}, \bar{y}) \psi_v^j(\bar{z}), \tag{9b}$$

$$W(\bar{x}, \bar{y}, \bar{z}) = \sum_{i=1}^m \sum_{j=1}^n C_w^{ij} \phi_w^i(\bar{x}, \bar{y}) \psi_w^j(\bar{z}), \tag{9c}$$

in which C_u^{ij} , C_v^{ij} , and C_w^{ij} are the unknown coefficients.

An energy functional is defined as the difference of the maximum strain and kinetic energy components

$$\Pi = U_{\max} - T_{\max}. \tag{10}$$

Numerical frequency solutions can be obtained by minimizing this energy functional with respect to the unknown coefficients in accordance with the Ritz procedure

$$\frac{\partial \Pi}{\partial C_\alpha^{ij}} = 0; \quad \alpha = u, v, \text{ and } w, \tag{11}$$

which leads to the governing eigenvalue equation

$$(K - \lambda^2 M)\{C\} = \{0\}, \tag{12}$$

where

$$\lambda = \omega a \sqrt{\frac{\rho(1+\nu)}{E}} \tag{13}$$

is the dimensionless frequency parameter. The derivatives of U_{\max} and T_{\max} with respect to the unknown coefficients are presented in Appendix A.

The stiffness and mass matrices are

$$K = \begin{bmatrix} \mathbf{k}_{uu} & \mathbf{k}_{uv} & \mathbf{k}_{uw} \\ & \mathbf{k}_{vv} & \mathbf{k}_{vw} \\ \text{sym} & & \mathbf{k}_{ww} \end{bmatrix}, \tag{14}$$

$$M = \begin{bmatrix} \mathbf{m}_{uu} & [0] & [0] \\ & \mathbf{m}_{vv} & [0] \\ \text{sym} & & \mathbf{m}_{ww} \end{bmatrix}, \tag{15}$$

and the vector of unknown coefficients is

$$C = \begin{Bmatrix} \{C_u\} \\ \{C_v\} \\ \{C_w\} \end{Bmatrix}. \tag{16}$$

The elements in the stiffness submatrix are

$$\begin{aligned}
k_{uu}^{ikjl} = & \frac{1-\nu}{1-2\nu} I_{\phi_{uu}^{ik}}^{1010} J_{\psi_{uu}^{jl}}^{00} + \frac{1}{2} \left[\left(\frac{a}{c} \right)^2 I_{\phi_{uu}^{ik}}^{0000} J_{\psi_{uu}^{jl}}^{11} \right. \\
& \left. + \left(\frac{a}{b} \right)^2 I_{\phi_{uu}^{ik}}^{0101} J_{\psi_{uu}^{jl}}^{00} \right], \tag{17a}
\end{aligned}$$

$$k_{uv}^{ikjl} = \frac{\nu}{1-2\nu} \frac{a}{b} I_{\phi_{uv}^{ik}}^{1001} J_{\psi_{uv}^{jl}}^{00} + \frac{a}{2b} I_{\phi_{uv}^{ik}}^{0110} J_{\psi_{uv}^{jl}}^{00}, \tag{17b}$$

$$k_{uw}^{ikjl} = \frac{\nu}{1-2\nu} \frac{a}{c} I_{\phi_{uw}^{ik}}^{1000} J_{\psi_{uw}^{jl}}^{01} + \frac{a}{2c} I_{\phi_{uw}^{ik}}^{0010} J_{\psi_{uw}^{jl}}^{10}, \tag{17c}$$

$$\begin{aligned}
k_{vv}^{ikjl} = & \frac{1-\nu}{1-2\nu} \left(\frac{a}{b} \right)^2 I_{\phi_{vv}^{ik}}^{0101} J_{\psi_{vv}^{jl}}^{00} + \frac{1}{2} \left[\left(\frac{a}{c} \right)^2 I_{\phi_{vv}^{ik}}^{0000} J_{\psi_{vv}^{jl}}^{11} \right. \\
& \left. + I_{\phi_{vv}^{ik}}^{1010} J_{\psi_{vv}^{jl}}^{00} \right], \tag{17d}
\end{aligned}$$

$$k_{vw}^{ikjl} = \frac{\nu}{1-2\nu} \frac{a^2}{bc} I_{\phi_{vw}^{ik}}^{0100} J_{\psi_{vw}^{jl}}^{01} + \frac{a^2}{2bc} I_{\phi_{vw}^{ik}}^{0001} J_{\psi_{vw}^{jl}}^{10}, \tag{17e}$$

$$\begin{aligned}
k_{ww}^{ikjl} = & \frac{1-\nu}{1-2\nu} \left(\frac{a}{c} \right)^2 I_{\phi_{ww}^{ik}}^{0000} J_{\psi_{ww}^{jl}}^{11} + \frac{1}{2} \left[\left(\frac{a}{b} \right)^2 I_{\phi_{ww}^{ik}}^{0101} J_{\psi_{ww}^{jl}}^{00} \right. \\
& \left. + I_{\phi_{ww}^{ik}}^{1010} J_{\psi_{ww}^{jl}}^{00} \right], \tag{17f}
\end{aligned}$$

and the elements in the mass submatrix are

$$m_{uu}^{ikjl} = I_{\phi_{uu}^{ik}}^{0000} J_{\psi_{uu}^{jl}}^{00}, \tag{18a}$$

$$m_{vv}^{ikjl} = I_{\phi_{vv}^{ik}}^{0000} J_{\psi_{vv}^{jl}}^{00}, \tag{18b}$$

$$m_{ww}^{ijkl} = I_{\phi_{ww}^{ik}}^{0000} J_{\psi_{ww}^{jl}}^{00}, \quad (18c)$$

in which

$$I_{\phi_{\alpha\beta}^{abcd}} = \int \int_{\bar{A}} \frac{\partial^{a+b} \phi_{\alpha}^i(\bar{x}, \bar{y})}{\partial \bar{x}^a \partial \bar{y}^b} \frac{\partial^{c+d} \phi_{\beta}^k(\bar{x}, \bar{y})}{\partial \bar{x}^c \partial \bar{y}^d} d\bar{x} d\bar{y}, \quad (19a)$$

$$J_{\psi_{\alpha\beta}^{ef}} = \int_{\bar{c}} \frac{\partial^e \psi_{\alpha}^j(\bar{z})}{\partial \bar{z}^e} \frac{\partial^f \psi_{\beta}^l(\bar{z})}{\partial \bar{z}^f} d\bar{z}, \quad (19b)$$

where $\alpha, \beta = u, v, w$; $i, k = 1, 2, \dots, m$, and m is the total number of terms employed in the two-dimensional p -Ritz shape functions for planes parallel to the xy -midsurface; $j, l = 1, 2, \dots, n$, and n is the total number of terms employed in the one-dimensional p -Ritz shape functions in the thickness z -direction. The normalized midsurface area is denoted as \bar{A} and the normalized thickness is \bar{c} .

D. Boundary conditions and p -Ritz admissible functions

In the Ritz method, we ensure the satisfaction of geometric boundary conditions such as displacements and rotations at the boundary surfaces. Although satisfaction of natural boundary conditions such as shear forces and moments is not required at the outset, accurate computation shows that stresses at the free boundaries approach zero if accurate frequencies (for vibration) or buckling loads (for buckling) are obtained. For a parallelepiped, no geometric boundary condition is required for a free boundary surface. For a clamped boundary surface at $\bar{x} = -0.5$, the geometric boundary conditions are

$$U(\bar{x}, \bar{y}, \bar{z})|_{\bar{x}=-0.5} = V(\bar{x}, \bar{y}, \bar{z})|_{\bar{x}=-0.5} = W(\bar{x}, \bar{y}, \bar{z})|_{\bar{x}=-0.5} = 0. \quad (20)$$

The displacement components denoted by $U(\bar{x}, \bar{y}, \bar{z})$, $V(\bar{x}, \bar{y}, \bar{z})$, and $W(\bar{x}, \bar{y}, \bar{z})$ are truncated finite series expressed in Eqs. (5a)–(5c). The two-dimensional deformation admissible functions $\phi_u(\bar{x}, \bar{y})$, $\phi_v(\bar{x}, \bar{y})$, and $\phi_w(\bar{x}, \bar{y})$ are geometrically compliant polynomial functions derived such that the geometric boundary conditions (20) are satisfied at the outset (Lim *et al.*, 1998a). They are composed of the product of a series of simple two-dimensional polynomials $F_{\phi}(\bar{x}, \bar{y})$ and boundary-compliant basic functions $\phi_u^b(\bar{x}, \bar{y})$, $\phi_v^b(\bar{x}, \bar{y})$, and $\phi_w^b(\bar{x}, \bar{y})$. The basic functions are geometric expressions of the parallelepiped boundaries raised to an appropriate basic power in accordance with various boundary constraints (Lim *et al.*, 1998a). For a cantilevered parallelepiped clamped at $\bar{x} = -0.5$, the two-dimensional deformation admissible functions are

$$\phi_u(\bar{x}, \bar{y}) = (\bar{x} + 0.5)F_{\phi}(\bar{x}, \bar{y}), \quad (21a)$$

$$\phi_v(\bar{x}, \bar{y}) = (\bar{x} + 0.5)F_{\phi}(\bar{x}, \bar{y}), \quad (21b)$$

$$\phi_w(\bar{x}, \bar{y}) = (\bar{x} + 0.5)F_{\phi}(\bar{x}, \bar{y}), \quad (21c)$$

where

$$F_{\phi}(\bar{x}, \bar{y}) = \sum_{q=0}^{p_{xy}} \sum_{i=0}^q \bar{x}^q \bar{y}^i, \quad (22)$$

in which p_{xy} is the highest degree of the two-dimensional polynomial in the admissible functions and $m = (p_{xy} + 1) \times (p_{xy} + 2)/2$ is the number of the 2D terms.

Similarly, the one-dimensional thickness admissible functions $\psi_u(\bar{z})$, $\psi_v(\bar{z})$, and $\psi_w(\bar{z})$ are the products of sets of one-dimensional polynomial functions $F_{\psi}(\bar{z})$ and appropriate basic functions $\psi_u^b(\bar{z})$, $\psi_v^b(\bar{z})$, and $\psi_w^b(\bar{z})$. For a cantilevered parallelepiped clamped at $\bar{x} = -0.5$, the one-dimensional thickness deformation admissible functions are

$$\psi_u(\bar{z}) = F_{\psi}(\bar{z}), \quad (23a)$$

$$\psi_v(\bar{z}) = F_{\psi}(\bar{z}), \quad (23b)$$

$$\psi_w(\bar{z}) = F_{\psi}(\bar{z}), \quad (23c)$$

where

$$F_{\psi}(\bar{z}) = \sum_{i=0}^{p_z} \bar{z}^i, \quad (24)$$

in which p_z is the highest degree of the one-dimensional polynomial in the admissible functions and $n = p_z + 1$ indicates the number of 1D terms.

II. RESULTS AND DISCUSSION

A. Mode classification and convergence of eigenvalues

To improve computational efficiency without sacrificing numerical accuracy, classification of vibration modes is possible by grouping terms with odd and even powers of \bar{x} , \bar{y} and \bar{z} in $F_{\phi}(\bar{x}, \bar{y})$, and $F_{\psi}(\bar{z})$ in Eqs. (22) and (24), respectively. This tremendously reduces the number of terms in each series and thus the determinant size of the eigenvalue equation is considerably smaller. Huge computational effort can be saved without affecting the numerical accuracy because the odd and even terms of \bar{x} , \bar{y} , and \bar{z} only contribute to specific modes and they are trivial in other modes (Lim *et al.*, 1998a).

Classification of vibration modes depends on symmetry of geometry and boundary conditions. Eight mode classes are possible for a parallelepiped with perfect symmetry in geometry and boundary conditions. For a cantilevered parallelepiped, mode classification can be referred to the xy - and xz -planes (see Fig. 1) perpendicular to the clamped surface at $\bar{x} = -0.5$. Four mode classes exist as the symmetric–symmetric (SS), symmetric–antisymmetric (SA), antisymmetric–symmetric (AS), and antisymmetric–antisymmetric (AA) modes.

Convergence of eigenvalues presented in Table I for a cantilevered cube is investigated with respect to increases in p_z and p_{xy} for the 1D and 2D polynomial functions, respectively, in accordance with various symmetry classes. The polynomial degrees are increased from 7 to 10 for p_{xy} and from 7 to 9 for p_z . As observed in Table I, downward convergence is obvious. It is a unique numerical feature of the

TABLE I. Convergence of $\lambda = \omega a \sqrt{\rho/E}$ for a cantilevered cube with $\nu = 0.3$.

Symmetry Class	$p_{xy} \times p_z$	Mode number				
		M-1	M-2	M-3	M-4	
SS	7×7	1.5977	2.5805	2.9126	3.0508	
	8×7	1.5972	2.5805	2.9123	3.0508	
	9×7	1.5969	2.5804	2.9120	3.0507	
	10×7	1.5967	2.5803	2.9119	3.0507	
	10×8	1.5965	2.5803	2.9118	3.0507	
	10×9	1.5965	2.5803	2.9118	3.0507	
	SA	7×7	0.668 51	1.7671	2.7523	3.0578
		8×7	0.668 34	1.7670	2.7523	3.0575
		9×7	0.668 24	1.7669	2.7522	3.0573
10×7		0.668 16	1.7669	2.7522	3.0573	
10×8		0.667 96	1.7668	2.7521	3.0567	
10×9		0.667 96	1.7668	2.7521	3.0567	
AS		7×7	0.671 02	1.7696	2.7529	3.0670
		8×7	0.669 93	1.7679	2.7525	3.0629
		9×7	0.669 27	1.7675	2.7523	3.0597
	10×7	0.668 85	1.7672	2.7523	3.0589	
	10×8	0.668 85	1.7672	2.7523	3.0589	
	10×9	0.668 81	1.7672	2.7522	3.0589	
AA	7×7	0.909 09	2.1790	2.6909	2.7474	
	8×7	0.908 53	2.1788	2.6879	2.7467	
	9×7	0.908 32	2.1786	2.6870	2.7465	
	10×7	0.908 21	2.1786	2.6865	2.7464	
	10×8	0.908 21	2.1786	2.6865	2.7464	
	10×9	0.908 15	2.1786	2.6863	2.7463	

Ritz procedure which overestimates vibration frequency and buckling load and underestimates bending deflection. Convergence of eigenvalues can be ensured by including an adequate number of terms in the admissible shape functions. As seen in Table I, the eigenvalues are converged to at least three significant figures and in most cases more than three significant figures.

B. Comparison of solutions

As described in the Introduction, the Kirchhoff–Love thin-plate theory and most of the existing thick-plate models neglect transverse normal stress σ_{zz} . Although the omission of σ_{zz} for thin-plate theory has been examined and verified in many publications (Leissa, 1969), it is rather hypothetical in Reissner–Mindlin and Levinson–Reddy thick-plate models and has been questioned by a number of researchers such as Gould (1988). These thick-plate models further assume constant transverse normal displacement (w) through the thickness, or transverse inextensibility. It is easily verified that a trivial σ_{zz} implies

$$\epsilon_{zz} = -\frac{\nu}{1-\nu}(\epsilon_{xx} + \epsilon_{yy}), \quad (25)$$

in accordance with the generalized Hooke’s law. Therefore, to neglect σ_{zz} and at the same time keep w constant through the thickness (transverse inextensibility) is an inconsistent thick-plate model. Although some early higher-order thick-plate models (Whitney and Sun, 1973, 1974; Iyengar *et al.*, 1974 and Lo *et al.*, 1977a,b) had suggested nonlinear (quadratic and cubic) expressions for w , these models had not gained much popularity over the Reissner–Mindlin and Levinson–Reddy thick-plate models. It is, therefore, the pri-

mary objective of this paper to address the importance of σ_{zz} and the consistency of a thick-plate model with nonlinear w .

To simulate a thick-plate model neglecting σ_{zz} , Eq. (25) is substituted into Eq. (1) to derive the strain energy as

$$U = \frac{1}{2} \int \int \int_V \left[\frac{E}{1-\nu^2} (\epsilon_{xx}^2 + \epsilon_{yy}^2) + \frac{2\nu E}{1-\nu^2} \epsilon_{xx} \epsilon_{yy} + G(\gamma_{yz}^2 + \gamma_{zx}^2 + \gamma_{xy}^2) \right] dx dy dz, \quad (26)$$

while the kinetic energy expression (4) is still valid. The maximum strain energy is presented in Appendix B. The formulation from Eqs. (5a)–(16) is repeated to obtain a corresponding eigenvalue equation. The stiffness matrix element expressions for a parallelepiped neglecting σ_{zz} are presented in Appendix C.

For the purpose of comparison and to simulate constant w through thickness in Reissner–Mindlin and Levinson–Reddy thick-plate models, $\psi_w(\bar{z}) = F_{\psi}(\bar{z}) = 1$ is set in Eq. (23c) for solutions with $\sigma_{zz} \approx 0$ such that the transverse normal displacement

$$W(\bar{x}, \bar{y}, \bar{z}) = W(\bar{x}, \bar{y}) = \sum_{i=1}^m C_w^{i1} \phi_w^i(\bar{x}, \bar{y}) \quad (27)$$

in Eq. (9c) is independent of \bar{z} .

A comparison of free vibration frequency solutions with the results of Leissa and Zhang (1983) (denoted by A) is presented in Table II for parallelepipeds with various aspect ratios. The vibration frequencies are classified into four symmetry classes with respect to xy - and xz -planes. Solutions (denoted by B) by solving the full three-dimensional energy functional using the Ritz energy approach as governed by Eq. (12) are presented. In addition, solutions (denoted by C) neglecting σ_{zz} and assuming constant w through thickness with reference to Eqs. (26) and (27) are also included. Overall, excellent agreement of solutions between A and B is observed, while the agreement with respect to C is generally satisfactory. The excellent agreement of A and B solutions is expected, as Leissa and Zhang (1983) also solved the full three-dimensional energy functional. The most obvious discrepancy between solutions of A and C or B and C happens in the SS modes for a cantilevered parallelepiped with $a/b = 0.5$ and $b/c = 1$. One of the reasons the largest discrepancy is observed in this case is because this parallelepiped has a thickness twice the length ($c/a = 2$) and the effect of σ_{zz} is expected to be more obvious as the thickness increases. The other parallelepiped configurations have c/a either 0.5 or 1. We will see the effects of σ_{zz} more closely in the next section.

C. Effects of transverse normal stress σ_{zz}

A set of first-known solutions for examining the effects of σ_{zz} for various parallelepiped configurations is presented. Figures 2–5 present the nondimensional vibration frequency parameters λ with varying thickness ratio c/b in four distinct mode classes for a cantilevered rectangular parallelepiped with aspect ratio $a/b = 0.5$.

TABLE II. Comparison of $\lambda = \omega a \sqrt{\rho/E}$ for cantilevered rectangular parallelepiped ($\nu=0.3$) with xy - and xz -symmetric planes.

Aspect ratios			Mode frequencies											
a/b	b/c	Sources	SS-1	SS-2	SS-3	SA-1	SA-2	SA-3	AS-1	AS-2	AS-3	AA-1	AA-2	AA-3
1	1	A	1.6000	2.5812	2.9154	0.670 87	1.7695	2.7562	0.670 87	1.7695	2.7562	0.909 30	2.1801	2.7011
		B	1.5965	2.5803	2.9118	0.667 96	1.7668	2.7521	0.668 81	1.7672	2.7522	0.908 15	2.1786	2.6863
		C	1.5800	2.8160	3.2228	0.662 26	1.7595	2.7518	0.658 93	1.7726	3.0386	0.909 42	2.1982	2.7132
2	1	A	1.5938	4.5811	5.0646	0.443 71	1.6711	3.7237	0.443 71	1.6711	3.7237	0.904 02	2.7191	4.1787
		B	1.5888	4.5496	5.0520	0.440 21	1.6612	3.4626	0.440 72	1.6629	3.4654	0.902 08	2.7012	4.1711
		C	1.5778	4.5898	5.2746	0.436 72	1.6516	3.4475	0.435 04	1.6592	3.4757	0.902 54	2.7082	4.1953
1	2	A	1.5962	2.7974	3.1994	0.447 33	1.6642	2.2777	0.667 44	1.7744	3.0680	0.788 31	2.2196	3.4387
		B	1.5920	2.7959	3.1946	0.444 13	1.6551	2.2733	0.664 96	1.7721	3.0436	0.785 86	2.2094	3.4007
		C	1.5800	2.8160	3.2228	0.440 30	1.6455	2.2634	0.658 93	1.7726	3.0386	0.784 53	2.2063	3.3960
0.5	1	A	1.4670	1.5623	1.7967	0.830 04	1.5317	1.7647	0.830 04	1.5317	1.7647	0.916 36	1.3550	1.9061
		B	1.4679	1.5588	1.8423	0.827 32	1.5293	1.7628	0.828 31	1.5289	1.7629	0.914 85	1.3530	1.9052
		C	1.5281	1.6889	2.0438	0.826 18	1.5281	1.7523	0.818 63	1.6839	2.3306	0.918 70	1.3646	1.8903
0.5	2	A	1.5325	1.6835	2.0337	0.674 84	1.3538	1.8070	0.827 12	1.6911	2.3128	0.821 31	1.7307	2.1420
		B	1.5300	1.6812	2.0282	0.672 00	1.3507	1.8037	0.825 31	1.6892	2.3020	0.819 37	1.7284	2.1363
		C	1.5208	1.6889	2.0438	0.665 72	1.3496	1.7933	0.818 63	1.6839	2.3306	0.818 29	1.7263	2.1193

^aLeissa and Zhang (1983).

^bPresent (including σ_{zz}).

^cPresent (excluding σ_{zz} and assuming constant w through thickness).

Frequency solutions for the first four SS modes with c/b ranging from 0.2 to 2.0 are presented in Fig. 2. As expected, excellent agreement for solutions including and excluding σ_{zz} (constant w through thickness) is achieved for small thickness ratio. The discrepancy of solutions becomes more apparent as c/b increases. The onsets of distinct discrepancy between the solutions are associated with smaller c/b for higher vibration modes. For instance, distinct discrepancy of SS-1 solutions can be observed for $c/b > 0.9$, while for SS-4 the onset reduces to $c/b \approx 0.4$. The physical implication is as follows. It is widely known that higher vibration modes pos-

sess more nodal lines (Lim *et al.*, 1998b) for plates; or more nodal surfaces for parallelepipeds. A nodal line (surface) is a line (surface) with zero vibration amplitude. We may treat a nodal line (surface) as a boundary line (surface) with certain constraints which should not be seen as a simply supported or clamped boundary constraints. As a result, the effective region is smaller for higher vibration modes with more nodal lines (surfaces) or, equivalently, the effective thickness ratio becomes larger than the overall thickness ratio c/b . Consequently, the onset of distinct discrepancy between solutions

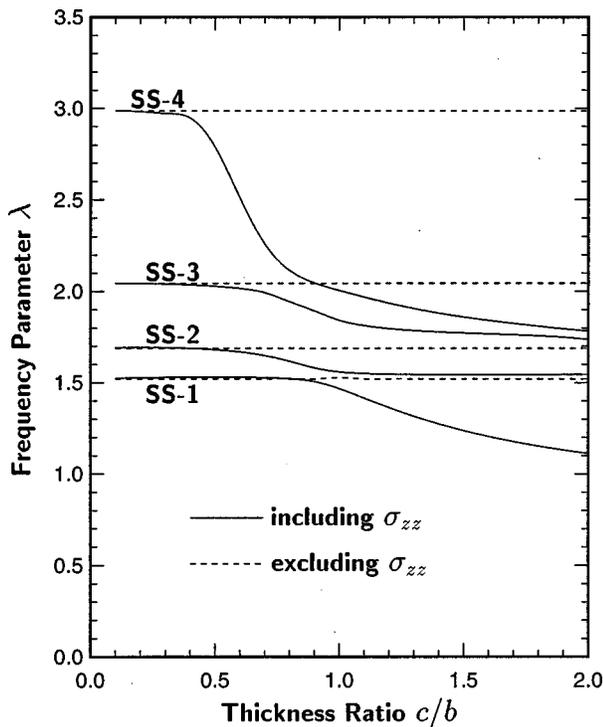


FIG. 2. Effect of thickness ratio on the SS frequencies for a thick cantilevered rectangular parallelepiped with $\nu=0.3$ and $a/b=0.5$.

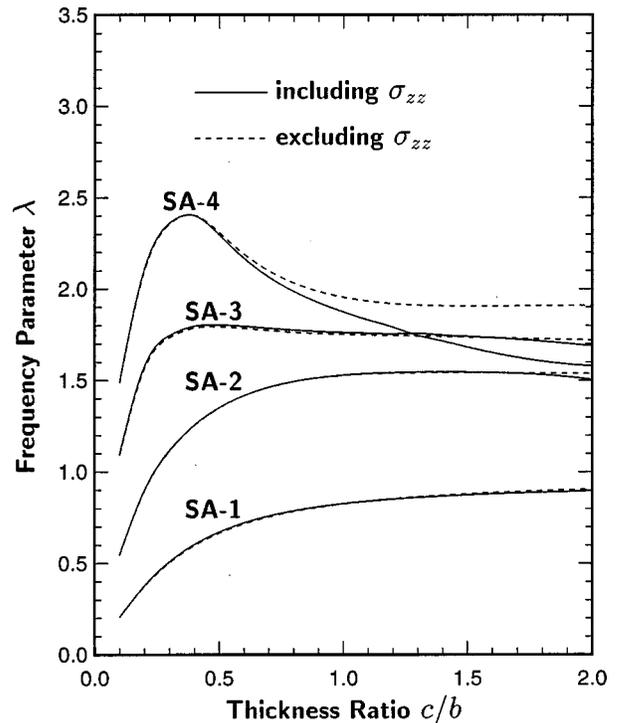


FIG. 3. Effect of thickness ratio on the SA frequencies for a thick cantilevered rectangular parallelepiped with $\nu=0.3$ and $a/b=0.5$.

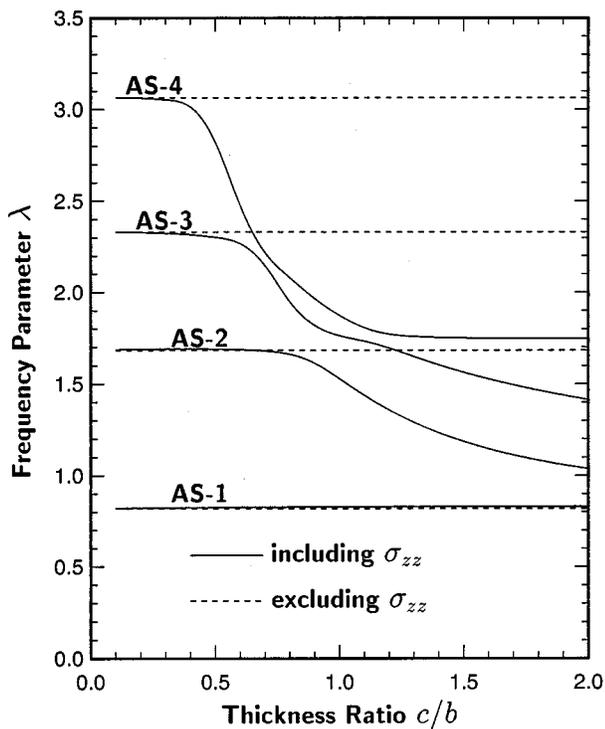


FIG. 4. Effect of thickness ratio on the AS frequencies for a thick cantilevered rectangular parallelepiped with $\nu=0.3$ and $a/b=0.5$.

with respect to c/b becomes smaller for higher vibration modes. The effects of σ_{zz} on other classes of vibration modes are illustrated in Figs. 3–5.

It is also noticed that solutions considering σ_{zz} are always lower when the effect of σ_{zz} becomes significant for large c/b . Although $\epsilon_{zz} \neq 0$, as determined by Eq. (25), even if σ_{zz} is neglected, the consideration of σ_{zz} allows transverse extensibility not governed by Eq. (25), thus providing a further degree of flexibility in the transverse normal direction. Vibration frequency is smaller if σ_{zz} is considered because flexibility reduces the structural stiffness for the parallelepiped.

In Figs. 2–5, more apparent discrepancy of solutions for the SS modes compared to the other modes is observed. This is true only for a cantilevered parallelepiped and it is inconclusive for other parallelepiped configurations as the significance and effects of σ_{zz} depend not only on geometry but also boundary conditions. However, it is reasonable to conclude that the effect of σ_{zz} for $c/b < 0.5$ is rather insignificant for some lower modes in the aspect of free vibration of thick plates and parallelepipeds. Such conclusion was also reached by Hutchinson (1979, 1984) in analyzing vibrations of thick, free circular plates using exact and approximate methods. These analyses concluded that, for thick, free circular plates, the approximate solution yields frequencies of sufficient accuracy for most engineering application within the range of applicability of the approximate theory. Based on the conclusions, accurate solutions employing the Levinson-Reddy higher-order plate theory have been reported (Liew and Lim, 1996; Lim and Liew, 1995, 1996; Lim *et al.*, 1998a,b).

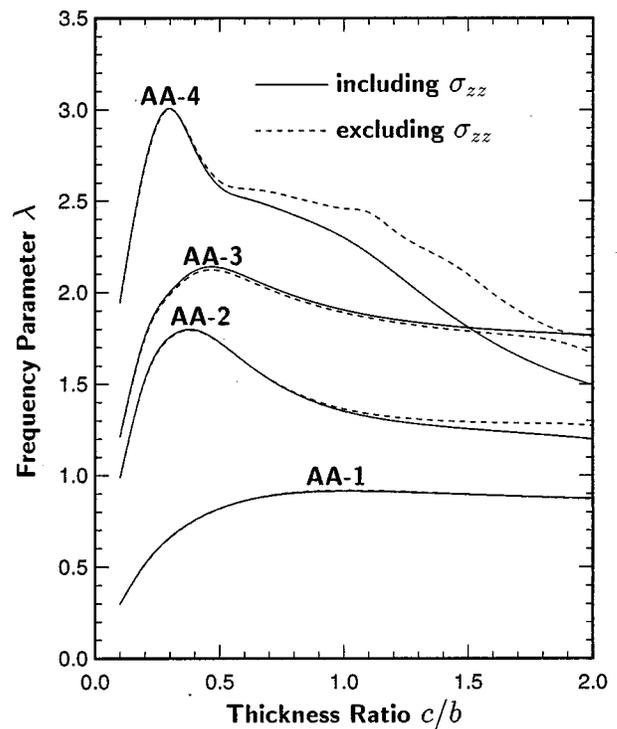


FIG. 5. Effect of thickness ratio on the AA frequencies for a thick cantilevered rectangular parallelepiped with $\nu=0.3$ and $a/b=0.5$.

III. CONCLUSIONS

New solutions and benchmarks for examining the effects of transverse normal stress σ_{zz} in the context of vibration of a cantilevered parallelepiped are presented and analyzed. The hypothetical assumption of neglecting σ_{zz} while maintaining constant transverse normal displacement w through thickness in existing thick-plate models is investigated by solving the full three-dimensional energy functional. The extent to which the assumption is applicable is verified.

When the effect of transverse normal stress becomes significant for large thickness ratio, vibration frequency is always lower when transverse normal stress is considered because transverse extensibility provides a further degree of flexibility in the transverse normal direction and thus reduces the stiffness of a parallelepiped. For higher vibration modes, the onset of significant contribution of transverse normal stress is associated with a lower thickness ratio.

ACKNOWLEDGMENT

A research fellowship for the author from The University of Hong Kong is gratefully acknowledged.

APPENDIX A

The derivatives of strain and kinetic energy integrals with respect to the unknown coefficients for Eq. (11) are as follows:

$$\begin{aligned} \frac{\partial U_{\max}}{\partial C_u^{ij}} = & \frac{bcE}{a(1+\nu)} \sum_{k=1}^m \sum_{l=1}^n \left\{ \frac{1-\nu}{1-2\nu} C_u^{kl} I_{\phi_{uu}^{ik}}^{1010} J_{\psi_{uu}^{jl}}^{00} \right. \\ & + \frac{\nu}{1-2\nu} \left[\frac{a}{b} C_v^{kl} I_{\phi_{uv}^{ik}}^{1001} J_{\psi_{uv}^{jl}}^{00} + \frac{a}{c} C_w^{kl} I_{\phi_{uw}^{ik}}^{1000} J_{\psi_{uw}^{jl}}^{01} \right] \\ & + \frac{1}{2} \left[\left(\frac{a}{c} \right)^2 C_u^{kl} I_{\phi_{uu}^{ik}}^{0000} J_{\psi_{uu}^{jl}}^{11} + \frac{a}{c} C_w^{kl} I_{\phi_{uw}^{ik}}^{0010} J_{\psi_{uw}^{jl}}^{10} \right. \\ & \left. \left. + \left(\frac{a}{b} \right)^2 C_u^{kl} I_{\phi_{uu}^{ik}}^{0101} J_{\psi_{uu}^{jl}}^{00} + \frac{a}{b} C_v^{kl} I_{\phi_{uv}^{ik}}^{0110} J_{\psi_{uv}^{jl}}^{00} \right] \right\}, \quad (\text{A1a}) \end{aligned}$$

$$\begin{aligned} \frac{\partial U_{\max}}{\partial C_v^{ij}} = & \frac{bcE}{a(1+\nu)} \sum_{k=1}^m \sum_{l=1}^n \left\{ \frac{1-\nu}{1-2\nu} \left(\frac{a}{b} \right)^2 C_v^{kl} I_{\phi_{vv}^{ik}}^{0101} J_{\psi_{vv}^{jl}}^{00} \right. \\ & + \frac{\nu}{1-2\nu} \left[\frac{a}{b} C_u^{kl} I_{\phi_{vu}^{ik}}^{0110} J_{\psi_{vu}^{jl}}^{00} + \frac{a^2}{bc} C_w^{kl} I_{\phi_{vw}^{ik}}^{0100} J_{\psi_{vw}^{jl}}^{01} \right] \\ & + \frac{1}{2} \left[\left(\frac{a}{c} \right)^2 C_v^{kl} I_{\phi_{vv}^{ik}}^{0000} J_{\psi_{vv}^{jl}}^{11} + \frac{a^2}{bc} C_w^{kl} I_{\phi_{vw}^{ik}}^{0001} J_{\psi_{vw}^{jl}}^{10} \right. \\ & \left. \left. + \frac{a}{b} C_u^{kl} I_{\phi_{vu}^{ik}}^{1001} J_{\psi_{vu}^{jl}}^{00} + C_v^{kl} I_{\phi_{vv}^{ik}}^{1010} J_{\psi_{vv}^{jl}}^{00} \right] \right\}, \quad (\text{A1b}) \end{aligned}$$

$$\begin{aligned} \frac{\partial U_{\max}}{\partial C_w^{ij}} = & \frac{bcE}{a(1+\nu)} \sum_{k=1}^m \sum_{l=1}^n \left\{ \frac{1-\nu}{1-2\nu} \left(\frac{a}{c} \right)^2 C_w^{kl} I_{\phi_{ww}^{ik}}^{0000} J_{\psi_{ww}^{jl}}^{11} \right. \\ & + \frac{\nu}{1-2\nu} \left[\frac{a^2}{bc} C_v^{kl} I_{\phi_{vw}^{ik}}^{0001} J_{\psi_{vw}^{jl}}^{10} + \frac{a}{c} C_u^{kl} I_{\phi_{wu}^{ik}}^{0010} J_{\psi_{wu}^{jl}}^{10} \right] \\ & + \frac{1}{2} \left[\frac{a^2}{bc} C_v^{kl} I_{\phi_{vw}^{ik}}^{0100} J_{\psi_{vw}^{jl}}^{01} + \left(\frac{a}{b} \right)^2 C_w^{kl} I_{\phi_{ww}^{ik}}^{0101} J_{\psi_{ww}^{jl}}^{00} \right. \\ & \left. \left. + \frac{a}{c} C_u^{kl} I_{\phi_{wu}^{ik}}^{1000} J_{\psi_{wu}^{jl}}^{01} + C_w^{kl} I_{\phi_{ww}^{ik}}^{1010} J_{\psi_{ww}^{jl}}^{00} \right] \right\}, \quad (\text{A1c}) \end{aligned}$$

$$\frac{\partial T_{\max}}{\partial C_u^{ij}} = \rho \omega^2 abc \sum_{k=1}^m \sum_{l=1}^n C_u^{kl} I_{\phi_{uu}^{ik}}^{0000} J_{\psi_{uu}^{jl}}^{00}, \quad (\text{A2a})$$

$$\frac{\partial T_{\max}}{\partial C_v^{ij}} = \rho \omega^2 abc \sum_{k=1}^m \sum_{l=1}^n C_v^{kl} I_{\phi_{vv}^{ik}}^{0000} J_{\psi_{vv}^{jl}}^{00}, \quad (\text{A2b})$$

$$\frac{\partial T_{\max}}{\partial C_w^{ij}} = \rho \omega^2 abc \sum_{k=1}^m \sum_{l=1}^n C_w^{kl} I_{\phi_{ww}^{ik}}^{0000} J_{\psi_{ww}^{jl}}^{00}, \quad (\text{A2c})$$

where the integral notations $I_{\phi_{\alpha\beta}^{abcd}}$ and $J_{\psi_{\alpha\beta}^{ef}}$ are given in Eqs. (19a), (19b).

APPENDIX B

The maximum strain energy in a vibration cycle for a parallelepiped neglecting σ_{zz} is as follows:

$$\begin{aligned} U_{\max} = & \frac{bcE}{a(1+\nu)} \int \int \int_V \left\{ \frac{1}{2(1-\nu)} \left[\left(\frac{\partial U}{\partial \bar{x}} \right)^2 \right. \right. \\ & + \left(\frac{a}{b} \right)^2 \left(\frac{\partial V}{\partial \bar{y}} \right)^2 \left. \left. + \frac{a\nu}{b(1-\nu)} \frac{\partial U}{\partial \bar{x}} \frac{\partial V}{\partial \bar{y}} \right] \right. \\ & + \frac{1}{4} \left[\left(\frac{a}{c} \right)^2 \left(\frac{\partial V}{\partial \bar{z}} \right)^2 + \frac{2a^2}{bc} \frac{\partial V}{\partial \bar{z}} \frac{\partial W}{\partial \bar{y}} + \left(\frac{a}{b} \right)^2 \left(\frac{\partial W}{\partial \bar{y}} \right)^2 \right. \\ & + \left(\frac{a}{c} \right)^2 \left(\frac{\partial U}{\partial \bar{z}} \right)^2 + \frac{2a}{c} \frac{\partial U}{\partial \bar{z}} \frac{\partial W}{\partial \bar{x}} + \left(\frac{\partial W}{\partial \bar{x}} \right)^2 + \left(\frac{a}{b} \right)^2 \left(\frac{\partial U}{\partial \bar{y}} \right)^2 \\ & \left. \left. + \frac{2a}{b} \frac{\partial U}{\partial \bar{y}} \frac{\partial V}{\partial \bar{x}} + \left(\frac{\partial V}{\partial \bar{x}} \right)^2 \right] \right\} d\bar{x} d\bar{y} d\bar{z}. \quad (\text{B1}) \end{aligned}$$

APPENDIX C

The stiffness matrix element expressions for a parallel-piped neglecting σ_{zz} are as follows:

$$k_{uu}^{ijkl} = \frac{1}{1-\nu} I_{\phi_{uu}^{ik}}^{1010} J_{\psi_{uu}^{jl}}^{00} + \frac{1}{2} \left[\left(\frac{a}{c} \right)^2 I_{\phi_{uu}^{ik}}^{0000} J_{\psi_{uu}^{jl}}^{11} + \left(\frac{a}{b} \right)^2 I_{\phi_{uu}^{ik}}^{0101} J_{\psi_{uu}^{jl}}^{00} \right], \quad (\text{C1a})$$

$$k_{uv}^{ijkl} = \frac{\nu}{1-\nu} \frac{a}{b} I_{\phi_{uv}^{ik}}^{1001} J_{\psi_{uv}^{jl}}^{00} + \frac{a}{2b} I_{\phi_{uv}^{ik}}^{0110} J_{\psi_{uv}^{jl}}^{00}, \quad (\text{C1b})$$

$$k_{uw}^{ijkl} = \frac{a}{2c} I_{\phi_{uw}^{ik}}^{0010} J_{\psi_{uw}^{jl}}^{10}, \quad (\text{C1c})$$

$$k_{vv}^{ijkl} = \frac{1}{1-\nu} \left(\frac{a}{b} \right)^2 I_{\phi_{vv}^{ik}}^{0101} J_{\psi_{vv}^{jl}}^{00} + \frac{1}{2} \left[\left(\frac{a}{c} \right)^2 I_{\phi_{vv}^{ik}}^{0000} J_{\psi_{vv}^{jl}}^{11} + I_{\phi_{vv}^{ik}}^{1010} J_{\psi_{vv}^{jl}}^{00} \right], \quad (\text{C1d})$$

$$k_{vw}^{ijkl} = \frac{a^2}{2bc} I_{\phi_{vw}^{ik}}^{0001} J_{\psi_{vw}^{jl}}^{10}, \quad (\text{C1e})$$

$$k_{ww}^{ijkl} = \frac{1}{2} \left[\left(\frac{a}{b} \right)^2 I_{\phi_{ww}^{ik}}^{0101} J_{\psi_{ww}^{jl}}^{00} + I_{\phi_{ww}^{ik}}^{1010} J_{\psi_{ww}^{jl}}^{00} \right]. \quad (\text{C1f})$$

- Bert, C. W., and Chen, T. L. C. (1978). "Effect of shear deformation on vibration of antisymmetric angle-ply laminated rectangular plates," *Int. J. Solids Struct.* **14**, 465–473.
- Fromme, J. A., and Leissa, A. W. (1970). "Free vibration of the rectangular parallelepiped," *J. Acoust. Soc. Am.* **48**, 290–298.
- Gould, P. L. (1988). *Analysis of Shells and Plates* (Springer, New York). p. 507.
- Hutchinson, J. R. (1967). "Axisymmetric vibrations of a solid elastic cylinder encased in a rigid container," *J. Acoust. Soc. Am.* **42**, 398–402.
- Hutchinson, J. R. (1971). "Axisymmetric vibration of a free finite-length rod," *J. Acoust. Soc. Am.* **51**, 223–240.
- Hutchinson, J. R. (1979). "Axisymmetric flexural vibrations of a thick free circular plate," *Trans. ASME, J. Appl. Mech.* **46**, 139–144.
- Hutchinson, J. R. (1980). "Vibrations of solid cylinders," *Trans. ASME, J. Appl. Mech.* **47**, 901–907.
- Hutchinson, J. R. (1981). "Transverse vibration of beams: Exact versus approximate solutions," *Trans. ASME, J. Appl. Mech.* **48**, 923–928.
- Hutchinson, J. R. (1984). "Vibrations of thick free circular plates, exact versus approximate solutions," *Trans. ASME, J. Appl. Mech.* **51**, 581–585.
- Hutchinson, J. R., and Zillmer, S. D. (1983). "Vibration of a free rectangular parallelepiped," *Trans. ASME, J. Appl. Mech.* **50**, 123–130.
- Iyengar, K. T. S., Chandrashekhara, K., and Sebastian, V. K. (1974). "On the analysis of thick rectangular plates," *Arch. Appl. Mech.* **43**, 317–330.

- Leissa, A. W. (1969). *Vibration of Plates*, NASA SP-160.
- Leissa, A. W., and Zhang, Z. D. (1983). "On the three-dimensional vibrations of the cantilevered rectangular parallelepiped," *J. Acoust. Soc. Am.* **73**, 2013–2021.
- Leissa, A. W., and So, J. (1995a). "Comparisons of vibration frequencies for rods and beams from one-dimensional and three-dimensional analyses," *J. Acoust. Soc. Am.* **98**, 2122–2135.
- Leissa, A. W., and So, J. (1995b). "Accurate vibration frequencies of circular cylinders from three-dimensional analysis," *J. Acoust. Soc. Am.* **98**, 2136–2141.
- Leissa, A. W., and So, J. (1995c). "Three-dimensional vibrations of truncated hollow cones," *J. Vibration Control* **1**, 145–158.
- Levinson, M. (1980). "An accurate simple theory of the statics and dynamics of elastic plates," *Mech. Res. Commun.* **7**, 343–350.
- Liew, K. M., Hung, K. C., and Lim, M. K. (1995a). "Free vibration studies on stress-free three-dimensional elastic solids," *Trans. ASME, J. Appl. Mech.* **62**(1), 159–165.
- Liew, K. M., Hung, K. C., and Lim, M. K. (1995b). "Vibration of stress-free hollow cylinders of arbitrary cross section," *Trans. ASME, J. Appl. Mech.* **62**(3), 718–724.
- Liew, K. M., and Lim, C. W. (1996). "A higher-order theory for vibration of doubly-curved shallow shells," *Trans. ASME, J. Appl. Mech.* **63**(3), 587–593.
- Lim, C. W., and Liew, K. M. (1995). "A higher order theory for vibration of shear deformable cylindrical shallow shells," *Int. J. Mech. Sci.* **37**(3), 277–295.
- Lim, C. W., and Liew, K. M. (1996). "Vibration of moderately thick cylindrical shallow shells," *J. Acoust. Soc. Am.* **100**, 3665–3673.
- Lim, C. W., Liew, K. M., and Kitipornchai, S. (1998a). "Numerical aspects for free vibration of thick plates. Part I: Formulation and verification," *Comput. Methods Appl. Mech. Eng.* **156**(1–4), 15–29.
- Lim, C. W., Kitipornchai, S., and Liew, K. M. (1998b). "Numerical aspects for free vibration of thick plates. Part II: Numerical efficiency and vibration frequencies," *Comput. Methods Appl. Mech. Eng.* **156**(1–4), 31–44.
- Lim, C. W., Liew, K. M., and Kitipornchai, S. (1998c). "Vibration of open cylindrical shells: A three-dimensional elasticity approach," *J. Acoust. Soc. Am.* (in press).
- Lo, K. H., Christensen, R. M., and Wu, E. M. (1977a). "A higher-order theory of plate deformation, Part I: Homogeneous plates," *Trans. ASME, J. Appl. Mech.* **44**, 663–668.
- Lo, K. H., Christensen, R. M., and Wu, E. M. (1977b). "A higher-order theory of plate deformation, Part 2: Laminated plates," *Trans. ASME, J. Appl. Mech.* **44**, 669–676.
- Mindlin, R. D. (1951). "Influence of rotatory inertia and shear on flexural motions of isotropic, elastic plates," *Trans. ASME, J. Appl. Mech.* **18**, 31–38.
- Qatu, M. S. (1993). "Theories and analyses of thin and moderately thick laminated composite curved beams," *Int. J. Solids Struct.* **30**(20), 2743–2756.
- Reddy, J. N. (1984a). *Energy and Variational Methods in Applied Mechanics* (Wiley, New York).
- Reddy, J. N. (1984b). "A simple higher-order theory for laminated composite plates," *Trans. ASME, J. Appl. Mech.* **51**, 745–752.
- Reddy, J. N., and Khdeir, A. A. (1989). "Buckling and vibration of laminated composite plates using various plate theory," *Amer. Inst. Aero Astro J.* **27**, 1808–1817.
- Reddy, J. N., and Phan, N. D. (1985). "Stability and vibration of isotropic, orthotropic and laminated plates according to a higher-order shear deformation theory," *J. Sound Vib.* **98**, 157–170.
- Reissner, E. (1945). "The effect of transverse shear deformation on the bending of elastic plates," *Trans. ASME, J. Appl. Mech.* **12**, A69–A77.
- So, J., and Leissa, A. W. (1997). "Free vibrations of thick hollow circular cylinders from three-dimensional analysis," *Trans. ASME, J. Vib. Acoust.* **119**, 89–95.
- Soler, A. (1968). "Higher order effects in thick rectangular elastic beams," *Int. J. Solids Struct.* **4**, 723–739.
- Washizu, K. (1982). *Variational Methods in Elasticity and Plasticity*, 3rd ed. (Pergamon Oxford). pp. 276–279, 599–600.
- Whitney, J. M., and Sun, C. T. (1973). "A higher order theory for extensional motion of laminated composites," *J. Sound Vib.* **30**, 85–97.
- Whitney, J. M., and Sun, C. T. (1974). "A refined theory for laminated, anisotropic cylindrical shells," *Trans. ASME, J. Appl. Mech.* **41**(2), 471–476.

The sensitivity of structural acoustic response to attachment feature scale representation

Kenneth A. Cunefare^{a)}

G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405

Sergio De Rosa^{b)}

Dipartimento di Progettazione Aeronautica, University of Naples "Federico II," Via Claudio 21, 80125 Naples, Italy

(Received 10 January 1999; accepted for publication 9 August 1999)

This paper presents a technique to assess the impact on model fidelity introduced through discretization of attachments on harmonically forced fluid-loaded structural acoustic models. While fluid loading is included, it is not a requirement or restriction to the methods presented. The perspective taken is one of knowledge of a reference state, with a desire to determine the impact on the total radiated acoustic power due to perturbations in the reference state. Such perturbations change the predicted resonance frequencies of a structure under consideration and, hence, change the predicted response amplitudes. The method uses a single degree of freedom response model in the local region of each fluid-loaded resonance, coupled with eigenvalue sensitivities, to estimate the perturbation impact. The sensitivity of the eigenvalues to changes in model detail is derived based on variations in the spatial representation of attached features (e.g., point versus distributed attachments). Elements of the analysis method are not necessarily restricted to model perturbations nor acoustic power, rather they may be used to assess the perturbation of any quadratic response quantity of interest due to changes in resonance frequency. The analysis reveals that the bandwidth of response perturbation increases with increasing resonance frequency. For frequencies within $\pm 5\%$ of a resonance frequency, the amount of damping in the system determines and limits the magnitude of the response perturbation. The perturbation outside the range of $\pm 5\%$ of the resonance frequency is relatively insensitive to damping. The SDOF analysis method is limited by its assumption of constant modal forcing between the reference and perturbed states. © 1999 Acoustical Society of America. [S0001-4966(99)01112-1]

PACS numbers: 43.40.Dx, 43.40.Rj [CBB]

INTRODUCTION

The question of the degree of fidelity required for accurate structural acoustic modeling continues to be of significant interest. A related issue is to what degree will manufacturing variability cause differences in the in-service response of otherwise identical structures. Both of these issues, model fidelity and manufacturing variability, represent two different perspectives to the same problem, the impact of parameter variation on the response of a system. From the numerical perspective, the amount of detail incorporated in the modeling of structural features (e.g., use of lumped versus distributed representations) lead to different possible levels of model refinement, with associated differences in the predicted response. From the manufacturing perspective, build variability leads to differences between actual structural properties and modeled properties, leading to differences in the response.

In the work at hand, we seek to determine bounds or estimates on the response perturbation due to the sorts of variations described above. These bounds provide the means to determine if additional effort is required in the modeling,

and should serve as an aid to identify those regions of structures that strongly influence response variability due to build variability or modeling fidelity. Prior work of Shepard¹ and Shepard and Cunefare² has demonstrated that the sorts of response perturbations we seek to bound are most prevalent about resonance conditions. The response perturbations may be attributed to contributions due to the change in the structure, and in the resonance frequencies, i.e., as a given mode's natural frequency is perturbed closer or farther away from a fixed excitation frequency, the response amplitude of that mode increases or decreases accordingly.

Our emphasis here is related to the concept of mesh refinement and design sensitivity analyses commonly used in finite element method (FEM) or boundary element method (BEM) techniques. In mesh refinement, a metric of interest (e.g., strain energy) is monitored while increasingly fine discretizations of the model are analyzed. The model is assumed to be sufficiently refined once the metric is observed to have converged by some set of criteria. Here, we take the perspective that we have an adequately converged reference model, but wish to investigate the significance of perturbations about that reference model.

The most closely related work to the perturbation aspect of what we present here is that of Hahn and Ferri.³ Hahn and Ferri developed a sensitivity approach, and assumed a per-

^{a)}Electronic mail: ken.cunefare@me.gatech.edu

^{b)}Electronic mail: derosa@unina.it

turbation that is on the same order of magnitude as the underlying dynamical system. Here, we explicitly model small perturbations, and determine when these perturbations become significant.

Other closely related work is that of Cuschieri and Feit.^{4,5} In their work, the authors consider the influence that attachments of finite span with the same total impedance but with different local impedance distributions within the span have on the scattering of sound from an infinite fluid-loaded elastic plate. Of relevance to the work at hand, the authors conclude that the distribution details of the attachment do impact the scattered response. Cuschieri and Feit considered infinite fluid-loaded plates.

Prior work of Shepard¹ and Shepard and Cunefare² indicate that attachment feature variation has the most significant impact only within a narrow bandwidth about the system resonances. As compared to the work of Cuschieri and Feit,^{4,5} Cunefare and Shepard considered a finite one-dimensional plate in an infinite baffle. Prior work by Davies⁶ indicates that near a fluid-loaded resonance, the mode in resonance is principally responsible for the acoustic radiation. Finally, Giordano and Koopmann⁷ and Cunefare and De Rosa⁸ have developed state-space methods for the direct solution of fluid-loaded resonances. In light of this enabling prior art, in the work at hand we focus our attention strictly to the regions in the frequency domain near to fluid-loaded resonances. Further, since only one mode dominates the response at resonance, we will use a simple single-degree-of-freedom (SDOF) model to estimate the change in the response due to a perturbation in the structural model. We will demonstrate that this SDOF model, when combined with an estimate of the change in a mode's resonance frequency due to the structural perturbation, provides an accurate bound on the response perturbation.

In the following, we first develop the general structure for our analysis. We then use a baffled, fluid-loaded one-dimensional infinite plate as a specific example for the application of the analysis. The particular structure, and the method and technique of obtaining the system of equations, is not fundamental to the development. Rather, it is what we do once we have such a system that is the focus of this work. Our technique should be applicable to any dynamical system, not just baffled panels.

I. GENERAL DEVELOPMENT OF RESPONSE BOUNDS

We use the convention of bold capital letters as representing matrices, and bold lower case letters as representing vectors. Throughout this work, we assume a harmonic time dependency of $e^{-j\omega t}$. In the following, we first introduce the system forced response in terms of a reference and perturbed system dynamic stiffness. We then obtain the eigensolution for the reference fluid-loaded system, and estimate the eigensolution for the perturbed system. We then introduce our SDOF model for the response perturbation.

II. FLUID-STRUCTURE MODELING AND PERTURBATION

We presume that we have available a structural acoustic model represented as

$$\mathbf{Z}_{\text{sys}}\mathbf{w}=\mathbf{f}^e, \quad (1)$$

where the system dynamic stiffness matrix \mathbf{Z}_{sys} is of size $M \times M$, where M is the number of degrees of freedom, \mathbf{w} is a vector of displacements, and \mathbf{f}^e is a vector of imposed forces (including fluid-loading effects). We write the system dynamic stiffness matrix as being composed of contributions from the fluid loading and from the structure as

$$\mathbf{Z}_{\text{sys}}=\mathbf{Z}_{\text{fl}}+\mathbf{Z}_{\text{st}}, \quad (2)$$

where \mathbf{Z}_{fl} is the fluid-loading dynamic stiffness and \mathbf{Z}_{st} is the structural dynamic stiffness. The fluid-loading dynamic stiffness, \mathbf{Z}_{fl} , is only a function of the method used to model the fluid-structure coupling, and will not be presumed to be subject to variation (we do not include variations in the wetted surface geometry). Indeed, for non-fluid-loaded problems, this dynamic stiffness term may be set to zero, without impacting the subsequent development.

To continue, we presume that the structural dynamic stiffness may be represented as

$$\mathbf{Z}_{\text{st}}=\mathbf{Z}_0+\mathbf{Z}_{\delta}, \quad (3)$$

where \mathbf{Z}_0 defines a presumably known reference dynamic stiffness state of the structure, and \mathbf{Z}_{δ} is a perturbation about that state. We may further consider the reference state to include contributions from a base, or parent, structure and its attachments,

$$\mathbf{Z}_0=\mathbf{Z}_p+\mathbf{Z}_a, \quad (4)$$

where \mathbf{Z}_p is the dynamic stiffness of the parent structure, and \mathbf{Z}_a is the dynamic stiffness of the attachments. Now, as we wish to consider the impact of alternative representations for the attachments, we may write the perturbing dynamic stiffness as

$$\mathbf{Z}_{\delta}=\mathbf{Z}'_a-\mathbf{Z}_a, \quad (5)$$

where \mathbf{Z}'_a is an attachment dynamic stiffness obtained through an alternative representation. Note that unlike the approach of Hahn and Ferri,³ our perturbation to the system need not be of the same order of magnitude as the original dynamic stiffness.

III. DETERMINATION OF FLUID-LOADED SYSTEM RESONANCES

We cannot directly solve the homogeneous form of Eq. (1),

$$\mathbf{Z}_{\text{sys}}\mathbf{w}=[\mathbf{Z}_{\text{fl}}(\omega)+\mathbf{Z}_{\text{st}}]\mathbf{w}=0, \quad (6)$$

to determine the fluid-loaded resonances and corresponding mode shapes, as the fluid dynamic stiffness matrix is itself a function of frequency, as evident in Eq. (6). However, we may circumvent this problem through the use of the state-space method of Giordano and Koopmann,⁷ as revised by Cunefare and De Rosa⁸ to directly compute the fluid-loaded system resonance frequencies. We briefly sketch the state

space method below. More detail on the method may be found in the papers of Giordano and Koopman⁷ and Cunefare and De Rosa.⁸

The state space method uses a polynomial fit in ka , where k is the acoustic wave number and a is a characteristic dimension, to approximate the individual elements of the fluid dynamic stiffness matrix,

$$\mathbf{Z}_{\text{fl}}(ka)_{ij} = \sum_{l=0}^L C_{lij}(ka)^l. \quad (7)$$

There will be $L+1$ \mathbf{C} matrices containing the coefficients of the interpolation for each term of the fluid matrix, where L is the order of the polynomial expansion. Here, we use a third-order least-square polynomial fit, such that

$$\mathbf{Z}_{\text{fl}} = \mathbf{C}_0 + \mathbf{C}_1(ka) + \mathbf{C}_2(ka)^2 + \mathbf{C}_3(ka)^3. \quad (8)$$

The fitted coefficient matrices \mathbf{C} depend on the range of ka over which the fit is performed as well as the number of discrete frequencies within the range for which \mathbf{Z}_{fl} is available. The coefficient matrices are complex, and may be generated by any suitable curve fitting algorithm. To perform the fit, we first compute the full fluid dynamic stiffness matrix at a number of discrete frequencies. The selection of these frequency points is at our discretion, subject only to the necessity of obtaining a valid fit. The product of Eq. (8) and the displacement vector \mathbf{w} is then

$$\mathbf{Z}_{\text{fl}}(ka)\mathbf{w} = \mathbf{C}_0\mathbf{w} + \mathbf{C}_1(ka)\mathbf{w} + \mathbf{C}_2(ka)^2\mathbf{w} + \mathbf{C}_3(ka)^3\mathbf{w}. \quad (9)$$

With

$$\dot{\mathbf{w}} = -j\mathbf{w}(ka), \quad \ddot{\mathbf{w}} = -\mathbf{w}(ka)^2, \quad \dddot{\mathbf{w}} = j\mathbf{w}(ka)^3, \quad (10)$$

then Eq. (9) becomes

$$\mathbf{Z}_{\text{fl}}(ka)\mathbf{w} = \mathbf{C}_0\mathbf{w} + j\mathbf{C}_1\dot{\mathbf{w}} - \mathbf{C}_2\ddot{\mathbf{w}} - j\mathbf{C}_3\ddot{\mathbf{w}}. \quad (11)$$

We use Eq. (11) and the structural dynamic stiffness matrix \mathbf{Z}_{st} and construct a state-space equivalent to Eq. (1) as

$$\begin{bmatrix} -j\mathbf{C}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{Bmatrix} \ddot{\mathbf{w}} \\ \dot{\mathbf{w}} \\ \mathbf{w} \end{Bmatrix} + \begin{bmatrix} -\mathbf{C}_2 + \mathbf{M} & j\mathbf{C}_1 & \mathbf{C}_0 + \mathbf{K} \\ -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \ddot{\mathbf{w}} \\ \dot{\mathbf{w}} \\ \mathbf{w} \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{Bmatrix}, \quad (12)$$

where \mathbf{I} is the identity matrix. In Eq. (12), we have decomposed \mathbf{Z}_{st} into its mass and stiffness components. Note that we may introduce structural damping into this representation through the use of a complex stiffness \mathbf{K} . We note that such a complex stiffness would be represented as $\mathbf{K}(1-j\eta)$, where the minus sign on the structural damping ratio is a consequence of our assumed time dependency of $e^{-j\omega t}$. Note that such a representation is a uniform application of damping to the entire structure under consideration. We apply standard eigenvalue routines to Eq. (12) to determine the fluid-loaded resonance frequencies and corresponding mode shapes.

IV. DETERMINATION OF RESONANCE PERTURBATION

We use the eigenvalue perturbation development of Rogers⁹ and Meirovitch¹⁰ to estimate the change in the system resonances due to perturbations in the system dynamic stiffness matrix. We consider a matrix \mathbf{A} of size $M \times M$, composed of a matrix \mathbf{A}_0 and some perturbation \mathbf{A}_1 ,

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1, \quad (13)$$

and seek to predict the eigenvalues of \mathbf{A} , given those of \mathbf{A}_0 . As it relates to our problem, we identify \mathbf{A}_1 as \mathbf{Z}_{δ} , and \mathbf{A}_0 as $\mathbf{Z}_{\text{fl}} + \mathbf{Z}_{\text{ref}}$. The eigensolution for \mathbf{A}_0 (obtained through the state space method) is

$$\mathbf{A}_0\mathbf{x}_{0i} = \lambda_{0i}\mathbf{x}_{0i}, \quad \mathbf{y}_{0i}^T\mathbf{A}_0 = \lambda_{0i}\mathbf{y}_{0i}^T \quad \text{with } i \in M, \quad (14)$$

$$\mathbf{y}_{0j}^T\mathbf{x}_{0i} = \delta_{ij} \quad \text{with } i, j \in M,$$

$$\mathbf{y}_{0j}^T\mathbf{A}\mathbf{x}_{0i} = \lambda_{0i}\delta_{ij} \quad \text{with } i, j \in M, \quad (15)$$

where \mathbf{x} and \mathbf{y} are the right- and left-hand eigenvectors, respectively, and λ are the eigenvalues (in the state-space method, the eigenvalues are directly the resonance frequencies). The eigenvalues for the perturbed matrix \mathbf{A} are then represented as

$$\lambda_i = \lambda_{0i} + \lambda_{1i}, \quad (16)$$

where the λ_{0i} are from Eq. (14) and the λ_{1i} are found from the perturbation \mathbf{A}_1 as

$$\lambda_{1i} = \mathbf{y}_{0i}^T\mathbf{A}_1\mathbf{x}_{0i} \quad \text{with } i \in M. \quad (17)$$

If \mathbf{A}_1 is due to some perturbation Δ_X , then the λ_{1i} represent $\partial\lambda/\partial\Delta_X$.

The eigenvalues obtained through the state space method are fully complex, with the imaginary part corresponding to the resonance frequency, and the real part corresponding to a damping coefficient. \mathbf{Z}_{δ} may be obtained through the explicit difference between two model feature representations, or through the use of design sensitivity methods as are currently implemented in many structural analysis codes.

V. 1-DOF PERTURBATION MODEL

As noted in the Introduction, the extensive prior art in the analysis of fluid-loaded structures leads us to focus our attention on the use of a single degree of freedom model for the response in the region of the fluid-loaded resonances. From basic vibration theory, we may state the amplitude response of the i th mode at the reference dynamic stiffness state is

$$x_{0,i} = \frac{F_i m_i^{-1}}{\omega_{0,i}^2 - \omega^2 - j\eta_{0,i}\omega_{0,i}^2} = \frac{F_i m_i^{-1} \omega_{0,i}^{-2}}{1 - r_{0,i}^2 - j\eta_{0,i}}, \quad (18)$$

where F_i , m_i , $\eta_{0,i}$ and $\omega_{0,i}$ are the i th modal force, modal mass, modal damping, and reference resonance frequency, respectively, and $r_{0,i} = \omega/\omega_{0,i}$. The minus sign before the damping term in Eq. (18) is a consequence of our assumed time dependency. We have introduced the frequency ratio r_0 as it will be convenient in our later analysis to deal with a

nondimensional frequency parameter. For the i th mode in the perturbed dynamic stiffness state,

$$x_{\delta,i} = \frac{F_i m_i^{-1}}{\omega_{\delta,i}^2 - \omega^2 - j \eta_{\delta,i} \omega_{\delta,i}^2} = \frac{F_i m_i^{-1} \omega_{\delta,i}^{-2}}{1 - r_{\delta,i}^2 - j \eta_{\delta,i}}, \quad (19)$$

where we have intentionally used the same symbol for the modal forcing and modal mass as for the reference state, and where the subscript δ indicates the perturbed parameter. We assume that the modal mass and modal force are unchanged by the dynamic stiffness perturbation. With this assumption, and comparing Eqs. (18) and (19), the difference in the response between the reference and perturbed states is due solely to the differences between the reference and perturbed states' resonances frequencies and modal dampings. The assumption that the modal mass and forcing are not changed by the perturbation is a limitation to the method developed here, with the assumption on unchanged modal forcing being more restrictive. While small changes in a system's mass will not alter its modal mass significantly (unless the mode shapes change dramatically), small changes in a mode shape may lead to large changes in the modal forcing if the applied force is on or near a mode's nodal position. A modal forcing very small compared to others for a given structure is indicative of weak modal coupling such as occurs with a forcing at a nodal position, and may be used as an indicator of potential unreliable results for the analysis developed here, but only for that specific mode or modes with weak coupling. Note also that the assumption of unchanged modal forcing does not permit this analysis to be used to consider the impact of dramatically different forcing function distributions.

We now exploit the fact that at and near a resonance the radiated acoustic power is proportional through some constant to the magnitude squared of the modal response,^{2,6,11-13} such that the ratio of the power radiated at the perturbed state to the power at the reference state is simply the ratio of the squared magnitudes of the modal amplitudes at the reference and perturbed states. Using Eqs. (18) and (19), we thus obtain

$$\frac{W_{\delta,i}}{W_{0,i}} = \frac{|x_{\delta,i}|^2}{|x_{0,i}|^2} = r_{0\delta}^4 \frac{(1 - r_{0,i}^2)^2 + \eta_{0,i}^2}{(1 - r_{0\delta}^2 r_{0,i}^2)^2 + \eta_{\delta,i}^2}, \quad (20)$$

where we define the resonance ratio $r_{0\delta}$ as the ratio between the reference and perturbed resonance frequencies of the i th mode, $r_{0\delta} = \omega_{0,i} / \omega_{\delta,i}$. In the limit of zero damping, Eq. (20) becomes

$$\frac{W_{\delta,i}}{W_{0,i}} = r_{0\delta}^4 \frac{(1 - r_{0,i}^2)^2}{(1 - r_{0\delta}^2 r_{0,i}^2)^2}. \quad (21)$$

It is evident from Eq. (21) that the response perturbation ratio for undamped systems is unbounded at $r_{0,i} = 1/r_{0\delta}$ (which corresponds to the condition $\omega = \omega_{\delta,i}$), and has a minimum of zero at $r_{0,i} = 1$ (which corresponds to the condition $\omega = \omega_{0,i}$).

Equation (20) is our 1-DOF model for the perturbation in the radiated acoustic power due to a perturbation in the system. It is a function of frequency through $r_{0,i}$, with dependence on the resonance ratio and the reference and per-

turbed modal damping. Equation (20) is actually quite general, describing the perturbation of any squared quantity related to the displacement of a given mode where the perturbation manifests itself solely in the eigenvalue (both its real and imaginary parts, i.e., resonance frequency and damping).

We may determine the maximum and minimum values of the perturbation response ratio in terms of the parameters through the usual practice of setting the differential of Eq. (20) with respect to $r_{0,i}$ equal to zero, and solving the resulting equation for $r_{0,i}$. This process yields two frequency ratios (one for the maximum, one for the minimum),

$$r_{1,2}^2 = \frac{\alpha r_{0\delta}^4 - \beta \pm [(\beta - \alpha r_{0\delta}^4)^2 - 4 r_{0\delta}^2 (r_{0\delta}^2 - 1)(\alpha r_{0\delta}^2 - \beta)]^{1/2}}{2 r_{0\delta}^2 (r_{0\delta}^2 - 1)}, \quad (22)$$

where we retain only the positive roots of the left hand side, and where we have defined

$$\alpha = 1 + \eta_{0,i}^2 \quad \text{and} \quad \beta = 1 + \eta_{\delta,i}^2. \quad (23)$$

The maximum and minimum perturbation may then be determined by substituting the two frequency ratios obtained from Eq. (22) into Eq. (20).

We may further exploit Eq. (20) to answer the question: At what frequency ratio does the perturbed response fall within a certain dB difference, if at all, of the unperturbed response? This is accomplished by setting the left-hand side of Eq. (20) to the desired value of the ratio, and solving for the frequency ratio that satisfies that value. If we define μ as the desired value of the response ratio, Eq. (20), we find

$$\begin{aligned} \mu + \mu \eta_{\delta}^2 - r_{0\delta}^4 - \eta_{0}^2 r_{0\delta}^4 - 2 \mu r_{0\delta}^2 r_0^2 + 2 r_{0\delta}^4 r_0^2 - r_{0\delta}^4 r_0^4 \\ + \mu r_{0\delta}^4 r_0^4 = 0, \end{aligned} \quad (24)$$

which is a quadratic equation in r_0^2 . Only the real positive roots are of physical significance here. Depending on the values of the damping ratio and desired response ratio μ , Eq. (24) will yield two, one or no real positive roots corresponding to two, one or no frequency ratios that yield the desired response ratio.

We may use Eq. (20) to explore the influence of the resonance ratio and modal damping parameters on the acoustic power response ratio in the vicinity of a resonance. Figure 1 presents plots of Eq. (20) as a function of frequency ratio for a number of damping values and values of $r_{0\delta}$. We have assumed the same value for η_0 and η_{δ} for simplicity. Note that $r_{0\delta} < 1$ indicates that the reference resonance frequency is less than the perturbed resonance frequency, and $r_{0\delta} > 1$ indicates that the reference resonance frequency is greater than the perturbed resonance frequency. Figure 1 clearly indicates that the magnitude of the response perturbation is a strong function of the amount of damping in the system. Note that we have not plotted the no-damping condition [Eq. (21)], as it is virtually indistinguishable from the $\eta = 0.001$ line, except in the immediate vicinity of the minimum and maximum. In this approximation, the source of the damping, be it acoustic radiation damping or structural damping, is not important. We also note in Fig. 1 that the greater the difference (positive or negative) between $r_{0\delta}$ and unity, the greater

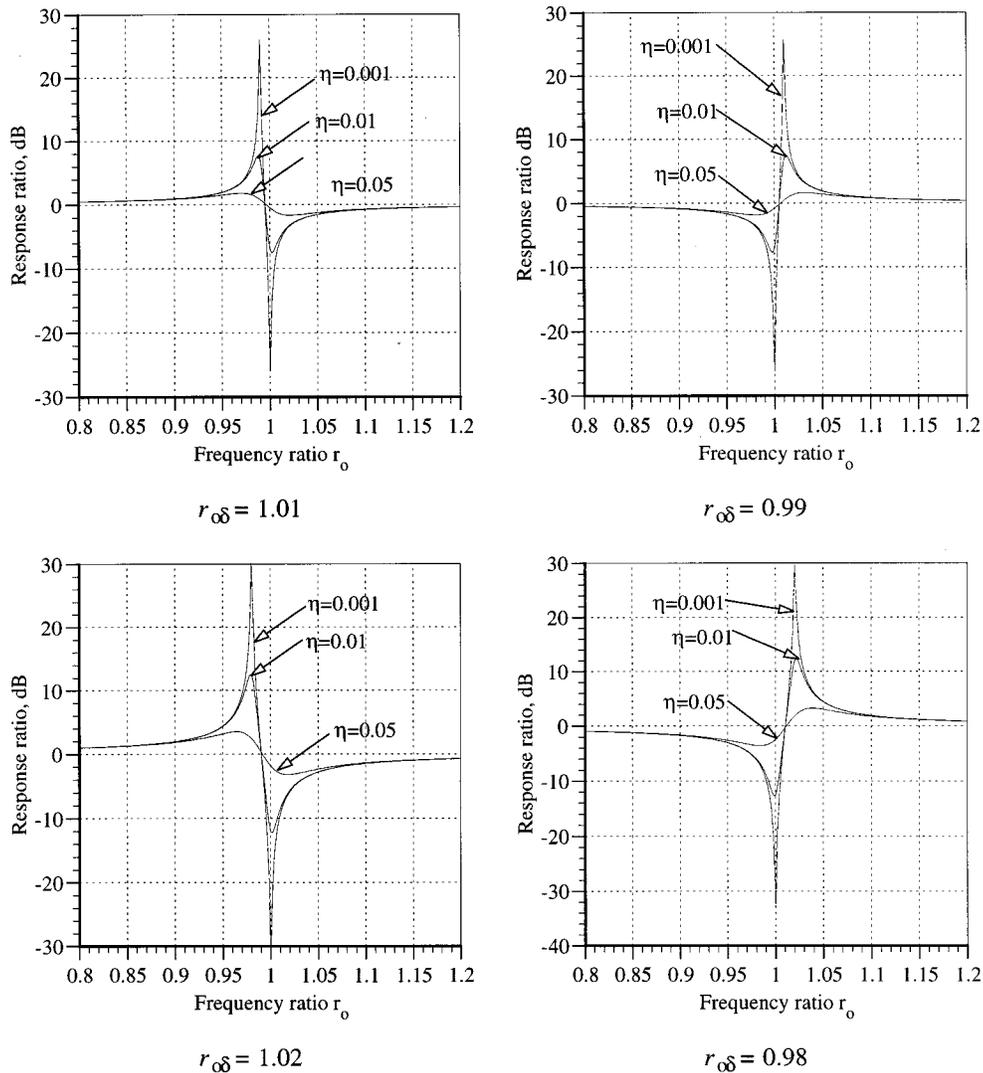


FIG. 1. Influence of resonance frequency ratio and damping on response ratio.

the maximum magnitude of the response ratio, as well as the lesser the minimum magnitude of the response ratio. Finally, we note that the response ratio curves for the different damping values are nearly coincident once the frequency ratio exceeds $\pm 5\%$ of unity. This indicates that outside of the range of frequency ratios $r_0 = 1 \pm 0.05$, the amount of damping is no longer the critical influence on the response ratio.

The above example leads rather naturally into yet another perspective for which the analysis we have developed would be useful: suppose that the eigenvalues are known within a certain precision, say $\pm 1\%$. For this situation, the curves defined by the $r_{0\delta} = 0.99$ and $r_{0\delta} = 1.01$ plots in Fig. 1 represent the potential variation in the response ratio in the vicinity of the eigenvalue. This actually has practical relevance, for example, in modal analysis where response eigenvalues are known within a certain precision: the development here then provides the means to bound the response.

Finally, and significantly, note that Eq. (20) clearly indicates that the bandwidth of frequencies influenced by a change in a given mode's resonance frequency is proportional to the resonance frequency. To clarify this statement, consider two modes of identical damping but with one hav-

ing a resonance frequency of 10 Hz and the other of 100 Hz. Equation (20) implies that these two modes will have the same response perturbation at the same values of their frequency ratios, but, in terms of the absolute frequency, the frequency span of the perturbed response for the 100-Hz mode will be ten times that of the 10-Hz mode. The implication is clear: perturbation or imprecision in the eigenvalues of higher frequency modes has much greater impact (in terms of frequency span) than lower frequency modes. This increasing perturbation bandwidth with increasing resonance frequency may have relevance to what the acoustics modeling community terms the "mid-frequency" range: that troublesome region in between low frequencies where determinist methods (finite element and/or boundary element methods) yield acceptable results, and high frequencies where statistical methods (statistical energy analysis) yield acceptable results. In the region where deterministic methods are valid, one may presume that resonance frequencies are known with sufficient precision that the perturbation bandwidth is small. As the frequency of interest increases, the bandwidth of perturbation increases. This behavior is observed in practice: for products subject to manufacturing

variability (e.g., automobiles¹⁴); small differences between products become significant in terms of the response of that product as the frequency increases. The analysis developed here helps explain the cause for this variability: what is modeled is not exactly what is built, and each item that is built is not exactly like the others. The differences between what is modeled and what is built, and the sample-to-sample variation then leads to different responses than what was modeled, with the response variation due to the differences increasing with increasing frequency.

VI. BOUNDS DEFINITION

We now have the elements required to bound the perturbation in the structural acoustic response. We consider two approaches to the development of bounds. First, we may use the single-degree-of-freedom model, Eq. (20), as a continuous function in the vicinity of the resonances. The resulting values are then the variations that may be expected from the reference state. The second approach is to simply develop a ‘‘bounding box’’ that sets bounding limits on the frequency range and magnitude of the perturbation. This is accomplished through the use of Eqs. (20) and (24) to determine the frequency range required for the perturbed response to fall within some limiting value of the unperturbed response, and Eqs. (20) and (22) to determine the maximum and minimum magnitude perturbations and their frequencies. The frequency range and the magnitude response then permits us to develop a ‘‘perturbation envelope’’ in both frequency and magnitude within which we expect the perturbed response to fall.

There are no particular advantages of one bounding method over the other, with the exception that the second approach is somewhat less computationally expensive. Even so, because both methods use rather simple equations, the computational expense for both methods is small compared to the larger effort of solving the underlying unperturbed structural model.

The challenge to apply the above analysis is to obtain and evaluate Eq. (20) for the particular response variable and perturbation of interest. We note that many commercially available FEM codes provide the means to obtain design sensitivity information. Such design sensitivities are typically proposed for use in design optimization or design tradeoff studies, rather than for considering the influence that model detail has on predicted response for a specific design.

VII. EXAMPLE OF A COUPLED STRUCTURAL ACOUSTIC MODEL

We consider the semi-infinite simply supported plate shown in Fig. 2. To demonstrate the methodology, we consider the variation introduced in the response through different representations of a mass attached to the plate. In the context of the development presented earlier, the plate with no attachments is the parent structure, the plate with a line-mass attachment is the reference structure, and the perturbation will be the difference between the line-mass attachment and a distributed attachment. Therefore, one perspective on

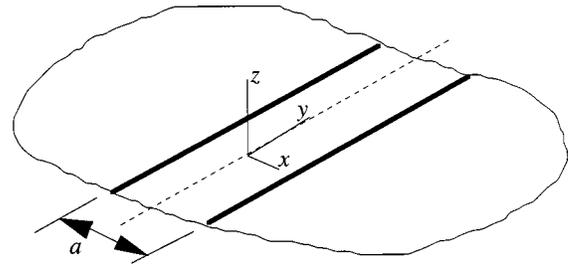


FIG. 2. Semi-infinite fluid-loaded elastic plate of width a in rigid baffle.

the example problem is that we are examining the perturbation that one may expect by modeling a particular distributed attachment feature as a lumped feature.

In each of the following subsections, we develop the model components necessary to perform our analysis, that is, \mathbf{Z}_{fl} , \mathbf{Z}_p , \mathbf{Z}_a , and \mathbf{Z}_δ .

VIII. FLUID MODEL

We use the surface variational principle (SVP) and the method of assumed modes to represent the fluid and structural responses, respectively. Note that this structure, the use of the SVP, and the method of assumed modes are only the means to the end of obtaining a coupled fluid-structure system. The ultimate focus of this work is not dependent on such considerations. Therefore, and rather than replicate material that may be found in great detail elsewhere,¹⁵ we present here only those elements of the system equations necessary to illustrate the application of our methods to the example structure.

The SVP and method of assumed modes permit us to relate pressure and displacement as

$$\mathbf{p} = 2\pi(ka)^2 \mathbf{A}^{-1} \mathbf{R}^T \mathbf{w} = \mathbf{Z}_{\text{ac}} \mathbf{w}, \quad \mathbf{Z}_{\text{ac}} = 2\pi(ka)^2 \mathbf{A}^{-1} \mathbf{R}^T, \quad (25)$$

where \mathbf{A} is a matrix relating coupling between pressure basis functions, \mathbf{R} represents a matrix of interactions between pressure and displacement basis functions, \mathbf{p} is a vector of unknown pressure coefficients, and \mathbf{w} is a vector of unknown displacement coefficients. The fluid dynamic stiffness matrix is then

$$\mathbf{Z}_{\text{fl}} = \mathbf{R} \mathbf{Z}_{\text{ac}}. \quad (26)$$

IX. PARENT STRUCTURE MODEL

The elements of the parent structure’s dynamical dynamic stiffness matrix are

$$Z_{p,jk} = \frac{\rho_s h}{\rho a} \left(\left(\frac{c_s}{c} \right)^2 \kappa_{jk} - (ka)^2 \mu_{jk} \right), \quad (27)$$

where h is the plate thickness, ρ_s is the plate density, c_s is the characteristic wave speed for the structure and depends on the bending stiffness of the plate, $D = Eh^3/(12(1-\nu^2))$ with ν Poisson’s ratio. In Eq. (27), μ and κ are generalized inertia and stiffness coefficients obtained from the assumed mode’s basis functions.

We presume a harmonic forcing function applied to the system. For the work presented here, the harmonic forcing

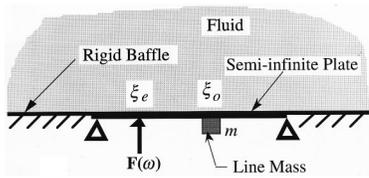


FIG. 3. Semi-infinite plate with line mass attachment.

function is a line excitation, $f(\xi) = f\delta(\xi - \xi_e)$, where ξ_e is the location of the force of magnitude f (note that $\xi = x/a$, where a is the width of the plate).

X. ATTACHMENT MODEL

For the purposes of the bounding analysis, the perturbation comes down to the differences in how we model an attachment feature. We consider a line mass attached to the semi-infinite plate as shown in Fig. 3. The attached mass is presumed to be a locally reacting structure, in that it can only exert forces normal to the surface of the plate. The mass does not react to bending forces. We will consider two different representations of the attachment: a line distribution for the reference state, and a finite-width distribution as the perturbation.

For our reference attachment state, we chose to model the mass as a lumped element that confines all of the mass to a single line parallel to the edge of the plate,

$$m_p(\xi) = m_{rat} \frac{\delta(\xi - \xi_0)}{a}, \quad (28)$$

where m_{rat} is the total added mass of the attachment divided by the plate mass, $m_{rat} = m_{added}/\rho_s h a$, and where $\delta(\xi - \xi_0)$ represents the Dirac delta function at ξ_0 .

For the finite width distribution, we represent the attached mass as

$$m(\xi) = \frac{m_{rat}}{\Delta_m} [u(\xi - (\xi_0 - \Delta_m/2)) - u(\xi - (\xi_0 + \Delta_m/2))], \quad (29)$$

where $u(\xi)$ is the unit step function. We may change the distribution of the attached mass by changing the width of the distribution Δ_m . While the value of Δ_m is changed, we hold the total mass constant. A few example distributions are shown in Fig. 4 for a mass located at $\xi = 0.2071$. As Δ_m decreases the mass becomes more centralized around the at-

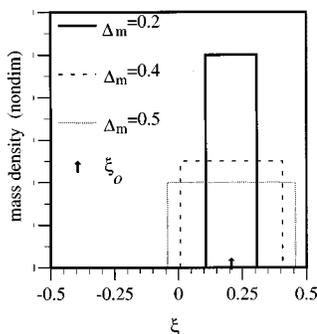


FIG. 4. Example step distributions for mass located at $\xi_0 = 0.2071$.

tachment position, ξ_0 . Conversely, as Δ_m increases the mass becomes more distributed over the plate.

The attachment's dynamic stiffness is

$$Z_{a,jk} = -\frac{\rho_s h}{\rho a} (ka)^2 \eta_{jk}, \quad (30)$$

where the attachment's inertial coupling coefficient, η_{jk} , depends on the type of distribution used. For the lumped representation in Eq. (28),

$$\eta_{jk} = m_{rat} \phi_j(\xi_0) \phi_k(\xi_0), \quad (31)$$

while for the step distribution in Eq. (29),

$$\eta_{jk} = \frac{m_{rat}}{\Delta_m} \int_{\xi_0 - \Delta_m/2}^{\xi_0 + \Delta_m/2} \phi_j \phi_k d\xi. \quad (32)$$

In Eqs. (31) and (32), the ϕ functions are basis functions used in the SVP development, and their detail is not important to the development at hand. From Eqs. (31) and (32), we note that the perturbation is fundamentally the difference between the inertial coupling coefficients for the two representations.

We build Z_a using Eqs. (30) and (31), and Z_δ using Eq. (30) and the difference between Eqs. (32) and (31). We build the reference state, Z_0 , using Z_a and Eq. (27).

XI. RESULTS

We are interested in the response perturbation with respect to a reference state, and, as such, it is the ratio of the response between the reference and perturbed states that is of importance. Therefore, we will present our results in terms of dB differences with respect to the reference state. One may always perform a complete analysis, using a detailed model of different attachments, and thereby directly calculating the perturbed response. In the following, we will use the term "direct analysis" to refer to this method of assessing the response perturbation. Our single-degree-of-freedom method will be referred to as the "SDOF analysis."

We chose plate and fluid properties to represent a steel plate in water (heavy fluid loading). Using the modeling methods developed above, we consider an attached line mass representing 25% of the mass of the parent plate structure. We will assess the impact of modeling the attached mass as either a lumped element, or as distributed uniformly over 5% of the span of the plate. The mass will be modeled as being attached at $\xi = 0.2071$. We will use a line force located at $\xi = 0.25$ (magnitude is not relevant to this analysis, as we are interested in the ratio of perturbed to unperturbed responses). We will assess the impact of heavy fluid and light fluid loading, as well as the presence or absence of structural damping. The underlying model is the same as that used by Shepard and Cunefare,² and more explicit detail may be found in that reference.

For each case considered below, we use the state-space method to locate the fluid-loaded resonances and associated damping. We then apply the direct and SDOF analyses to determine the response perturbation. For the direct analysis, we compute the ratio of the unperturbed and perturbed responses at discrete ka values ranging from 0.05 to 10.0 with

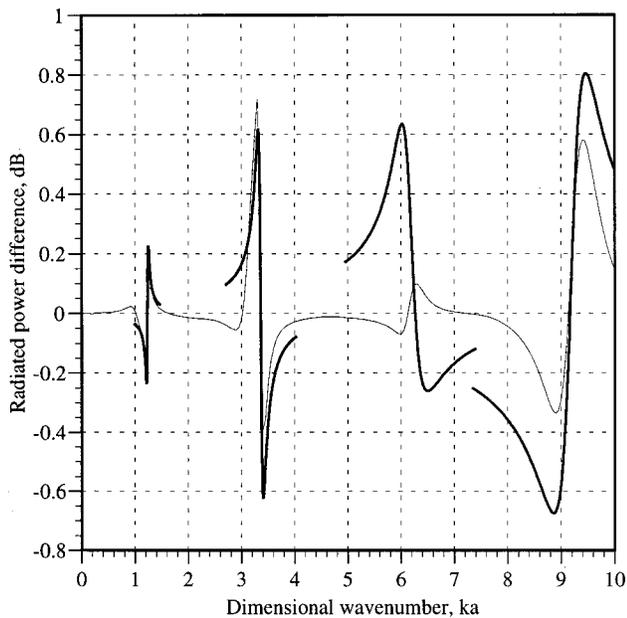


FIG. 5. Perturbation in response ratio of lumped element versus distributed mass with $\Delta_m = 0.05$, determined through direct analysis and through single degree of freedom estimate, undamped, heavy fluid loaded; — direct analysis, — SDOF analysis.

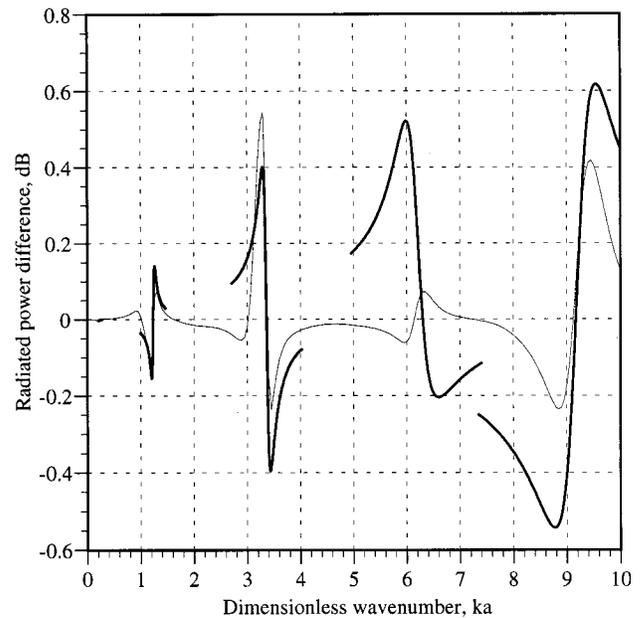


FIG. 6. Perturbation in response ratio of lumped element versus distributed mass with $\Delta_m = 0.05$, determined through direct analysis and through single degree of freedom estimate, damped, heavy fluid loaded; — direct analysis, — SDOF analysis.

a step size of 0.05 in ka . We use the individual resonance frequencies (expressed in terms of wave number, ka) and damping values for the SDOF analysis, at frequencies representing a span of $\pm 20\%$ of the resonance frequency. This choice of a percentage-of-resonance-frequency for the span of the SDOF analysis was made to illustrate the increasing span of perturbed response with increasing frequency, as noted in the discussion at the end of the section where we developed Eq. (20).

Figures 5 and 6 are for the model as described above with heavy fluid loading (water), with Fig. 5 representing the results of the analysis with no structural damping and Fig. 6 the results with a structural loss factor of 2%. We note that in heavy fluid, the inclusion of structural damping in the plate does not influence the response perturbation to any great extent. For the specific set of analysis parameters chosen here, the response perturbation is limited to less than 1 dB over the range of wave numbers considered. The SDOF analysis method quite ably captures the response perturbation for each of the resonances, with the exception of the one near $ka = 6.2$. It turns out that the selected position for the line force is nearly coincident with a nodal point of the mode whose resonance is near $ka = 6.2$. As a consequence, the forcing cannot efficiently excite this mode, and small changes in the mode shape can lead to large changes in the modal forcing. Recall that one of the underlying assumptions of the SDOF analysis was that the modal forcing does not change between the perturbed and unperturbed states. In our example problem near $ka = 6.2$, this assumption is not valid. This failure illustrates a drawback of the method as it currently stands: If a given mode's modal forcing is itself sensitive to the modeling perturbation, then the SDOF model will not yield accurate results for that mode. We will address this issue in our future work.

Figures 7 and 8 are for the model as described above with light fluid loading, with Fig. 7 representing the results of the analysis with no structural damping and Fig. 8 the results with a structural loss factor of 2%. To model the light fluid loading case, we arbitrarily multiplied the fluid dynamic stiffness matrix derived assuming water loading by 0.001. The effect of this is to represent a fluid with a density of 0.1% of water, but the same sound speed (e.g., a low molecular weight gas). Unlike the heavy fluid analysis, the in-

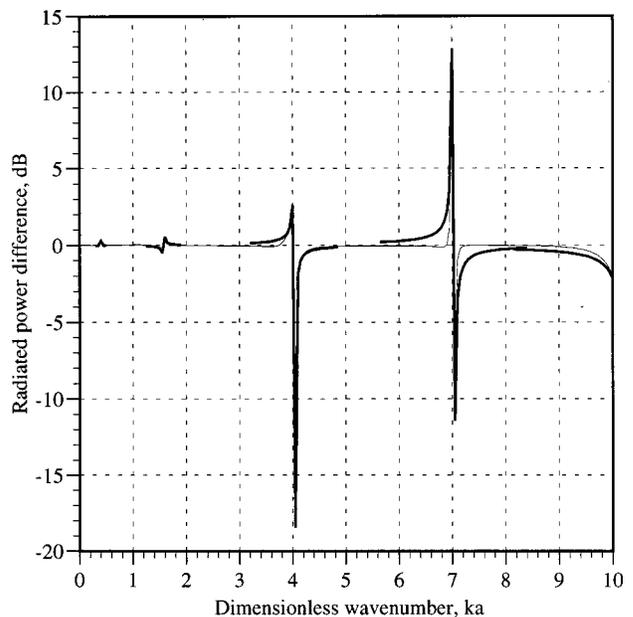


FIG. 7. Perturbation in response ratio of lumped element versus distributed mass with $\Delta_m = 0.05$, determined through direct analysis and through single degree of freedom estimate, undamped, light fluid loaded; — direct analysis, — SDOF analysis.

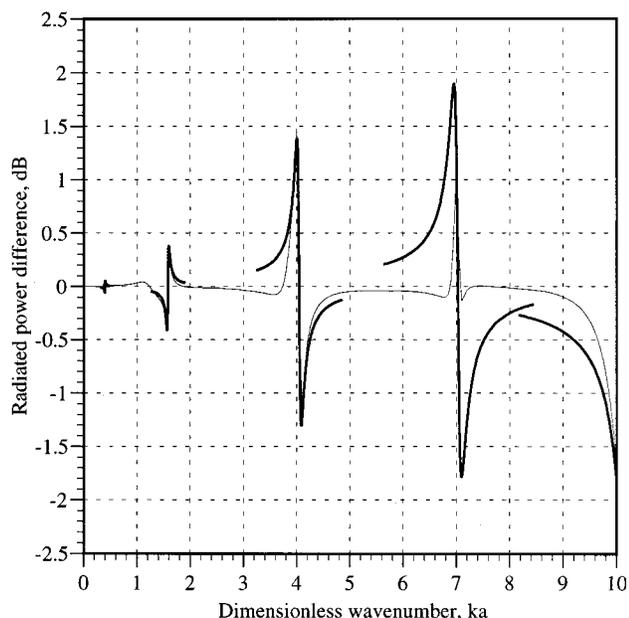


FIG. 8. Perturbation in response ratio of lumped element versus distributed mass with $\Delta_m = 0.05$, determined through direct analysis and through single degree of freedom estimate, damped, light fluid loaded; — direct analysis, — SDOF analysis.

clusion of structural damping in the plate strongly influences the response perturbation. In the absence of structural damping, the only damping present in the system is that due to the radiation of acoustic energy to the far field. Nonetheless, the SDOF analysis is able to well represent the perturbation, again with one exception and for the same reasons as discussed above. Note that Figs. 7 and 8 illustrate the strong influence of damping on the response perturbation near resonances, as generally discussed with respect to Fig. 1.

In each of Figs. 5–8, it is evident that the span of the SDOF analysis increases with wave number, even though the underlying SDOF analysis is in terms of a constant percentage of frequency ratio. Each figure also shows that for the chosen frequency ratio span, the physical wave number spans for the analysis for adjacent modes begin to overlap for the highest two resonances considered in this analysis. Again, this serves to illustrate that the resonance frequency perturbation has greater impact for higher frequency modes than lower frequency modes.

Figures 5–8 also are consistent with the prior results of Shepard and Cunefare,² which demonstrated that the response of this system is sensitive to perturbations only near the fluid loaded resonances (in fact, this prior work, which demonstrated localized impact only near the resonances, is what prompted the development at hand).

XII. CONCLUSIONS

The analysis method demonstrates that a simple SDOF approximation may be used to assess the perturbation in the radiated acoustic power due to perturbations in an underlying structural acoustic model. Elements of the analysis method are not necessarily restricted to model perturbations nor

acoustic power; rather, they may be used to assess the perturbation of any quadratic response quantity of interest due to changes in resonance frequency.

In general terms, the SDOF analysis reveals that the bandwidth of response perturbation increases with increasing resonance frequency. For frequencies within about $\pm 5\%$ of a resonance frequency, the amount of damping in the system determines and limits the magnitude of the response perturbation. The perturbation outside the range of $\pm 5\%$ of the resonance frequency is relatively insensitive to damping.

The example fluid-loaded plate model validated the SDOF analysis method, and reinforced the role of damping on limiting the response perturbation. Further, the example illustrated the bandwidth of response perturbation increasing with increasing resonance frequency, and demonstrated that for adjacent resonances the bandwidths can overlap.

The SDOF analysis method, at present, is limited by its assumption of constant modal forcing and modal mass between the reference and perturbed states. The assumption on unchanged modal forcing is more restrictive than that of unchanged modal mass. A modal forcing very small compared to others for a given structure is indicative of weak modal coupling such as occurs with a forcing at a nodal position, and may be used as an indicator of potential unreliable results for the analysis developed here, but only for that specific mode or modes with weak coupling. We will address the significance of these assumptions in future work.

ACKNOWLEDGMENTS

The genesis of the work presented in this paper was performed during S. De Rosa's three month visiting research period at the Georgia Institute of Technology, from 20 September to 19 December 1996. The visit was made possible by a Fulbright Grant. The authors would like to acknowledge the Council For The International Exchange of Students and the Commissione Per Gli Scambi Culturali Fra L'Italia E Gli Stati Uniti for having selected and approved the research plan. These organizations gave the authors the opportunity to increase their knowledge and experience by working together. The second author wishes also to thank Dr. T. Curry, President and C.E.O. of MSC/NASTRAN Co., for partially supporting his visit to Georgia Tech and for the attention he is giving to this ongoing collaboration between Georgia Tech and the Dipartimento di Progettazione Aeronautica.

¹W. S. J. Shepard, "The impact of attached feature scales and spatial distributions on the response of structural-acoustic systems," Ph.D. Dissertation, The Georgia Institute of Technology, 1996.

²W. S. Shepard, Jr. and K. A. Cunefare, "Sensitivity of structural acoustic response to attachment feature scales," *J. Acoust. Soc. Am.* **102**, 1612–1619 (1997).

³S. R. Hahn and A. A. Ferri, "Sensitivity analysis of coupled structural-acoustic problems using perturbation techniques," *J. Acoust. Soc. Am.* **101**, 1–7 (1997).

⁴J. M. Cuschieri and D. Feit, "Scattering of sound by a fluid-loaded infinite plate with a distributed mass inhomogeneity," *J. Acoust. Soc. Am.* **99**, 2686–2700 (1996).

⁵J. 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).

- ⁶H. G. Davies, "Sound from turbulent-boundary-layer-excited panels," J. Acoust. Soc. Am. **49**, 878–889 (1971).
- ⁷J. A. Giordano and G. H. Koopmann, "State space boundary element-finite element coupling for fluid-structure interaction analysis," J. Acoust. Soc. Am. **98**, 363–372 (1995).
- ⁸K. A. Cunefare and S. De Rosa, "An improved state-space method for coupled fluid-structure interaction analysis," J. Acoust. Soc. Am. **105**, 206–210 (1998).
- ⁹L. C. Rogers, "Derivatives of eigenvalues and eigenvectors," AIAA J. **8**(5), 943–945 (1970).
- ¹⁰L. Meirovitch, *Computational Methods in Structural Dynamics* (Sijthoff & Noordhoff, Rockville, MD, 1980).
- ¹¹C. E. Wallace, "Radiation resistance of a rectangular panel," J. Acoust. Soc. Am. **51**, 946–952 (1972).
- ¹²C. E. Wallace, "Radiation resistance of a baffled beam," J. Acoust. Soc. Am. **51**, 936–945 (1972).
- ¹³H. G. Davies, "Low frequency random excitation of water-loaded rectangular plates," J. Sound Vib. **15**(1), 107–126 (1971).
- ¹⁴M. S. Kompella and R. J. Bernhard, "Variation of structural-acoustic characteristics of automotive vehicles," Noise Control Eng. J. **44**(Mar/Apr), 93–99 (1996).
- ¹⁵J. H. Ginsberg, K. A. Cunefare, and H. Pham, "A spectral description of inertial effects in fluid-loaded plates," Trans. ASME, J. Vib. Acoust. **117**(2), 206–212 (1995).

Near-field sensing strategies for the active control of the sound radiated from a plate

A. Berry

G.A.U.S., Department of Mechanical Engineering, University of Sherbrooke, Quebec J1K 2R1, Canada

X. Qiu and C. H. Hansen

Department of Mechanical Engineering, The University of Adelaide, South Australia 5005, Australia

(Received 3 March 1999; accepted for publication 24 July 1999)

When designing an active control system to globally control the *far-field* sound radiation from a vibrating surface, a challenging problem is to properly define the near-field acoustic sensing strategy and the type of cost function to be minimized by the controller. The strategy of sensing and minimizing the near-field active intensity at discrete locations in the active control of free field radiation from a vibrating plate is investigated in this paper. The optimal minimization of the sum of the near-field, normal active sound intensities at the error sensor locations using acoustic control sources is derived for this problem, and the results obtained are compared to the minimization of the sum of the near-field squared pressures. Some of the difficulties associated with sound intensity minimization are pointed out. © 1999 Acoustical Society of America. [S0001-4966(99)05211-X]

PACS numbers: 43.40.Rj, 43.40.Dx, 43.40.Vn, 43.50.Ki [CBB]

INTRODUCTION

This work is concerned with the active control of free-field acoustic radiation from extended vibrating surfaces. The generic problem has received considerable attention in the past, and is now well documented in textbooks on active control (Nelson *et al.*, 1992, Chap. 8; Hansen *et al.*, 1997, Chap. 8); an important and historical application of active control of free-field radiation is the attenuation of humming noise generated by large electrical transformers (see, e.g., Angevine, 1993).

When designing an active control system to globally control the *far-field* sound radiation from a vibrating surface using acoustic control sources and acoustic error sensors, it is desirable to keep the control sources and error sensors close to the primary vibrating surface: the system is then physically more compact, the control sources need to be located at a distance usually less than an acoustic wavelength from the primary source in order to couple efficiently with the primary sound field (Nelson *et al.*, 1987), otherwise the control sources arrangement is more complex (Kempton, 1976; Bolton *et al.*, 1995). Furthermore, using far-field error sensors usually has the effect of reducing the sound radiation in the direction of the sensors, but not necessarily reducing the sound field globally; far-field sensors also introduce control path transfer functions with large acoustic delays which can result in controller instability, and these control path transfer functions are more prone to time variations, thereby requiring continuous on-line system identification. For all these reasons, error sensors located in the near field of the primary source are more appropriate.

A challenging problem is then to properly define the near-field acoustic sensing strategy and the type of cost function to be minimized by the controller. Previous work that considered the simple problem of controlling the sound radiation from a primary monopole source using a secondary monopole source showed that minimizing near-field sound

pressure does not necessarily reduce far-field sound pressure levels (Hansen *et al.*, 1997). Some observations made on the problem of controlling the sound radiation from electrical transformers using arrays of acoustic control sources and error microphones showed that a better control performance is achieved when the error microphone array is moved to a distance sufficiently far from the transformer surface (Angevine, 1993). As near-field error microphones measure acoustic near-field components not necessarily related to far-field sound radiation, a near-field sensing strategy directly measuring the flow of energy to the far field would be more appropriate in this case. A number of alternatives to sensing and minimizing squared sound pressure have been suggested recently in active noise control applications: these concern the active control of duct noise (Kang *et al.*, 1997; Zander *et al.*, 1993), and more importantly the active control of enclosed sound fields, for which sensing strategies based on acoustic potential energy density minimization and total energy density minimization have been suggested (Park *et al.*, 1997; Sommerfeldt *et al.*, 1995; Cazzolato, 1998). Associated adaptation algorithms for the minimization of energy-based quantities have been derived (Sommerfeldt *et al.*, 1994). Recently, Qiu *et al.* (1998) have compared various acoustic near-field sensing strategies for the active control of the sound field radiated from a monopole source; it was found that sensing the radial active intensity at a number of positions close to the primary source usually gives better control performance (in terms of total sound power attenuation) than sensing the squared pressure, acoustic kinetic energy density, or total kinetic energy density.

This study is a continuation of the previous work of Qiu *et al.* (1998), for the case of a more complex radiator, a vibrating plate, and the strategy of sensing and minimizing the active intensity in the direction normal to the plate surface at a number of near-field points is investigated. The optimal minimization of the sum of the squared sound pressures, or the sum of the active sound intensities at a number

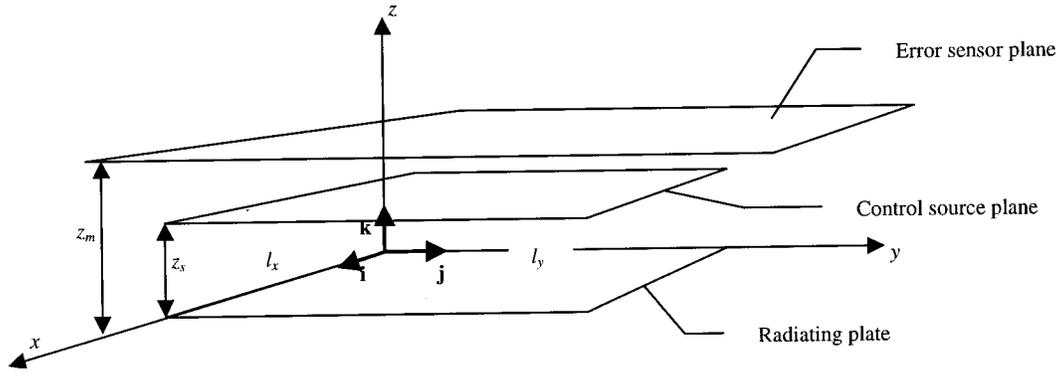


FIG. 1. Physical model description.

of near-field error sensor locations, using an array of control point sources is derived for a primary source consisting of a simply supported vibrating plate. Numerical simulations compare the two sensing strategies in a variety of situations, and some of the difficulties associated with sound intensity minimization are pointed out.

I. PHYSICAL MODEL

A. Description of the model

The physical model used here is a rectangular plate mounted in an infinite rigid baffle. The plate dimensions are l_x , and l_y , and it radiates sound into an unbounded fluid half space (density ρ_0 , sound speed c_0) (Fig. 1). The sound radiation from the plate is controlled by an array of S secondary point sources which lie in the plane of the control sources parallel to the plate, at a distance z_s from the plate. The control aims at minimizing a cost function provided by an array of M error sensors, located in the plane of the error sensors, also parallel to the plate, at a distance z_m from the plate. In the following, two active control strategies are investigated for minimizing the sound radiated from the plate—the minimization of the sum of the squared sound pressures at the error sensors (as would be provided by error microphones) and the minimization of the sum of the normal active sound intensities at the error sensors (as would be provided by sound intensity probes as error sensors).

B. Primary sound field

In general, the transverse displacement of a transversely vibrating plate can be expressed as

$$w(\mathbf{r}_0, t) = \mathbf{W}(t)^T \boldsymbol{\phi}(\mathbf{r}_0), \quad (1)$$

where $\mathbf{r}_0 = x_0 \mathbf{i} + y_0 \mathbf{j}$ is a point on the plate surface, $\mathbf{W}(t) = [W_{11}(t), \dots, W_{mn}(t)]^T$ is the vector of modal displacements, and $\boldsymbol{\phi}(\mathbf{r}_0) = [\phi_{11}(\mathbf{r}_0), \dots, \phi_{mn}(\mathbf{r}_0)]^T$ is the vector of the plate eigenfunctions; for a simply supported plate, $\phi_{mn}(\mathbf{r}_0) = \sin(m\pi x_0/l_x) \sin(n\pi y_0/l_y)$. In the following we assume that only a single mode m, n participates in the plate response, such that

$$w(\mathbf{r}_0, t) = W_{mn}(t) \phi_{mn}(\mathbf{r}_0). \quad (2)$$

In this study, the primary modal displacement W_{mn} is imposed, and is not affected by the acoustic control sources. Only a time-harmonic primary disturbance of the form

$W_{mn}(t) = W_{mn} e^{-j\omega t}$ is considered, where ω is the angular frequency of the disturbance. The time dependence $e^{-j\omega t}$ is omitted for brevity in the following.

The primary acoustic pressure radiated by the plate to point $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is given by the classical Rayleigh integral,

$$p_p(\mathbf{r}) = \omega^2 W_{mn} P_{mn}(\mathbf{r}), \quad (3)$$

where

$$P_{mn}(\mathbf{r}) = \rho_0 \int_{S_0} \phi_{mn}(\mathbf{r}_0) G(\mathbf{r}, \mathbf{r}_0) dS_0 \quad (4)$$

and $G(\mathbf{r}, \mathbf{r}_0) = e^{jk|\mathbf{r}-\mathbf{r}_0|} / (2\pi|\mathbf{r}-\mathbf{r}_0|)$ is the acoustic Green's function; $k = \omega/c_0$ is the acoustic wavenumber. Of interest here is also the primary acoustic velocity, given by

$$\mathbf{v}_p(\mathbf{r}) = \frac{1}{j\omega\rho_0} \nabla_r p_p(\mathbf{r}) = -j\omega W_{mn} \mathbf{V}_{mn}(\mathbf{r}) \quad (5)$$

with $\mathbf{V}_{mn}(\mathbf{r}) = \int_{S_0} \phi_{mn}(\mathbf{r}_0) \nabla_r G(\mathbf{r}, \mathbf{r}_0) dS_0$.

The active acoustic intensity is $\mathbf{i}_p(\mathbf{r}) = \frac{1}{2} \Re(p_p(\mathbf{r}) \mathbf{v}_p^*(\mathbf{r}))$, where \Re denotes the real part and $*$ is the complex conjugate. Substituting (3) and (5) in the above expression yields

$$\mathbf{i}_p(\mathbf{r}) = \frac{\omega^2}{2} |W_{mn}|^2 \Re(\mathbf{I}_{mn}(\mathbf{r}))$$

$$\text{with } \mathbf{I}_{mn}(\mathbf{r}) = j\omega P_{mn}(\mathbf{r}) \mathbf{V}_{mn}^*(\mathbf{r}). \quad (6)$$

The general expressions for the acoustic velocity and intensity are particularized to the z component of the acoustic velocity and intensity,

$$v_{zp}(\mathbf{r}) = \frac{1}{j\omega\rho_0} \frac{\partial p_p(\mathbf{r})}{\partial z} = -j\omega W_{mn} V_{mnz}(\mathbf{r}) \quad (7)$$

with

$$V_{mnz}(\mathbf{r}) = \int_{S_0} \phi_{mn}(\mathbf{r}_0) \frac{\partial G(\mathbf{r}, \mathbf{r}_0)}{\partial z} dS_0 \quad (8)$$

and

$$\frac{\partial G(\mathbf{r}, \mathbf{r}_0)}{\partial z} = \frac{z}{|\mathbf{r}-\mathbf{r}_0|} G(\mathbf{r}, \mathbf{r}_0) \left(jk - \frac{1}{|\mathbf{r}-\mathbf{r}_0|} \right), \quad (9)$$

where z is the height of the receiving point \mathbf{r} .

The first term in (9) contributes to a particle velocity component in phase with the sound pressure (resulting in acoustic energy flowing into the far field), while the second term contributes to a particle velocity component in quadrature with the sound pressure (resulting in zero net acoustic energy flow); this second term is confined to the near field of the plate, typically $kz \leq 1$. These remarks are important in the context of near-field sensing strategies directed towards an attenuation of the far-field sound.

For the z component of the active acoustic intensity,

$$i_{zp}(\mathbf{r}) = \frac{\omega^2}{2} |W_{mn}|^2 \Re(\mathbf{I}_{mnz}(\mathbf{r})) \quad (10)$$

with $\mathbf{I}_{mnz}(\mathbf{r}) = j\omega P_{mn}(\mathbf{r}) V_{mnz}^*(\mathbf{r})$.

In the numerical simulations, discretized versions of Eqs. (4) and (8) are used; the radiating surface is discretized into elementary areas ΔS_{oi} over which the transverse displacement is assumed constant, $\phi_{mn}(\mathbf{r}_{oi})$. Equations (4) and (8) become

$$\begin{aligned} P_{mn}(\mathbf{r}) &\approx \rho_0 \sum_i \phi_{mn}(\mathbf{r}_{oi}) G(\mathbf{r}, \mathbf{r}_{oi}) \Delta S_{oi} V_{mnz}(\mathbf{r}) \\ &\approx \sum_i \int_{S_0} \phi_{mn}(\mathbf{r}_{oi}) \frac{z}{|\mathbf{r} - \mathbf{r}_{oi}|} G(\mathbf{r}, \mathbf{r}_{oi}) \\ &\quad \times \left(jk - \frac{1}{|\mathbf{r} - \mathbf{r}_{oi}|} \right) \Delta S_{oi}. \end{aligned}$$

The radiated sound pressure and z component of the acoustic velocity and active acoustic intensity are evaluated at any arbitrary position \mathbf{r} in the acoustic medium using the above equations; however, due to the singularity of the Green's function $G(\mathbf{r}, \mathbf{r}_{oi})$ and of its normal derivative $\partial G(\mathbf{r}, \mathbf{r}_{oi})/\partial z$ when $\mathbf{r} = \mathbf{r}_{oi}$, these quantities are not evaluated on the plate surface.

In the following sections, active control schemes are considered for minimizing the sound field radiated from the vibrating plate, using two-dimensional arrays of control point-sources and error sensors, both located at a short distance from the plate surface. Essentially two control schemes are investigated—the minimization of the sum of the squared sound pressures at the error sensors, and the minimization of the sum of the active sound intensities at the error sensors.

C. Exact minimization of the near-field sound pressure

In a practical implementation, the control of the near-field sound pressure would require the use of microphones as error sensors. The radiation from a control point source located at $\mathbf{r}_s = x_s \mathbf{i} + y_s \mathbf{j} + z_s \mathbf{k}$ is given by

$$p_s(\mathbf{r}) = q_s G(\mathbf{r}, \mathbf{r}_s), \quad (11)$$

where q_s is the control source strength and $G(\mathbf{r}, \mathbf{r}_s) = e^{jk|\mathbf{r} - \mathbf{r}_s|} / (4\pi|\mathbf{r} - \mathbf{r}_s|) + e^{jk|\mathbf{r} - \mathbf{r}'_s|} / (4\pi|\mathbf{r} - \mathbf{r}'_s|)$ is the acoustic Green's function, and $\mathbf{r}'_s = x_s \mathbf{i} + y_s \mathbf{j} - z_s \mathbf{k}$ is the position of the mirror image of the control source with respect to the plate surface. The expression for the secondary sound field assumes that the plate vibration is not affected by the control

source, and that the plate is seen by the control source as a rigid surface. When several control source with positions $\mathbf{r}_{s1}, \dots, \mathbf{r}_{ss}, \dots, \mathbf{r}_{sS}$, are considered, the secondary sound field is given by

$$p_s(\mathbf{r}) = \mathbf{q}^T [G(\mathbf{r}, \mathbf{r}_{s1}), \dots, G(\mathbf{r}, \mathbf{r}_{ss}), \dots, G(\mathbf{r}, \mathbf{r}_{sS})]^T, \quad (12)$$

where $\mathbf{q} = [q_1, \dots, q_s, \dots, q_S]^T$ is the vector of the S control sources strengths.

The total sound field is the superposition of the primary and secondary sound fields,

$$\begin{aligned} p(\mathbf{r}) &= p_p(\mathbf{r}) + p_s(\mathbf{r}) \\ &= \omega^2 W_{mn} P_{mn}(\mathbf{r}) \\ &\quad + \mathbf{q}^T [G(\mathbf{r}, \mathbf{r}_{s1}), \dots, G(\mathbf{r}, \mathbf{r}_{ss}), \dots, G(\mathbf{r}, \mathbf{r}_{sS})]^T. \end{aligned} \quad (13)$$

The vector \mathbf{q} is adjusted in order to minimize the sum of the squared pressures at the error microphone locations; the cost function in this case can be written

$$J = \mathbf{p}^* \mathbf{p}^T, \quad (14)$$

where

$$\mathbf{p} = [p(\mathbf{r}_{m1}), \dots, p(\mathbf{r}_{mm}), \dots, p(\mathbf{r}_{mM})] \quad (15)$$

is the vector of the total sound pressures measured at the M error microphone positions. Substituting (13) into (15),

$$\mathbf{p} = \mathbf{q}^T \mathbf{G} + \omega^2 W_{mn} \mathbf{p}_p, \quad (16)$$

where

$$\mathbf{G} = \begin{bmatrix} G(\mathbf{r}_{m1}, \mathbf{r}_{s1}) & \cdots & G(\mathbf{r}_{mm}, \mathbf{r}_{s1}) & \cdots & G(\mathbf{r}_{mM}, \mathbf{r}_{s1}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ G(\mathbf{r}_{m1}, \mathbf{r}_{ss}) & & G(\mathbf{r}_{mm}, \mathbf{r}_{ss}) & & G(\mathbf{r}_{mM}, \mathbf{r}_{ss}) \\ \vdots & & \vdots & \ddots & \vdots \\ G(\mathbf{r}_{m1}, \mathbf{r}_{sS}) & \cdots & G(\mathbf{r}_{mm}, \mathbf{r}_{sS}) & \cdots & G(\mathbf{r}_{mM}, \mathbf{r}_{sS}) \end{bmatrix}$$

is the transfer function matrix between the S control sources and the M error microphones, and $\mathbf{p}_p = [P_{mn}(\mathbf{r}_{m1}), \dots, P_{mn}(\mathbf{r}_{mm}), \dots, P_{mn}(\mathbf{r}_{mM})]$ is the primary sound pressure vector at the M error microphones.

Substituting (16) into (14), the cost function is written as a Hermitian quadratic form of the control variable \mathbf{q} (Nelson *et al.*, 1992),

$$J = \mathbf{q}^H \mathbf{A}_p \mathbf{q} + \mathbf{q}^H \mathbf{b}_p + \mathbf{b}_p^H \mathbf{q} + c_p, \quad (17)$$

where

$$\mathbf{A}_p = \mathbf{G}^* \mathbf{G}^T, \quad (18a)$$

$$\mathbf{b}_p = \omega^2 W_{mn} \mathbf{G}^* \mathbf{p}_p^T, \quad (18b)$$

$$c_p = \omega^4 |W_{mn}|^2 \mathbf{p}_p^* \mathbf{p}_p^T, \quad (18c)$$

and H denotes the Hermitian transpose. The cost function J also depends on the amplitude of the primary vibration field, W_{mn} . The first term in (17) represents the sum of the squared pressures at the M error microphones due to the S control sources, and is therefore always positive. The matrix \mathbf{A}_p is thus positive definite; this ensures that J has a unique global minimum. The optimal control source strengths are given by (Nelson *et al.*, 1992)

$$\mathbf{q}_{\text{opt}} = -\mathbf{A}_p^{-1} \mathbf{b}_p$$

and the minimized cost-function is given by $J_{\text{opt}} = c_p + \mathbf{b}_p^H \mathbf{q}_{\text{opt}}$.

D. Exact minimization of the near-field sound intensity

The minimization of the sum of the normal active sound intensities at the error sensors is now considered. Such a control strategy involves the sensing of the normal active sound intensity above the plate, and therefore require intensity probes as error sensors, instead of microphones.

Again, when several control sources with positions $\mathbf{r}_{s1}, \dots, \mathbf{r}_{sS}, \dots, \mathbf{r}_{sS}$ are considered, the secondary sound field is given by

$$p_s(\mathbf{r}) = \mathbf{q}^T \mathbf{g}_s(\mathbf{r}),$$

where $\mathbf{q} = [q_1, \dots, q_s, \dots, q_S]^T$ is the vector of the S control sources strengths, and

$$\mathbf{g}_s(\mathbf{r}) = [G(\mathbf{r}, \mathbf{r}_{s1}), \dots, G(\mathbf{r}, \mathbf{r}_{sS}), \dots, G(\mathbf{r}, \mathbf{r}_{sS})]^T,$$

$$G(\mathbf{r}, \mathbf{r}_s) = e^{jk|\mathbf{r}-\mathbf{r}_s|} / (4\pi|\mathbf{r}-\mathbf{r}_s|) + e^{jk|\mathbf{r}-\mathbf{r}'_s|} / (4\pi|\mathbf{r}-\mathbf{r}'_s|).$$

The z component of the acoustic velocity due to the control sources is given by

$$\mathbf{v}_{zs}(\mathbf{r}) = \frac{1}{j\omega\rho} \frac{\partial p_s(\mathbf{r})}{\partial z} = \mathbf{q}^T \mathbf{U}_s(\mathbf{r}), \quad (19)$$

where

$$\begin{aligned} \mathbf{U}_s(\mathbf{r}) &= \frac{1}{j\omega\rho} \frac{\partial \mathbf{g}_s}{\partial z} \\ &= \frac{1}{j\omega\rho} \left[\frac{\partial G(\mathbf{r}, \mathbf{r}_{s1})}{\partial z}, \dots, \frac{\partial G(\mathbf{r}, \mathbf{r}_{sS})}{\partial z}, \dots, \frac{\partial G(\mathbf{r}, \mathbf{r}_{sS})}{\partial z} \right]^T, \end{aligned}$$

$$\begin{aligned} \frac{\partial G(\mathbf{r}, \mathbf{r}_s)}{\partial z} &= \frac{z-z_s}{|\mathbf{r}-\mathbf{r}_s|} \frac{e^{jk|\mathbf{r}-\mathbf{r}_s|}}{4\pi|\mathbf{r}-\mathbf{r}_s|} \left(jk - \frac{1}{|\mathbf{r}-\mathbf{r}_s|} \right) \\ &\quad + \frac{z+z_s}{|\mathbf{r}-\mathbf{r}'_s|} \frac{e^{jk|\mathbf{r}-\mathbf{r}'_s|}}{4\pi|\mathbf{r}-\mathbf{r}'_s|} \left(jk - \frac{1}{|\mathbf{r}-\mathbf{r}'_s|} \right). \end{aligned}$$

The total sound pressure and acoustic velocity are the superpositions of the primary and secondary sound fields,

$$p(\mathbf{r}) = p_p(\mathbf{r}) + p_s(\mathbf{r}) = \omega^2 W_{mn} P_{mn}(\mathbf{r}) + \mathbf{q}^T \mathbf{g}_s(\mathbf{r}). \quad (20a)$$

$$\mathbf{v}_z(\mathbf{r}) = \mathbf{v}_{zp}(\mathbf{r}) + \mathbf{v}_{zs}(\mathbf{r}) = -j\omega W_{mn} V_{mnz}(\mathbf{r}) + \mathbf{U}_s^T(\mathbf{r}) \mathbf{q}. \quad (20b)$$

The z component of the total active acoustic intensity is therefore

$$\begin{aligned} i_z(\mathbf{r}) &= \frac{1}{2} \Re(p(\mathbf{r}) \mathbf{v}_z^*(\mathbf{r})) \\ &= \frac{1}{2} \Re[\omega^2 W_{mn} P_{mn}(\mathbf{r}) + \mathbf{q}^T \mathbf{g}_s(\mathbf{r}) \\ &\quad \times (j\omega W_{mn}^* V_{mnz}^*(\mathbf{r}) + \mathbf{U}_s^H(\mathbf{r}) \mathbf{q}^*)]. \end{aligned}$$

Expanding this expression,

$$\begin{aligned} i_z(\mathbf{r}) &= \frac{1}{2} \Re[\mathbf{q}^T \mathbf{g}_s(\mathbf{r}) \mathbf{U}_s^H(\mathbf{r}) \mathbf{q}^*] \\ &\quad + \frac{1}{2} \Re[\mathbf{q}^T \mathbf{g}_s(\mathbf{r}) j\omega W_{mn}^* V_{mnz}^*(\mathbf{r})] \\ &\quad + \frac{1}{2} \Re[\omega^2 W_{mn} P_{mn}(\mathbf{r}) \mathbf{U}_s^H(\mathbf{r}) \mathbf{q}^*] + i_{zp}(\mathbf{r}). \end{aligned} \quad (21)$$

It is desirable to express the total active sound intensity in a Hermitian quadratic form similar to (17). This is possible by applying the relation $\Re(Z) = \frac{1}{2}(Z + Z^*)$ to the first three terms on the right-hand side of (21). After rearranging, this gives

$$\begin{aligned} i_z(\mathbf{r}) &= \frac{1}{4} \mathbf{q}^H [\mathbf{U}_s^*(\mathbf{r}) \mathbf{g}_s^T(\mathbf{r}) + \mathbf{g}_s^*(\mathbf{r}) \mathbf{U}_s^T(\mathbf{r})] \mathbf{q} \\ &\quad + \frac{1}{4} \mathbf{q}^H [\omega^2 W_{mn} P_{mn}(\mathbf{r}) \mathbf{U}_s^*(\mathbf{r}) - j\omega W_{mn} V_{mnz}(\mathbf{r}) \mathbf{g}_s^*(\mathbf{r})] \\ &\quad + \frac{1}{4} [\omega^2 W_{mn}^* P_{mn}^*(\mathbf{r}) \mathbf{U}_s^T(\mathbf{r}) + j\omega W_{mn}^* V_{mnz}^*(\mathbf{r}) \mathbf{g}_s^T(\mathbf{r})] \mathbf{q} \\ &\quad + i_{zp}(\mathbf{r}). \end{aligned}$$

After posing

$$\mathbf{a}(\mathbf{r}) = \frac{1}{4} [\mathbf{U}_s^*(\mathbf{r}) \mathbf{g}_s^T(\mathbf{r}) + \mathbf{g}_s^*(\mathbf{r}) \mathbf{U}_s^T(\mathbf{r})],$$

$$\mathbf{b}(\mathbf{r}) = \frac{1}{4} [\omega^2 W_{mn} P_{mn}(\mathbf{r}) \mathbf{U}_s^*(\mathbf{r}) - j\omega W_{mn} V_{mnz}(\mathbf{r}) \mathbf{g}_s^*(\mathbf{r})],$$

$$\mathbf{c}(\mathbf{r}) = i_{zp}(\mathbf{r}),$$

we obtain

$$i_z(\mathbf{r}) = \mathbf{q}^H \mathbf{a}(\mathbf{r}) \mathbf{q} + \mathbf{q}^H \mathbf{b}(\mathbf{r}) + \mathbf{b}^H(\mathbf{r}) \mathbf{q} + i_{zp}(\mathbf{r}). \quad (22)$$

In this case, the vector \mathbf{q} is adjusted to minimize the sum of the active sound intensities at the error sensor locations. The cost function is therefore

$$J = \sum_{k=1}^M i_z(\mathbf{r}_{mk}), \quad (23)$$

that is,

$$J = \mathbf{q}^H \mathbf{A}_i \mathbf{q} + \mathbf{q}^H \mathbf{b}_i + \mathbf{b}_i^H \mathbf{q} + c_i, \quad (24)$$

where

$$\mathbf{A}_i = \sum_{k=1}^M \mathbf{a}(\mathbf{r}_{mk}), \quad (25a)$$

$$\mathbf{b}_i = \sum_{k=1}^M \mathbf{b}(\mathbf{r}_{mk}), \quad (25b)$$

$$c_i = \sum_{k=1}^M i_{zp}(\mathbf{r}_{mk}). \quad (25c)$$

Provided the matrix \mathbf{A}_i is positive definite, J has a unique global minimum. The optimal control sources strengths are given by

$$\mathbf{q}_{\text{opt}} = -\mathbf{A}_i^{-1} \mathbf{b}_i,$$

and the minimized cost function is given by $J_{\text{opt}} = c_i + \mathbf{b}_i^H \mathbf{q}_{\text{opt}}$. It should be emphasized here that, contrary to the squared pressures minimization, the first term in (24), which represents the algebraic sum of the active intensities due to the control sources at the error sensors (in the absence of the primary vibration field), is not guaranteed to be positive; hence the matrix \mathbf{A}_i is not necessarily positive definite. In effect, the active intensity is a signed quantity, which can

locally take negative values indicating a net local acoustic energy flow in the direction of negative z . These sign fluctuations of the active intensities are especially important in the near field of the control-sources and at low frequencies. As a consequence, if the number of intensity sensors is small, the algebraic sum of the intensities measured at the error sensors may be negative. However, when the number of sensors and the extent of the error sensors plane become large, the first term in (24) should approach (to a multiplying constant) the sound power radiated from the control sources and should be positive. In the numerical simulations, the positivity of \mathbf{A}_i was checked by computing its eigenvalues; when all the eigenvalues of \mathbf{A}_i are real and positive, then \mathbf{A}_i is positive and the resulting extremum J_{opt} is a global minimum. Note that \mathbf{A}_i depends only on the control source arrangement, and not on the primary acoustic field.

Other potential difficulties may arise with the control of the acoustic intensity, even in cases where \mathbf{A}_i is positive. In contrast to squared pressures minimization, the cost function J for intensity minimization is not guaranteed to be positive; the minimization of J can result in a large and negative value of the minimum, which translates into poor or negative control performance. Such a situation may again occur when the number of intensity sensors is not sufficient. Therefore, it is safe to check if the cost function for intensity minimization has a positive or negative minimum; this is done by checking the sign of $J_{\text{opt}} = c_i + \mathbf{b}_i^H \mathbf{q}_{\text{opt}} = c_i - \mathbf{b}_i^H \mathbf{A}_i \mathbf{b}_i$, which depends on both the control source arrangement and the primary disturbance. Additionally, since the acoustic intensity is a signed quantity, a small value of the cost function after control does not necessarily imply a small value of the acoustic intensity at all error sensors. Some of the difficulties associated with active control of acoustic intensity have been discussed by Sommerfeldt and Nashif (1994).

E. Exact minimization of the sound power

The exact minimization of the total sound power radiated by the vibrating plate and the control sources is carried out hereafter as a reference solution, in order to compare the attenuations provided by the the two previous near-field control strategies to the maximum sound power attenuation achievable for a given control source configuration. To begin the superposition principle is applied to obtain the total sound pressure,

$$p(\mathbf{r}) = p_p(\mathbf{r}) + p_s(\mathbf{r}) = \omega^2 W_{mn} P_{mn}(\mathbf{r}) + \mathbf{q}^T \mathbf{g}_s(\mathbf{r}), \quad (26)$$

where $\mathbf{q} = [q_1, \dots, q_s, \dots, q_S]^T$ is the vector of the S control sources strengths, and $\mathbf{g}_s(\mathbf{r}) = [G(\mathbf{r}, \mathbf{r}_{s1}), \dots, G(\mathbf{r}, \mathbf{r}_{sS}), \dots, G(\mathbf{r}, \mathbf{r}_{sS})]^T$ is the vector of the sound pressures due to the individual control sources.

In the following, the sound power is evaluated from the far-field sound pressure. Considering a point \mathbf{r} in the far field of the plate, with spherical coordinates R, θ, ϕ $\{|R| \rightarrow \infty, \theta \in [0, \pi/2]$ is the angle between \mathbf{k} and \mathbf{r} , $\phi \in [0, 2\pi]$ is the angle between \mathbf{i} and the projection of \mathbf{r} in the (x, y) plane},

$$p(R, \theta, \phi) = \rho_0 \omega^2 W_{mn} \frac{e^{jkR}}{2\pi R} \check{\phi}_{mn}(\lambda, \mu) + \mathbf{q}^T \mathbf{g}_s(R, \theta, \phi), \quad (27)$$

where $\lambda = k \sin \theta \cos \phi$; $\mu = k \sin \theta \sin \phi$, and

$$\check{\phi}_{mn}(\lambda, \mu) = \frac{m \pi / l_x}{((m \pi / l_x)^2 - \lambda^2)} (1 - e^{-j\lambda l_x} \cos m \pi) \\ \times \frac{n \pi / l_y}{((n \pi / l_y)^2 - \mu^2)} (1 - e^{-j\mu l_y} \cos n \pi)$$

is the wavenumber transform of $\phi_{mn}(\mathbf{r}_0)$, representing the far-field directivity of the vibration mode m, n . Also,

$\mathbf{g}_s(R, \theta, \phi)$

$$= [G_{s1}(R, \theta, \phi), \dots, G_{ss}(R, \theta, \phi), \dots, G_{sS}(R, \theta, \phi)]^T$$

and $G_s(R, \theta, \phi) = (e^{jkR}/2\pi R) e^{-j\lambda x_s - j\mu y_s} \cos \delta z_s$ is the far-field directivity of the control source s , with $\delta = k \cos \theta$.

The total sound power is obtained by integrating the squared far-field sound pressure over a hemisphere of infinite radius $\Omega(|R| \rightarrow \infty, 0 \leq \theta \leq \pi/2, 0 \leq \phi \leq 2\pi)$:

$$\Pi = \frac{1}{2\rho_0 c_0} \int_{\Omega} p(R, \theta, \phi) p^*(R, \theta, \phi) R^2 |\sin \theta| d\theta d\phi. \quad (28)$$

Substituting (27) into (28),

$$\Pi = \frac{\rho_0 \omega^4}{8 \pi^2 c_0} \int_{\Omega} [\check{\phi}_{mn}(\lambda, \mu) + \mathbf{q}^T \mathbf{D}(\lambda, \mu)] \\ \times [\check{\phi}_{mn}^*(\lambda, \mu) + \mathbf{D}^H(\lambda, \mu) \mathbf{q}^*] |\sin \theta| d\theta d\phi, \quad (29)$$

where

$$\mathbf{D}(\lambda, \mu) = \frac{1}{\rho_0 \omega^2} [e^{-j\lambda x_{s1} - j\mu y_{s1}} \cos \delta z_{s1}, \dots, \\ e^{-j\lambda x_{ss} - j\mu y_{ss}} \cos \delta z_{ss}, \dots, \\ e^{-j\lambda x_{sS} - j\mu y_{sS}} \cos \delta z_{sS}]^T.$$

The total sound power can be expressed as an Hermitian quadratic form of the control variable \mathbf{q} ,

$$\Pi = \mathbf{q}^H \mathbf{A}_{\Pi} \mathbf{q} + \mathbf{q}^H \mathbf{b}_{\Pi} + \mathbf{b}_{\Pi}^H \mathbf{q} + c_{\Pi}, \quad (30)$$

where

$$\mathbf{A}_{\Pi} = \frac{\rho_0 \omega^4}{8 \pi^2 c_0} \int_{\Omega} \mathbf{D}^*(\lambda, \mu) \mathbf{D}^T(\lambda, \mu) |\sin \theta| d\theta d\phi,$$

$$\mathbf{b}_{\Pi} = \frac{\rho_0 \omega^4}{8 \pi^2 c_0} \int_{\Omega} \check{\phi}_{mn}(\lambda, \mu) \mathbf{D}^*(\lambda, \mu) |\sin \theta| d\theta d\phi,$$

and

$$c_{\Pi} = \frac{\rho_0 \omega^4}{8 \pi^2 c_0} \int_{\Omega} \check{\phi}_{mn}(\lambda, \mu) \check{\phi}_{mn}^*(\lambda, \mu) |\sin \theta| d\theta d\phi$$

is the primary sound power radiated by the plate in the absence of the control sources.

In the numerical simulations, the quantities \mathbf{A}_{Π} , \mathbf{b}_{Π} , c_{Π} were calculated using a numerical integration over the far-field directions θ, ϕ . In this case, the first term in (30), representing the sound power radiated by the control sources in the absence of the primary plate radiation, is a positive quantity; therefore \mathbf{A}_{Π} is a positive definite matrix, and the total sound power has a unique global minimum defined by

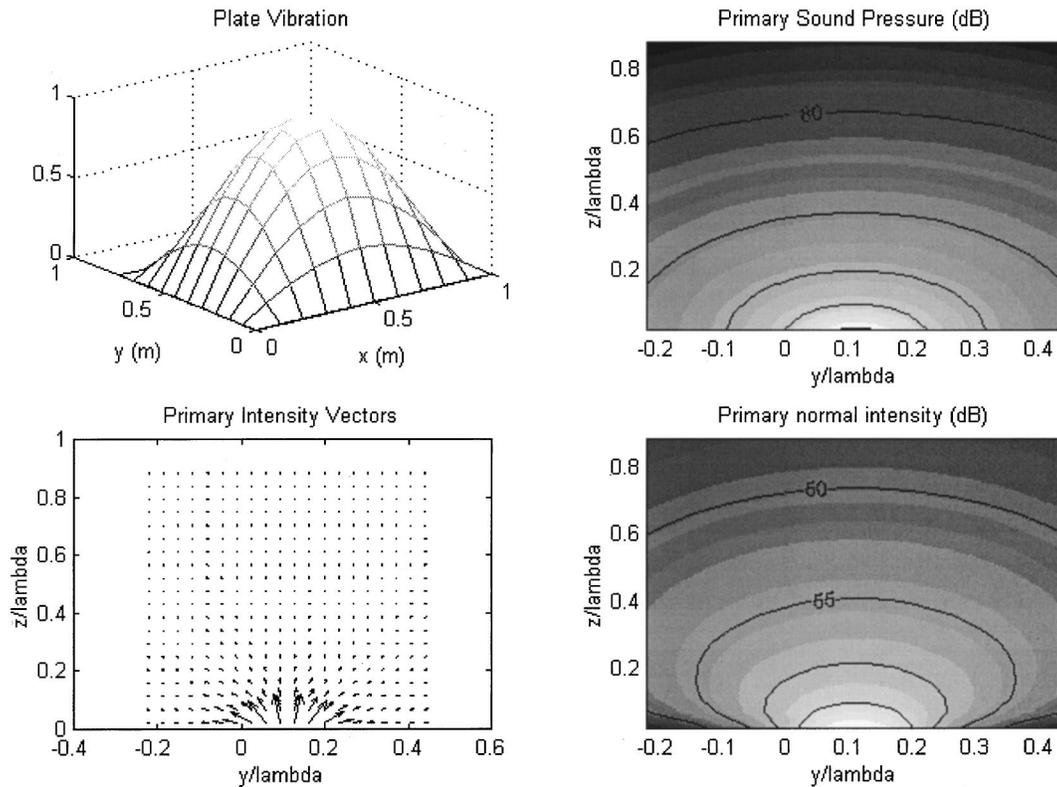


FIG. 2. Primary sound field in the plane $x=l_x/2$ for mode $(m,n)=(1,1)$, $W_{mn}=1$, $l_x=1$ m, $l_y=0.75$ m, and $f=100$ Hz.

$$\mathbf{q}_{\text{opt}} = -\mathbf{A}_{\text{II}}^{-1} \mathbf{b}_{\text{II}}, \quad (31a)$$

$$\Pi_{\text{opt}} = c_{\text{II}} + \mathbf{b}_{\text{II}}^H \mathbf{q}_{\text{opt}}. \quad (31b)$$

II. ACTIVE CONTROL SIMULATIONS

A. Nonefficient radiators

1. Case 1

A primary vibration field of the plate consisting of the $(m,n)=(1,1)$ mode driven at a unit modal displacement, $W_{mn}=1$, is considered first of all. The plate has dimensions $l_x=1$ m and $l_y=0.75$ m, and the excitation frequency is $f=100$ Hz. Figure 2 shows the primary sound field of the plate in the y - z plane passing through $x=l_x/2$. A total of 10×10 plate elements were used to calculate the sound field;

the sound radiation is calculated into air ($\rho_0=1.225$ kg/m³, $c_0=340$ m/s). In this figure, as well as in the next figures, the acoustic wavelength is used to scale the radiated acoustic field. For the $(1,1)$ mode and the frequency considered, the ratio of the acoustic wave number to the structural wave number is $k/k_{mn}=0.352 < 1$ [with $k_{mn} = \sqrt{(m\pi/l_x)^2 + (n\pi/l_y)^2}$], and the mode is therefore a poor radiator at this frequency, with almost omnidirectional radiation. Note also from Fig. 2 that the z component of the active acoustic intensity is always positive; this mode does not exhibit local negative energy flow.

Figures 3 and 4 show active control simulations based on squared pressure minimization and intensity minimization, respectively. In both cases the control source and error

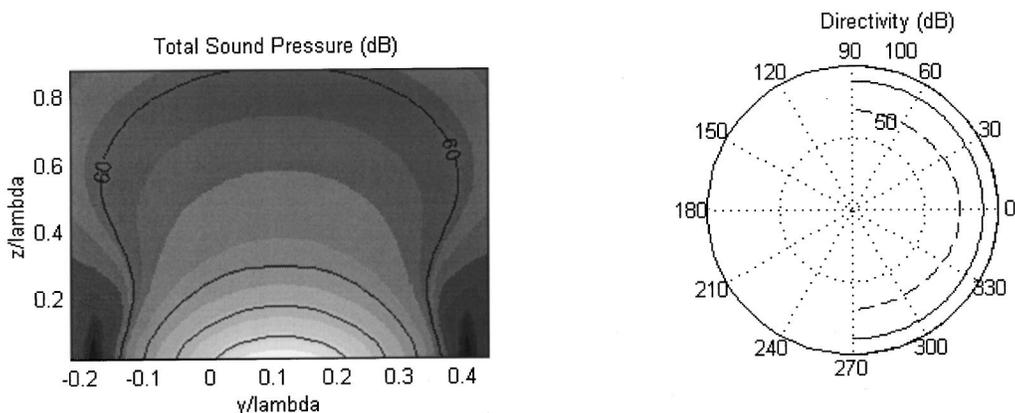


FIG. 3. Results of near-field intensity minimization, 2×2 control sources at $z_s=0.06\lambda$, 3×3 error sensors at $z_m=0.24\lambda$. The far-field directivity is in the y - z plane, —: before control, ---: after control.

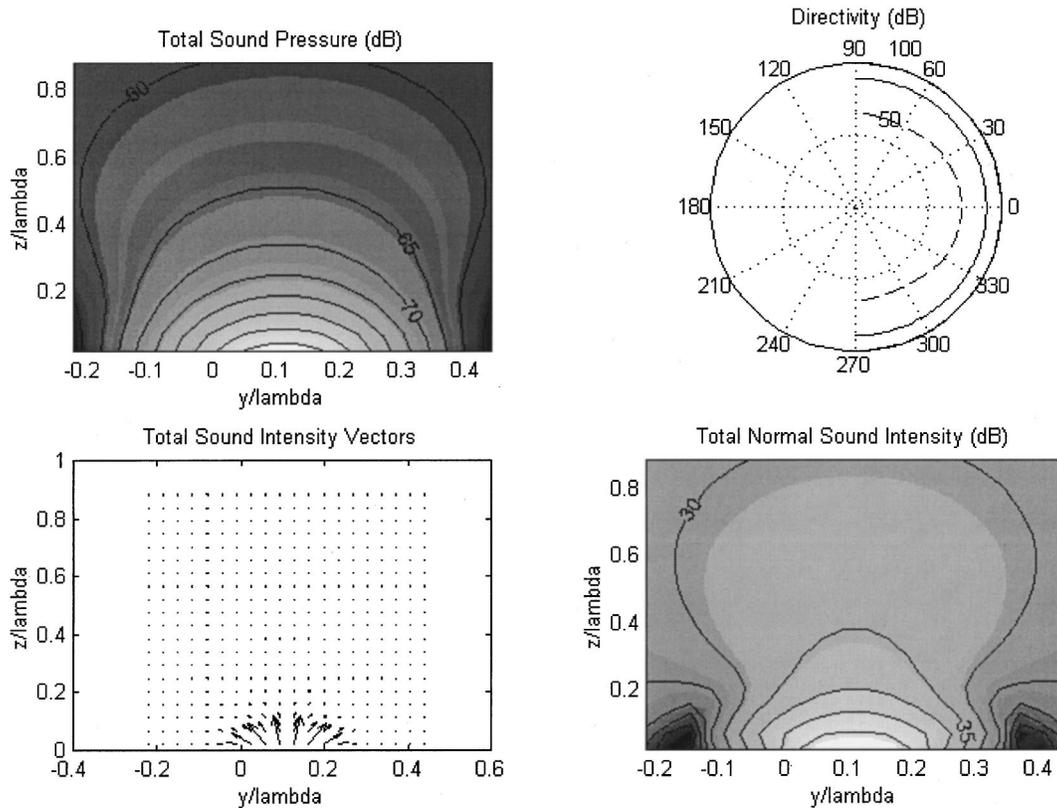


FIG. 4. Results of near-field intensity minimization, 2×2 control sources at $z_s = 0.06\lambda$, 3×3 error sensors at $z_m = 0.24\lambda$. The far-field directivity is in the y - z plane, —: before control, ---: after control.

sensor configurations are as follows: 2×2 control sources at height $z_s = 0.2$ m (corresponding to approximately $z_s = 0.06\lambda$, where λ is the acoustic wavelength at the frequency of interest) above the plate at positions $(x_s, y_s) = (0, 0); (0, l_y); (l_x, 0); (l_x, l_y)$; 3×3 error sensors at height $z_m = 0.8$ m ($z_m = 0.24\lambda$) above the plate, regularly positioned in a plane extending from $(x_m, y_m) = (-l_x/2, -l_y/2)$ to $(x_m, y_m) = (3l_x/2, 3l_y/2)$ (that is, a plane covering four times the area of the radiating plate). Such an arrangement satisfies the criterion of at least three control sources and three error sensors per acoustic wavelength. In this configuration, as well as in all cases reported in this section, it was verified that the matrix \mathbf{A}_i has all its eigenvalues real and positive, such that \mathbf{A}_i is positive. The optimal sound power attenuation calculated from (31) for this control source configuration is 21.2 dB; the minimization of the squared pressures at the error sensors gives a sound power attenuation of 17.9 dB, while the minimization of the active intensities gives a slightly better, almost optimal sound power attenuation of 19.7 dB. Note that the sound pressure in the near-field of the plate after minimizing the intensity (Fig. 4) is slightly larger than the sound pressure after minimizing the squared pressure (Fig. 3), even though the radiated sound power is smaller after minimizing the intensity. The far-field directivity of the plate in the y - z plane was plotted in Figs. 3 and 4, before and after control, based on (27). These directivity plots show that the two near-field control strategies yield almost uniform far-field attenuation in all directions. On the other hand, the intensity vector plots before and after minimization of the intensity (Figs. 2 and 4) show that the inten-

sity is confined to the near field of the plate after control, and is globally reduced downstream of the sensors plane, resulting in a large attenuation of the sound power.

The same control source and error sensor arrangement as used previously was again considered, with the error sensor plane being moved away from the control source plane, at a distance varying between 0 and $\lambda/2$. Figure 5 shows the sound power attenuation obtained for the two control strategies as a function of the distance (scaled to the acoustic

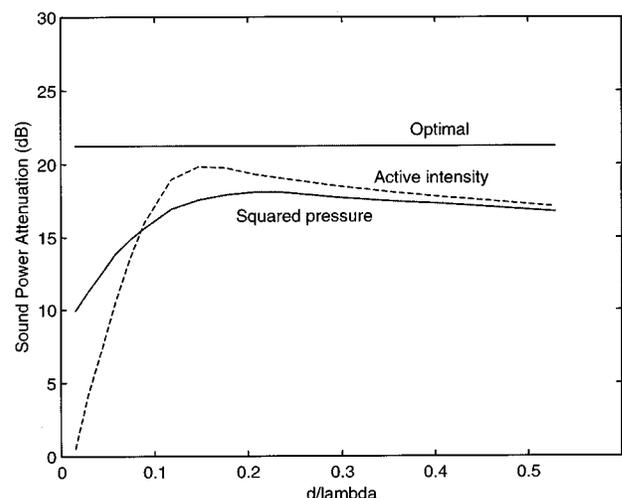


FIG. 5. Sound power attenuation as a function of the distance between the control sources plane and the error sensors plane, 2×2 control sources at $z_s = 0.06\lambda$, 3×3 error sensors, —: squared pressure minimization, ---: intensity minimization.

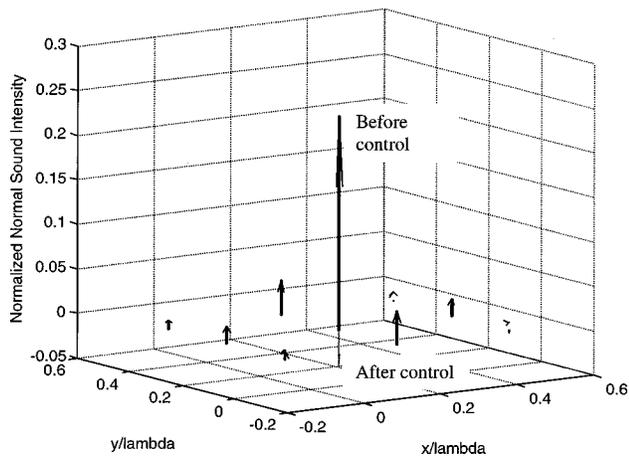


FIG. 6. Sound intensity at the error sensors before and after minimization of the intensity, 2×2 control sources at $z_s = 0.06\lambda$, 3×3 error sensors at $z_m = 0.12\lambda$.

wavelength) between the control source plane and the error sensor plane. The optimal attenuation of 21.2 dB, obtained from the minimization of the sound power, is reported on the figure as a reference value. A number of interesting features can be noted on this figure. First, intensity minimization yields slightly better control performance than squared pressure minimization for all of the distances considered except in the near-field of the control sources (below approximately $\lambda/10$); the difference, however, remains small and would not justify the practical complications involved in active intensity sensing, as compared to sound pressure sensing. The best control performance is achieved when minimizing the intensity, with the error sensor plane at 0.15λ from the control sources. At large distances, the two strategies give similar results; this is to be expected since the active intensity becomes proportional to the squared pressure at large distances from the source, and intensity minimization then becomes equivalent to squared pressure minimization. It should also be noted that squared pressure minimization gives a maximum sound power attenuation for a certain distance between the control sources plane and the error sensor plane (in this case 0.24λ). At larger distances, the control performance decreases because of the decreasing aperture of the error sensors array as seen from the plate; at smaller distances, there is also a rapid decrease of the control performance, because the control then aims at minimizing near-field components which are not radiated in the far field. The strategy of minimizing the active intensity instead of the squared pressure does not completely alleviate this difficulty; intensity minimization, like squared pressure minimization, suffers from a rapid decrease of the control performance when the sensors are in the near field of the sources. This is related to the signed nature of the cost function in the case of intensity minimization; the degradation of the control performance in the near field coincides with a negative and possibly large value of the sum of the intensities after control. As an example, Fig. 6 shows the z component of the active sound intensity at the error sensors before and after minimization of the intensity, for the same control source and error sensor arrangement as used previously (the separation between the

control source plane and the error sensor plane was 0.06λ). Even though the error sensors are located downstream the control sources, the control clearly creates negative sound intensity at some sensor locations, resulting in a nonoptimal control performance. The value of the optimized, intensity-based cost function $J_{\text{opt}} = c_i - \mathbf{b}_i^H \mathbf{A}_i \mathbf{b}_i$ was found to be negative in this configuration. The same observation was made for control sources–error sensors separation smaller than 0.1λ . Such inconvenience can be compensated by adding more intensity sensors, at the expense of a more complex system.

2. Case 2

We now consider a primary vibration field of the plate consisting of the $(m,n) = (1,3)$ mode driven at a unit modal displacement, $W_{mn} = 1$. The plate has dimensions $l_x = 4$ m and $l_y = 3$ m, and the frequency is $f = 100$ Hz. These data more accurately reflect the practical case of a transformer sidewall, at the fundamental frequency of excitation of the structure. Figure 7 shows the primary sound field of the plate in the y - z plane passing through $x = l_x/2$. For the $(1,3)$ mode and the frequency considered, the ratio of the acoustic wave number to the structural wave number is $k/k_{mn} = 0.57 < 1$, and the mode is therefore a nonefficient radiator at this frequency. Figure 7 shows that the acoustic near field is more complex than in the previous case, with strong spatial fluctuations of the sound pressure and active sound intensity (in some areas close to the plate, not shown on this figure, the sound intensity can be negative).

Figures 8 and 9 show active control results after minimization of the near-field squared pressures and near-field intensity, respectively. A total of 4×3 control sources at height $z_s = 0.2$ m ($z_s = 0.06\lambda$) above the plate surface were used, regularly positioned in a plane extending from $(x_s, y_s) = (0,0)$ to $(x_s, y_s) = (l_x, l_y)$ (a plane covering the area of the vibrating plate). Also, there are 8×6 error sensors at height $z_m = 0.8$ m ($z_m = 0.24\lambda$) above plate surface, regularly positioned in a plane extending from $(x_m, y_m) = (-l_x/2, -l_y/2)$ to $(x_m, y_m) = (3l_x/2, 3l_y/2)$ (that is, a plane covering four times the area of the radiating plate). Such an arrangement ensures that approximately two control sources and two error sensors are used per acoustic wavelength. The optimal sound power attenuation calculated from (31) for this control source configuration is 22.2 dB; the minimization of the squared pressures at the error sensors gives a sound power attenuation of 20.3 dB, while the minimization of the active intensities gives a significantly smaller sound power attenuation of 13.4 dB. In this case the poor performance of intensity control is due to the fact that the sum of the intensities at the error sensors is driven to a negative value after control. Note that the near-field sound pressure is significantly different when minimizing the near-field pressure or near-field intensity.

Figure 10 shows the sound power attenuation obtained for the two strategies as a function of the distance between the control source plane and the error sensor plane, for the same configuration of control sources and error sensors as previously. The optimal attenuation of 22.2 dB, obtained from the minimization of the sound power, is reported on the

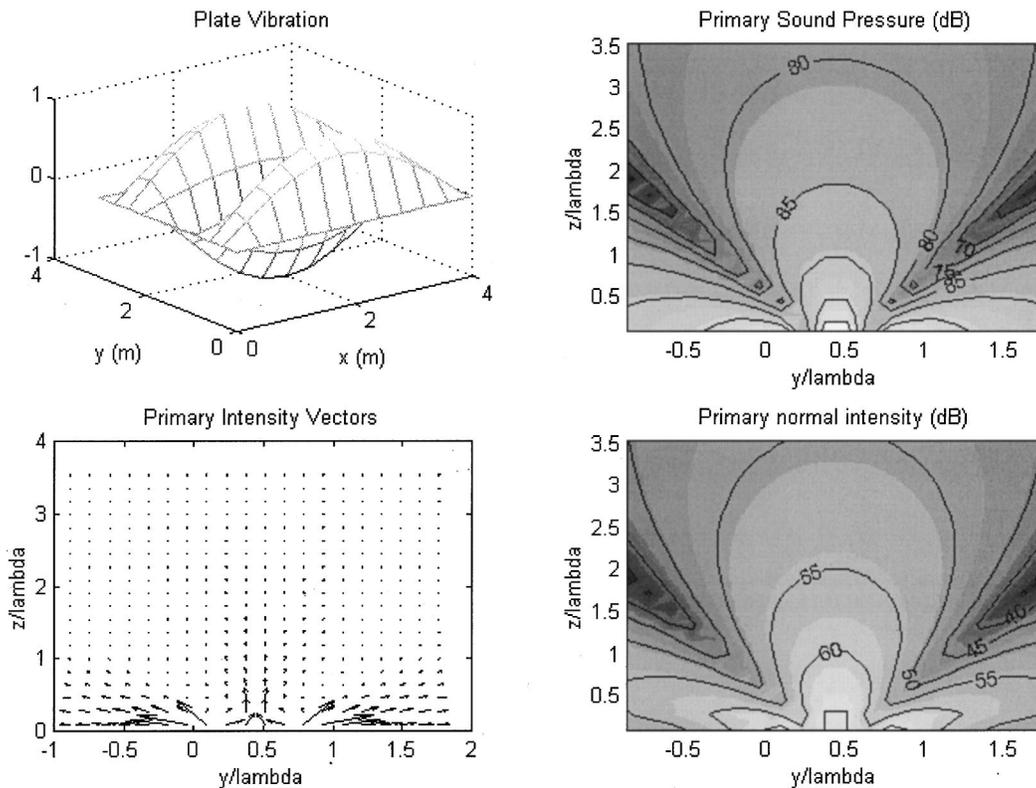


FIG. 7. Primary sound field in the plane $x=l_x/2$ for mode $(m,n)=(1,3)$, $W_{mn}=1$, $l_x=4$ m, $l_y=3$ m, and $f=100$ Hz.

figure as a reference value. In this case, intensity minimization provides lower attenuation than squared pressure minimization for all the distances considered, with similar attenuations obtained in the far field. The best control performance (20.3 dB sound power attenuation) is achieved when minimizing the squared pressure, with the error sensor plane at 0.18λ from the control sources. Figure 11 shows the z component of the active sound intensity at the error sensors before and after minimization of the intensity, for the same control source and error sensor arrangement as used previously (the separation between the control source plane and the error sensor plane was 0.06λ). The poor performance of intensity control in this configuration is again attributed to the fact that the intensity (and hence the sum of the intensities) at the error sensors is driven to a negative and possibly

large value after control. The value of the optimized, intensity-based cost function $J_{\text{opt}} = c_i - \mathbf{b}_i^H \mathbf{A}_i \mathbf{b}_i$ was found to be negative in this configuration. The same observation was made for control sources–error sensors separation smaller than 0.3λ . A possible solution to this problem is to add more error sensors, in order to observe a quantity more closely approximating the radiated sound power, hence a positive quantity. Figure 12 shows the sound power attenuation obtained for the same control source arrangement as previously, and a control sources–error sensors separation of 0.24λ , with now 24×18 error sensors in the error sensors plane. With more error sensors, intensity minimization performs better than squared pressure minimization in the near field. The best control performance (19.8 dB sound power

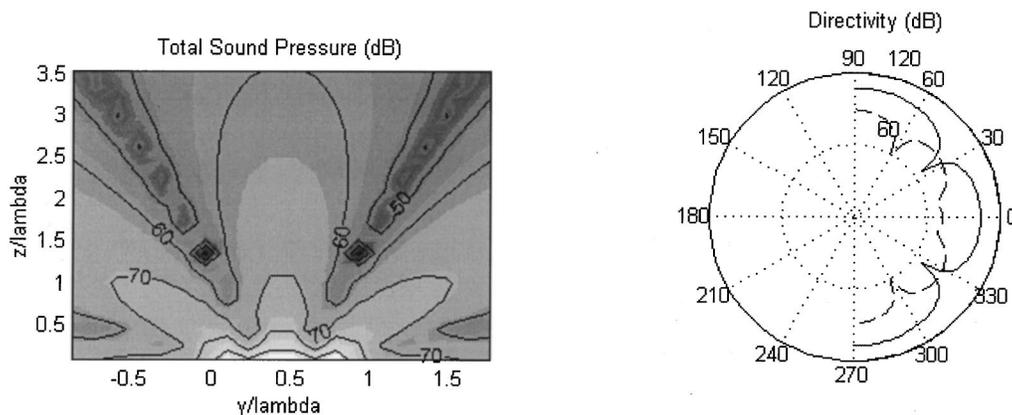


FIG. 8. Results of near-field squared pressure minimization, 4×3 control sources at $z_c=0.06\lambda$, 8×6 error sensors at $z_m=0.24\lambda$. The far-field directivity is in the y - z plane, —: before control, ---: after control.

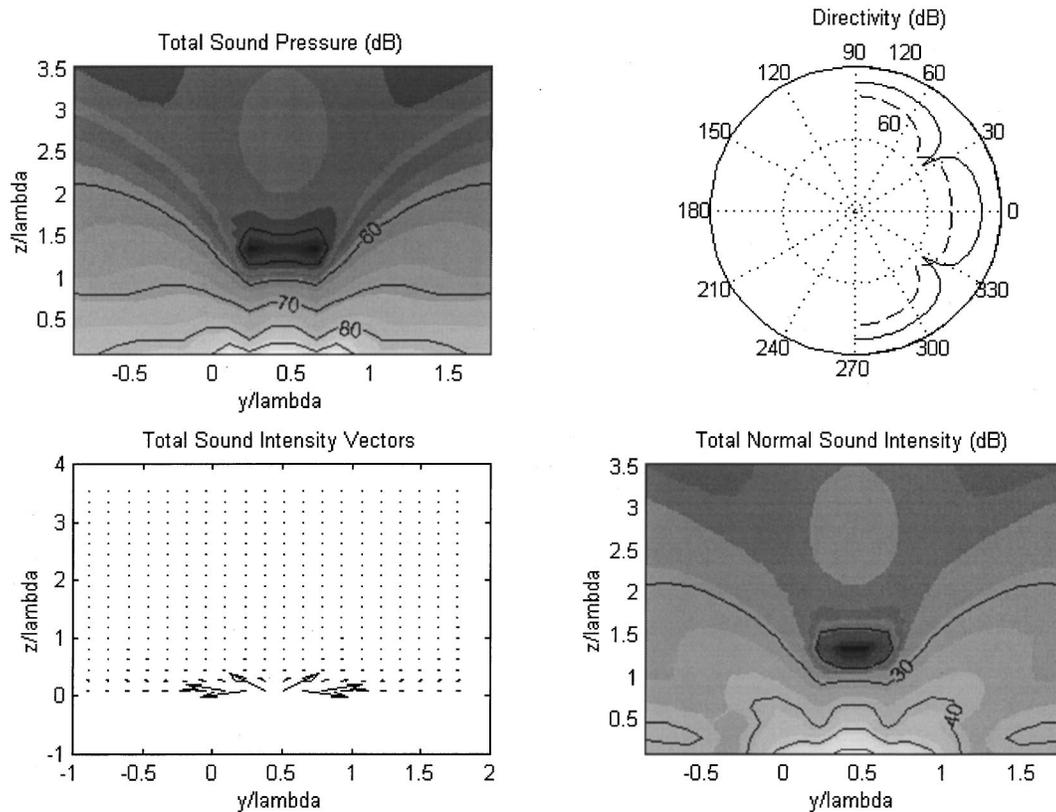


FIG. 9. Results of near-field intensity minimization, 4×3 control sources at $z_s = 0.06\lambda$, 8×6 error sensors at $z_m = 0.24\lambda$. The far-field directivity is in the y - z plane, —: before control, ---: after control.

attenuation) is achieved when minimizing the intensity, with the error sensor plane at 0.06λ from the control sources. In this configuration, the intensity-based cost function $J_{\text{opt}} = c_i - \mathbf{b}_i^H \mathbf{A}_i \mathbf{b}_i$ was found to be positive after minimization. A potential advantage of intensity minimization is thus the possibility of placing the error sensors very close to the primary source and control sources, but at the expense of a larger number of sensors required. The reduction obtained, however, is smaller than the maximum sound power attenuation

of 20.3 dB obtained when minimizing the squared pressure with only 8×6 error sensors at 0.18λ from the control sources. Figure 13 shows the sound power attenuation as a function of the density of the error sensor array, for the same control source arrangement as previously, and a control source–error sensors separation of 0.18λ . In this case, the extent of the error sensor plane is kept fixed (four times the area of the radiating plate), and the number of error sensors per acoustic wavelength is variable. An error sensor density of three to four sensors per acoustic wavelength is found to be necessary in order to obtain a good control performance for the control source arrangement considered. As far as the extent of the error-sensor plane is concerned, a rapid degradation of the control performance is observed (for both intensity minimization and squared pressure minimization) when the error sensor plane covers an area smaller than the primary radiator.

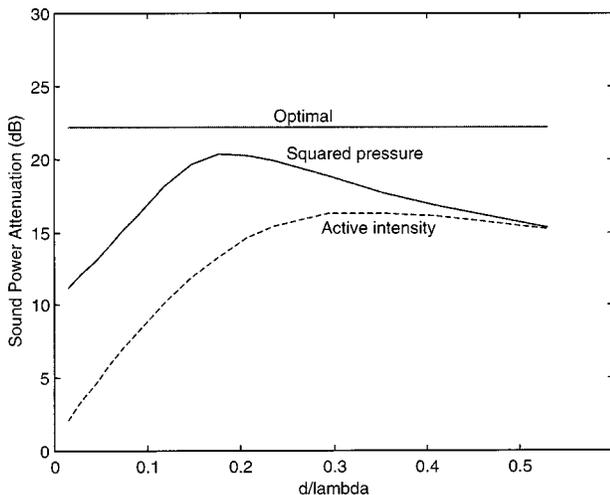


FIG. 10. Sound power attenuation as a function of the distance between the control sources plane and the error sensors plane, 4×3 control sources at $z_s = 0.06\lambda$, 8×6 error sensors, —: squared pressure minimization, ---: intensity minimization.

B. Efficient radiator

We now consider a similar radiator as previously (dimensions $l_x = 4$ m, $l_y = 3$ m), with a primary vibration field of the plate consisting of the $(m, n) = (1, 3)$ mode driven at a unit modal displacement, $W_{mn} = 1$, and at a frequency $f = 300$ Hz. In contrast to the previous case, the ratio of the acoustic wave number to the structural wave number is now $k/k_{mn} = 1.71 > 1$, and the mode is therefore an efficient radiator at this frequency. Figure 14 shows the primary sound field of the plate in the y - z plane passing through $x = l_x/2$. The acoustic field in this case reveals a clear energy flow

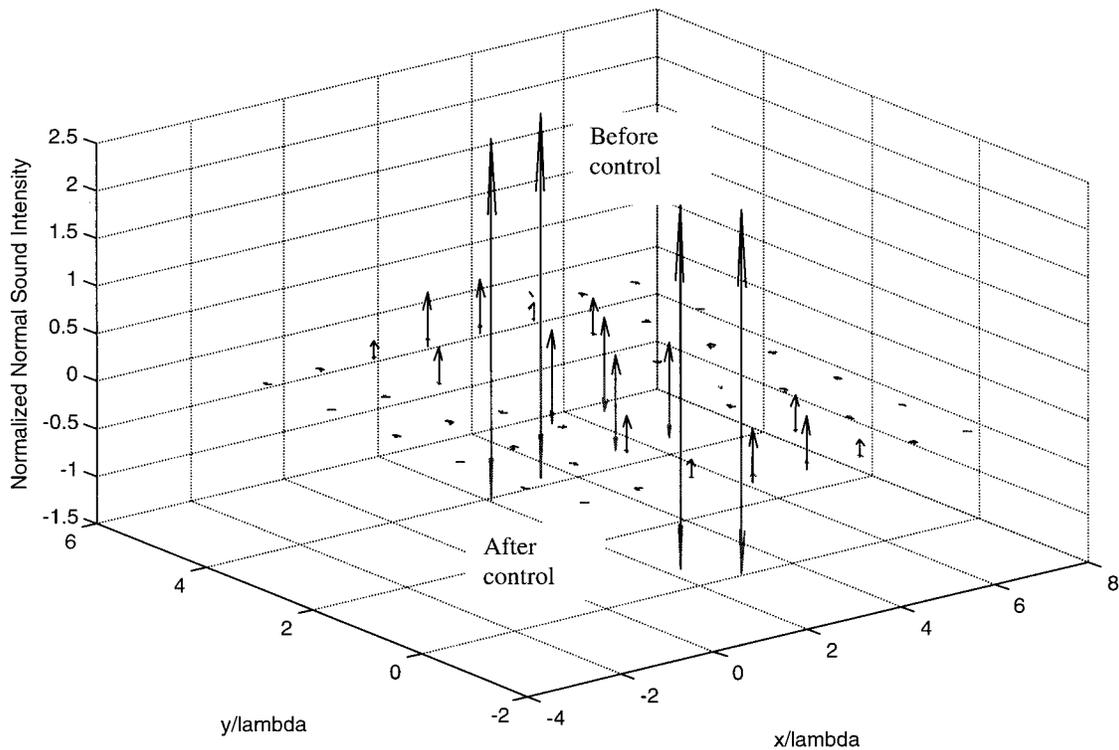


FIG. 11. Sound intensity at the error sensors before and after minimization of the intensity, 4×3 control sources at $z_s = 0.06\lambda$, 8×6 error sensors at $z_m = 0.12\lambda$.

from the radiator to the acoustic medium, with a strong directivity of the sound radiation and less local circulation of the acoustic intensity in the near field than in the nonefficient radiator case. The following simulations investigate whether active control of intensity outperforms active control of squared pressure in this case.

Figures 15 and 16 show active control results after minimization of the near-field squared pressures and near-field intensity, respectively. A total of 8×6 control sources at height $z_s = 0.2$ m ($z_s = 0.18\lambda$) above the plate surface were

used, regularly positioned in a plane extending from $(x_s, y_s) = (0, 0)$ to $(x_s, y_s) = (l_x, l_y)$. Also, there are 16×12 error sensors at height $z_m = 0.8$ m ($z_m = 0.71\lambda$) above plate surface, regularly positioned in a plane extending from $(x_m, y_m) = (-l_x/2, -l_y/2)$ to $(x_m, y_m) = (3l_x/2, 3l_y/2)$. The optimal sound power attenuation calculated from (31) for this control source configuration is 26.0 dB; the minimization of the squared pressures at the error sensors gives a sound power attenuation of 20.6 dB, while the minimization of the active intensities gives a sound power attenuation of

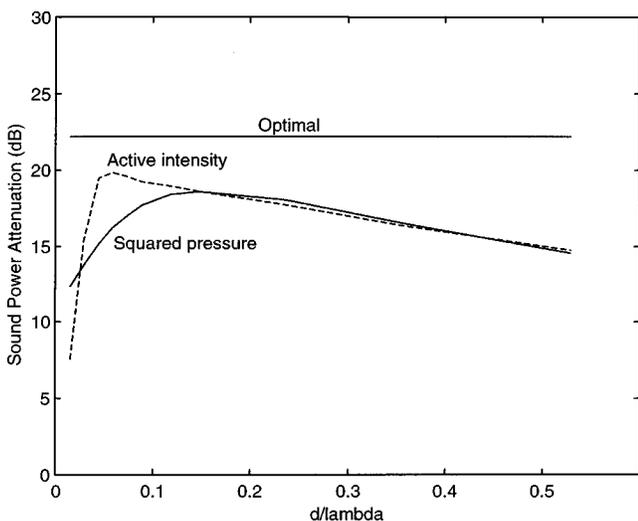


FIG. 12. Sound power attenuation as a function of the distance between the control sources plane and the error sensors plane, 4×3 control sources at $z_s = 0.06\lambda$, 24×18 error sensors, —: squared pressure minimization, ---: intensity minimization.

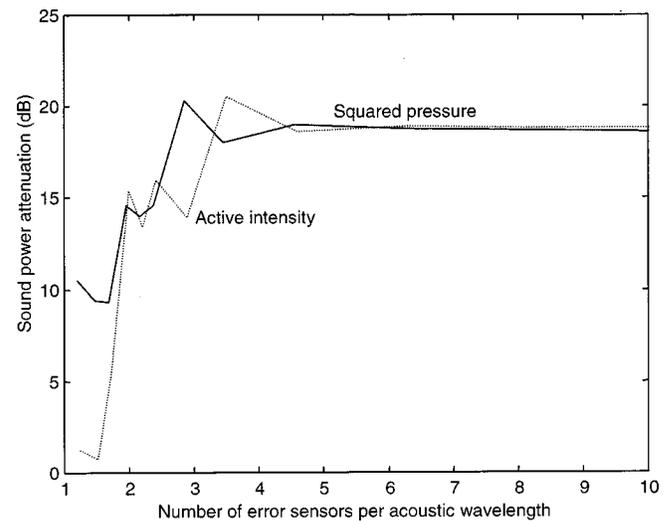


FIG. 13. Sound power attenuation as a function of the density of the error sensor array, 4×3 control sources at $z_s = 0.06\lambda$, 8×6 error sensors at $z_m = 0.24\lambda$, —: squared pressure minimization, ---: intensity minimization.

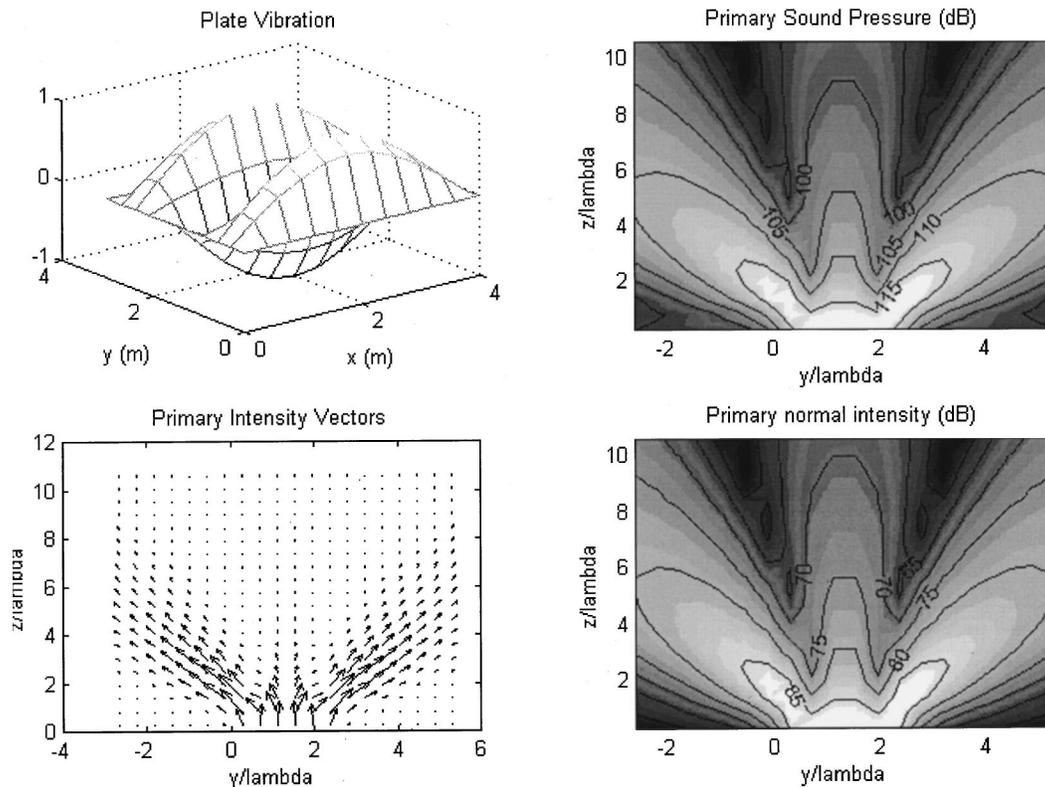


FIG. 14. Primary sound field in the plane $x=l_x/2$ for mode $(m,n)=(1,3)$, $W_{mn}=1$, $l_x=4$ m, $l_y=3$ m, and $f=300$ Hz.

19.4 dB. Again, intensity minimization does not outperform squared pressure minimization in this case, even though the primary intensity field shows a clear energy flow away from the plate. Similar observations were made for control sources—error sensors separation smaller than 0.3λ , for which the sum of the intensities was found to be negative after control.

III. CONCLUSION

The strategy of sensing and minimizing the near-field active intensity in the active control of free-field radiation from a vibrating plate has been investigated. The optimal minimization of the sum of the near-field, normal active sound intensity using secondary acoustic sources has been

derived for this problem, and the results obtained have been compared to the minimization of the sum of the near-field squared pressure. The strategy of minimizing the near-field sound intensity suffers from two main limitations, which are related to the fact that sound intensity is a signed quantity: (1) precautions need to be taken when deriving the optimal control source strengths in the minimization of the sound intensity, as the quadratic term in the associated cost function is not necessarily positive; (2) additionally, the cost function itself can take negative values after minimization, possibly resulting in poor control performance. The numerical simulations show that near-field sound pressure minimization offers decreasing performance when the sensors are very close to the primary source. However, near-field sound

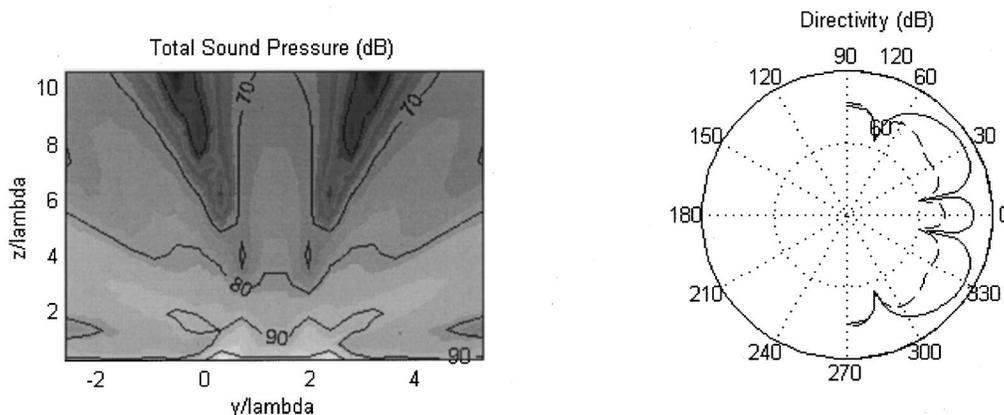


FIG. 15. Results of near-field squared pressure minimization, 4×3 control sources at $z_s=0.18\lambda$, 8×6 error sensors at $z_m=0.71\lambda$. The far-field directivity is in the y - z plane, —: before control, ---: after control.

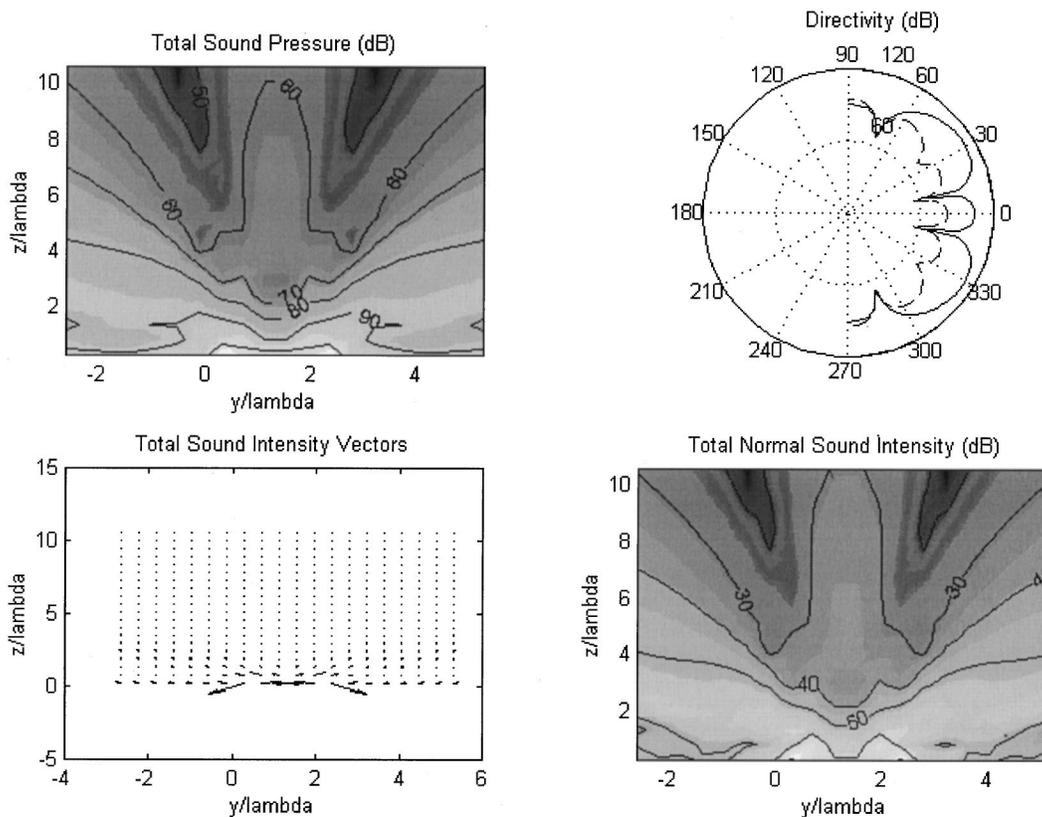


FIG. 16. Results of near-field intensity minimization, 4×3 control sources at $z_s = 0.18\lambda$, 8×6 error sensors at $z_m = 0.71\lambda$. The far-field directivity is in the y - z plane, —: before control, ---: after control.

intensity minimization does not, in general, provide significant improvements as compared to near-field squared pressure minimization, and may in some cases give smaller attenuations than pressure minimization. This is due to the fact that, when the intensity sensors are in the near field of the primary source, the sum of the intensities can be driven to large and negative values after control. A potential advantage of intensity minimization is the possibility of placing the error sensors very close to the primary source and control sources, but at the expense of a larger number of sensors required.

Angevine, O. L. (1993). "Active cancellation of hum of large power transformers: the real world," in *Proceedings of the Second Conference on Recent Advances in Active Control of Sound and Vibration*, pp. 279–290.
 Bolton, J. S., Gardner, B. K., and Beauvilain, T. A. (1995). "Sound cancellation by the use of secondary multipoles," *J. Acoust. Soc. Am.* **98**, 2343–2362.
 Cazzolato, B. S. (1998). "Sensing systems for active control of sound transmission into cavities," Submitted for Ph.D. degree, The University of Adelaide.

Hansen, C. H., and Snyder, S. D. (1997). *Active Control of Noise and Vibration* (E&FN Spon, London).
 Kang, S. W., and Kim, Y. H. (1997). "Active intensity control for the reduction of radiated duct noise," *J. Sound Vib.* **201**, 595–611.
 Kempton, A. J. (1976). "The ambiguity of acoustic sources: a possibility for active control?" *J. Sound Vib.* **48**, 475–483.
 Nelson, P. A., Curtis, A. R. D., Elliott, S. J., and Bullmore, A. J. (1987). "The minimum power output of free field point source and active control of sound," *J. Sound Vib.* **116**, 397–414.
 Nelson, P. A., and Elliot, S. J. (1992). *Active Control of Sound* (Academic, London).
 Park, Y. C., and Sommerfeldt, S. D. (1997). "Global attenuation of broadband noise fields using energy density control," *J. Acoust. Soc. Am.* **101**, 350–359.
 Qiu, X., Hansen, C. H., and Li, X. (1998). "A comparison of near-field acoustic error sensing strategies for the active control of harmonic free field sound radiation," *J. Sound Vib.* **215**, 81–103.
 Sommerfeldt, S. D., and Nashif, P. J. (1994). "An adaptive filtered-x algorithm for energy-base active control," *J. Acoust. Soc. Am.* **96**, 300–306.
 Sommerfeldt, S. D., Parkins, J., and Park, Y. (1995). "Global active noise control in rectangular enclosures," in *Proceedings of Active 95*, pp. 477–488.
 Zander, A. C., and Hansen, C. H. (1993). "A comparison of error sensing strategies for the active control of duct noise," *J. Acoust. Soc. Am.* **94**, 841–848.

Experimental demonstration of a band-limited actuator/sensor selection strategy for structural acoustic control

Robert L. Clark

Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708-0300

David E. Cox

NASA Langley Research Center, Hampton, Virginia 23681

(Received 25 November 1998; accepted for publication 14 September 1999)

A band-limited method of selecting actuators and sensors for structural acoustic control is reviewed, and experimental results are presented to demonstrate the approach. The selection methodology is based upon the decomposition of the Hankel singular values of a system model in terms of individual sensor and actuator configurations for lightly damped structures. The technique selects sensor and actuator combinations which couple well to structural modes that radiate efficiently. However, it rejects sensor and actuator combinations which couple well to modes that are inefficient acoustic radiators or are outside of the desired bandwidth of control. Selecting transducer combinations which filter modes outside of the desired bandwidth serves to minimize the potential for spillover and instability associated with unmodeled or poorly modeled dynamics. The approach is computationally efficient since it is based upon open-loop dynamics and does not require iterative nonlinear optimization. © 1999 Acoustical Society of America. [S0001-4966(99)06812-5]

PACS numbers: 43.40.Vn [PJR]

INTRODUCTION

Various approaches for optimizing the location of actuators and sensors for structural acoustic control have been proposed in the past (Snyder and Hansen, 1990; Wang *et al.*, 1991; Clark and Fuller, 1991, 1992a; Heck and Naghshineh, 1994; Ruckman and Fuller, 1995). The ultimate goal of such studies has been aimed at a reduction in the complexity associated with the compensator in favor of wave number filtering through the distributed actuator and/or sensor (Fuller and Burdisso, 1991; Clark and Fuller, 1992b; Clark *et al.*, 1993; Snyder *et al.*, 1993; Gu *et al.*, 1994; Fuller *et al.*, 1996; Johnson and Elliott, 1995; Maillarad and Fuller, 1998; Charette *et al.*, 1998). However, for many of these studies, the emphasis was placed upon adaptive feedforward control. For such control, the focus can be limited to spatial apertures which couple well to the performance path over the bandwidth of interest. Additionally, spillover into modes outside of the bandwidth of interest is of little concern since control effort is extended only over the bandwidth of the reference signal. However, for feedback control, a necessity for structural acoustic systems driven by exogenous inputs such as turbulent boundary layer noise, spillover is a primary concern (Thomas and Nelson, 1995; Clark and Frampton, 1997, 1998; Frampton and Clark, 1997).

Lim (1997) recently proposed a methodology for selecting actuator and sensor placement for disturbance rejection based upon the Hankel singular values of lightly damped structures. Compared to previous methodologies which emphasized the use of iterative, nonlinear optimization techniques (Wang *et al.*, 1991; Clark and Fuller, 1992a; Ruckman and Fuller, 1995; Smith and Clark, 1998; Smith *et al.*, 1998), the approach outlined by Lim (1997) is computationally efficient. For lightly damped structures, the Hankel sin-

gular values can be estimated from the modal properties of the discrete time system models. Thus, the participation of each structural mode in the control path (actuator to sensor) can be weighted by the participation of this mode on the disturbance to performance path. While this method is efficient and computationally direct (i.e., no iteration), it does lack a robustness metric.

For experimental implementation of feedback control, we often seek to design a controller for a reduced-order model of the system. Due to practical constraints imposed by finite bandwidths of digital signal processors, an accurate model of the structure is obtained over a limited bandwidth, and the compensator is designed to "roll-off" at frequencies outside of the bandwidth of interest. However, this roll-off and the application of low-pass filters to attenuate the coupling of higher-order modes complicates the design, leading to stability issues which limit closed-loop performance. In a recent study by Clark and Cox (1998), a methodology for band-limited actuator/sensor selection for disturbance rejection was proposed, extending the work of Lim (1997).

In this method, actuator/sensor pairs are selected based upon a combined metric which provides a tradeoff between the coupling to structural modes related to performance within the bandwidth of interest, and the *lack of coupling* to structural modes outside of the identified bandwidth of interest. An analytical example for structural acoustic control was provided by Clark and Cox (1998), and the complementary experimental demonstration is the subject of this work.

A brief review of the actuator/sensor selection methodology is provided in the subsequent section. Upon outlining the selection methodology, the experimental system is described including the optimal piezoceramic actuator and sensor for structural acoustic control of a panel radiating into a half-space. H_2 design methodologies were applied (Clark

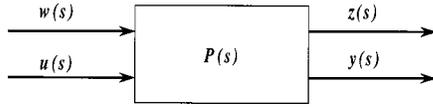


FIG. 1. Block diagram of generalized plant.

et al., 1998), and the experimental results are compared to the predicted results. The experimental results demonstrate that the band-limited transducer selection methodology for disturbance rejection provides the structural acoustic control system engineer with a powerful and simple technique for selecting transducers from analytical or experimental models of systems.

I. BAND-LIMITED TRANSDUCER SELECTION FOR DISTURBANCE REJECTION

In structural acoustic control we seek to reduce sound radiation from a structure (Fuller, 1988). Under such constraints, previous research has demonstrated that only some of the structural modes are efficient acoustic radiators within the low-frequency bandwidth targeted for such application (Wallace, 1972a, b; Lomas and Hayek, 1977). As such we typically modify models of structures to incorporate radiation filters which provide, as a performance output, a measure of sound power radiated (Baumann *et al.*, 1991; Cunnefare, 1991; Gibbs *et al.*, 1998). Such filters, when augmented to a structural model, yield what is frequently termed a generalized plant in control terminology, $P(s)$. A block diagram of the generalized plant is presented in Fig. 1. As illustrated, the plant is subjected to disturbance inputs, $w(s)$, and control inputs, $u(s)$. The outputs of the generalized plant are the performance variables, $z(s)$, and the sensed or measured variables, $y(s)$. The corresponding input–output equations can be expressed as follows (Clark *et al.*, 1998):

$$\begin{bmatrix} \mathbf{z}(s) \\ \mathbf{y}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{zw}(s) & \mathbf{P}_{zu}(s) \\ \mathbf{P}_{yw}(s) & \mathbf{P}_{yu}(s) \end{bmatrix} \begin{bmatrix} \mathbf{w}(s) \\ \mathbf{u}(s) \end{bmatrix}. \quad (1)$$

The performance path is defined by the transfer matrix existing between the exogenous input disturbances and the performance outputs, $P_{zw}(s)$. The control path is defined by the transfer matrix existing between the control inputs and the measured or sensed outputs, $P_{yu}(s)$. In structural acoustic control, the performance path is defined between the disturbances which serve to excite the structure and the radiation filters used to estimate the sound power radiated.

Lim (1997) proposed a method of determining the optimal subset of actuators and sensors for control based upon a weighted measure of the modal participation of each sensor/actuator combination. To determine the transducer pairs best suited for control, the Hankel singular values of the system through the performance path and the control path were determined. Lim and Gawronski (1996) noted that the Hankel singular values of a discrete time model of a lightly damped structure can be efficiently computed from a diagonalized form of the state space model of the system, denoted by the quadruple (A, B, C, D) . As noted by Lim and Gawronski (1996), a similarity transform can be constructed such that the transformed system matrix is block diagonal:

$$\tilde{A} = \begin{pmatrix} \tilde{A}_1 & 0 & \dots & 0 \\ 0 & \tilde{A}_2 & & \\ \vdots & \cdot & \ddots & \\ 0 & \dots & & \tilde{A}_n \end{pmatrix}, \quad (2)$$

where

$$\tilde{A}_i = \begin{bmatrix} \text{Re}(z_i) & -\text{Im}(z_i) \\ \text{Im}(z_i) & \text{Re}(z_i) \end{bmatrix} \quad (3)$$

and

$$z_i = \exp(\delta_i + j\psi_i)T \quad (4)$$

is the i th discrete eigenvalue of A with corresponding eigenvector, v_i . (Note that $\delta_i \pm j\psi_i$ are the eigenvalues of the corresponding continuous time system.) The similarity transform is defined by the transformation matrix,

$$V = [r_1, \dots, r_n], \quad (5)$$

where $r_i = [\text{Re}(v_i) - \text{Im}(v_i)]$. The resulting transformation produces states whose eigenvectors approximately correspond to principal directions. The corresponding transformations of the B , C , and D matrices are defined as follows:

$$\tilde{B} = V^{-1}B, \quad (6)$$

$$\tilde{C} = CV, \quad (7)$$

$$\tilde{D} = D. \quad (8)$$

Owing to the diagonal dominance of the discrete controllability and observability grammian for lightly damped structures, the square of the i th Hankel singular value (HSV) can be approximated as follows (Lim and Gawronski, 1996):

$$\gamma_1^4 \approx \frac{\text{tr}[\tilde{B}\tilde{B}^T]_{ii} \text{tr}[\tilde{C}^T\tilde{C}]_{ii}}{(4\delta_i T)^2}. \quad (9)$$

Note that $\delta_i \leftrightarrow \zeta_i \omega_i$, where ζ_i is the damping ratio of the i th mode and ω_i is the natural frequency of the i th mode. Thus, the i th HSV is directly related to the time scale, $1/(\zeta_i \omega_i)$, of the i th, lightly damped, structural mode. As detailed by Lim and Gawronski (1996), the approximation is accurate up to frequencies near 90% of the Nyquist frequency.

Defining the HSVs through the disturbance path as

$$\Gamma_{zw}^2 = \text{diag}(\gamma_{zw_1}^2, \dots, \gamma_{zw_n}^2), \quad (10)$$

and the HSVs through the control path as

$$\Gamma_{yu}^2 = \text{diag}(\gamma_{yu_1}^2, \dots, \gamma_{yu_n}^2), \quad (11)$$

one obtains a measure of the level of participation of each structural mode or state in the performance path, $P_{zw}(s)$, and control path, $P_{yu}(s)$, respectively. One can compute the HSVs defined in Eq. (11) for every possible input–output path (i.e., every possible combination of an array of transducers targeted for control). Additionally, one can compute the HSVs for the general vector of inputs and outputs (i.e., using all possible actuators and sensors) as a baseline:

$$\bar{\Gamma}_{yu}^2 = (\bar{\gamma}_{yu_1}^2, \dots, \bar{\gamma}_{yu_n}^2). \quad (12)$$

From Eqs. (10)–(12), Lim (1997) defined the following performance metric for sensor/actuator selection:

$$J_{qp} \equiv \sum_{i=1}^n \frac{\gamma_{y_q^u p_i}^4}{\bar{\gamma}_{y_{u_i}}^4} \gamma_{z_{w_i}}^4, \quad (13)$$

where J_{qp} is the performance metric associated with the q th sensor and p th actuator, and n is the number of complex conjugate pole pairs (i.e., modes) of the system. The performance metric defined by Eq. (13) thus provides a measure of the relative modal participation of each sensor/actuator combination over the relative modal participation of all possible sensor/actuator combinations weighted by the modal participation through the performance path. Obviously, if a mode (state) does not exist in the performance path, it is weighted by zero. The metric outlined by Lim (1997) thus offers a convenient and computationally efficient means of determining a subset of actuators and sensors which couple well to modes present in the performance path.

However, for feedback control applications, practical constraints impose limitations on bandwidth, and as such there are always modes outside of the bandwidth of interest which present problems with respect to the trade-off between stability and performance. We often seek to limit the participation of out-of-bandwidth modes by implementing low-pass filters or by forcing the compensator to “roll-off” at some predetermined frequency through loop-shaping techniques. However, such practices can serve to complicate the compensator design. Additionally, we often seek to design compensators based upon reduced-order models of the system, reducing the order of the resulting compensator for practical implementation. All of these practical design issues serve to introduce sensitivity to structural modes present outside of the bandwidth identified for control. Ideally, we would seek a set of sensor/actuator pairs which naturally filter the contribution of out-of-bandwidth modes while at the same time couple well to those modes which dominate the disturbance path.

As was demonstrated by Clark and Cox (1998), a performance metric can be developed which emphasizes band-limited coupling to structural modes important in the disturbance path and de-emphasizes coupling to structural modes outside of the identified bandwidth. To develop this metric, Clark and Cox (1998) distinguished between the number of in-bandwidth modes, n_{in} , and the number of out-of-bandwidth modes, $n_{out} = n - n_{in}$, used in their performance metric. For the in-bandwidth modes, the selection metric is identical to that presented by Lim (1997):

$$J_{qp}^{in} = \sum_{i=1}^{n_{in}} \left(\frac{\gamma_{y_q^u p_i}^4}{\bar{\gamma}_{y_{u_i}}^4} \right) \gamma_{z_{w_i}}^4. \quad (14)$$

However, to enhance stability, we prefer to limit the coupling through the control path, P_{yu} , to the out-of-bandwidth modes. As such, a metric is defined by weighting the normalized HSVs for the control path with themselves and taking the reciprocal of the result to emphasize poor coupling:

$$J_{qp}^{out} = \left(\sum_{i=n_{in}+1}^n \left(\frac{\gamma_{y_q^u p_i}^4}{\bar{\gamma}_{y_{u_i}}^4} \right) \gamma_{y_q^u p_i}^4 \right)^{-1}. \quad (15)$$

If each metric of Eqs. (14) and (15) are normalized with respect to their maximum entry, and the normalized metrics are defined by \bar{J}_{qp}^{in} and \bar{J}_{qp}^{out} , respectively, the sum of the two normalized metrics results in a selection methodology which provides a trade-off between in-bandwidth performance and out-of-bandwidth stability:

$$\hat{J}_{qp} = \bar{J}_{qp}^{in} + \bar{J}_{qp}^{out}. \quad (16)$$

Selection is based upon those sensor/actuator pairs that couple well to modes present in the disturbance path over the desired bandwidth of control and those that couple poorly to modes present in the control path outside of the desired bandwidth of control.

II. REVIEW OF STRUCTURAL ACOUSTIC MODEL

The emphasis of the work presented herein is placed upon the experimental demonstration of the band-limited sensor/actuator selection methodology for disturbance rejection. A test structure was machined from steel, measuring $18 \times 16 \times 0.1875$ in.³ thick. The boundaries were constructed to approximate that of a simply supported plate, and piezoceramic transducers measuring $6 \times 4 \times 0.0075$ in.³ thick were available for experimental implementation. These relatively large piezoceramic patches were not sized for control authority, rather they were selected based upon the results of a prior study which demonstrated that large aperture transducers serve as spatial wave number filters and provide greater control over the low-bandwidth modes required for structural acoustic control (Vipperman and Clark, 1999).

In a previous study by Clark and Cox (1998), the performance metrics defined in Eqs. (13) and (16) were used to select a single sensor/actuator pair devoted to controlling the sound power radiated from the panel, assuming that it radiates into a half-space. The objective was to determine the optimal transducer pair from a predetermined array of possible transducers for control of sound power radiated below approximately 1 kHz. To accomplish this task, radiation modal expansion (Gibbs *et al.*, 1998) was used to generate filters modeling the rms sound power radiated. For the bandwidth identified, only the first radiation mode was used. The first radiation mode captures the radiation of volumetric modes at low frequency. Essentially, sound radiation from the (1,1), (1,3), (3,1), and (3,3) modes is targeted.

For the chosen test structure, an analytic model was developed in an earlier study (Clark and Cox, 1998) using an assumed modes approach with simply supported boundary conditions and proportional damping. This model was suitable for the actuator/sensor optimization process, however, it neglected filter dynamics and nonideal boundary conditions present in the experimental setup. Therefore, for control design and evaluation an experimentally identified model was used, as described in the next section.

In the analytic model there are ten structural modes within the bandwidth of interest. A total of seven target locations were identified resulting in 49 possible sensor/

TABLE I. Center-points of piezoceramic transducers in nondimensional plate coordinates.

Transducer	x coordinate	y coordinate
1	0.3	0.3
2	0.4	0.4
3	0.5	0.5
4	0.3	0.5
5	0.4	0.5
6	0.5	0.3
7	0.4	0.3

actuator combinations including colocated options. Due to the use of piezoelectrics patches as both the sensor and actuator transducer many of these combinations are redundant. For example, using sensor-2 and actuator-3 is equivalent to using sensor-3 and actuator-2. This reciprocity is apparent in the results. The target center-point locations for each transducer in nondimensional plate coordinates are presented in Table I. Additionally, the center-points are indicated by asterisks on the schematic diagram of the plate illustrated in Fig. 2. Also shown in Fig. 2 are the resulting optimal locations for a single sensor/actuator pair based upon the selection metric defined by Clark and Cox (1998) of Eq. (16).

Upon applying the metric developed by Lim (1997), Eq. (13), the optimal sensor/actuator combination selected was a colocated transducer centered on the plate (sensor 3/actuator 3). Due to the efficient coupling of modes with odd indices, this location was anticipated for low-frequency (<1 kHz) structural acoustic control. However, the sensor/actuator pair (sensor 3/actuator 2) resulting from the metric defined by Clark and Cox (1998), Eq. (16), was nonintuitive. It was assumed *a priori* that a single patch centered on the plate would couple well to efficient radiators at low frequency and adequately filter modes at higher frequency due to the dimension of the transducer and its relative symmetry in location. To reconcile the two transducer pair options, the frequency response of each was plotted. As illustrated in Fig. 4, it is obvious why sensor 3/actuator 2 resulted as the optimal transducer pair based upon the metric of Eq. (16): the response to out-of-bandwidth modes (>1 kHz) is significantly less through this sensor/actuator pair than it is through the colocated transducer pair centered on the plate. However,

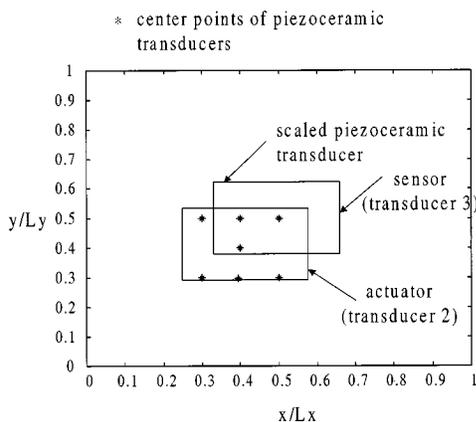


FIG. 2. Schematic diagram of test panel showing center-points of transducer options and optimal sensor/actuator location.

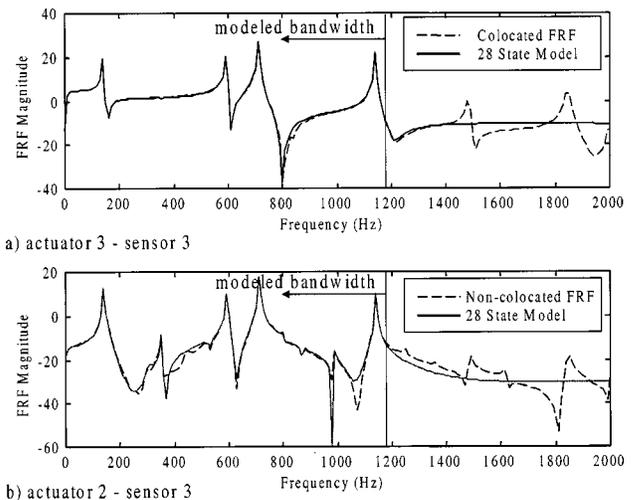


FIG. 3. Comparison of measured frequency response function and that predicted with model obtained through experimental system identification for both transducer pair options.

there is very little difference in the in-bandwidth (<1 kHz) response.

To test the sensor/actuator selections, two transducers were bonded to the test plate as illustrated in the schematic diagram of Fig. 2. Transducers were bonded to opposite sides of the plate due to the required overlap in the optimal locations from the metric of Eq. (16). The transducer centered on the structure was operated as an adaptive sensor-actuator (Vipperman and Clark, 1999) to achieve the colocated transducer pair selected by the metric of Eq. (13). An adaptive sensor-actuator is defined as a transducer used simultaneously for sensing and actuation and capable of continuously rejecting the electrical contribution of the output associated with the feedthrough from the applied electrical input. Specific details of implementation for piezoceramic transducers are outlined in the prior work of Vipperman and Clark (1999).

III. SYSTEM IDENTIFICATION AND CONTROL SYSTEM DESIGN

Prior to designing the control system, the experimental test structure was identified using the time-domain eigenvalue realization algorithm (ERA) (Juang and Pappa, 1985). Each transducer illustrated in Fig. 2 was operated as a sensor-actuator, and the multi-input, multi-output (2I2O) system was identified. A plot of the frequency responses, both measured and from the identified model, of the piezoceramic transducer centered on the plate (transducer 3 of Fig. 2) and operated as a sensor-actuator are presented in Fig. 3(a). As illustrated, the dynamic response of the system is accurately captured with a 28-state model over the bandwidth presented. In Fig. 3(b), the frequency responses of the measured and identified model of the structure using transducer 3 as a sensor and transducer 2 as an actuator are presented. Again, the measured frequency response and that resulting from the identified model correlate well below approximately 1200 Hz.

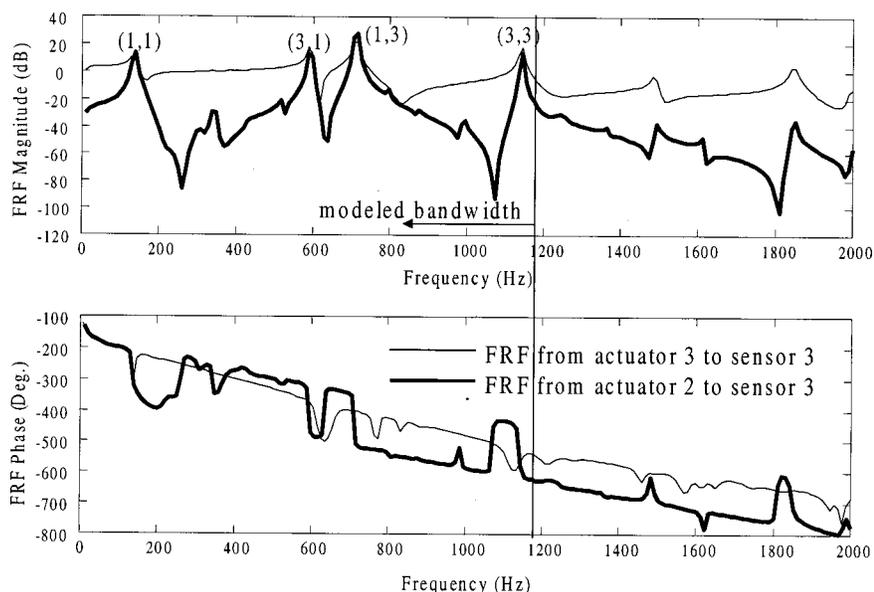


FIG. 4. Comparison of measured frequency response functions for each transducer pair, emphasizing roll-off resulting from spatial filtering through sensor 3–actuator 2..

The bode plots of the two transducer pair options are presented in Fig. 4. As illustrated, both the colocated transducer pair (sensor 3–actuator 3) and the non-colocated transducer pair (sensor 3–actuator 2) couple well to structural modes below 1 kHz. However, above 1 kHz, the non-colocated transducer pair (sensor 3–actuator 2) provides “roll-off” due to spatial filtering, which serves to reduce the effects of structural modes on the closed-loop response outside of the identified bandwidth for control. If a low-pass filter was introduced to achieve this same reduction in magnitude, an additional phase-lag would result. However, due to the spatial filtering provided by the distributed, non-colocated transducers, the reduction in magnitude is achieved at frequencies greater than 1200 Hz with no additional phase-lag. This crude “spatial loop-shaping” provides the control system designer with an added benefit since the in-bandwidth compensator can be designed for more aggressive control without increasing the loop-gain outside of the desired bandwidth for control.

For the experimental results presented, a two-pole Ithaco high-pass filter with a corner frequency of 10 Hz and a four-pole Ithaco low-pass filter with a corner frequency of 2 kHz were used to filter the sensor signals. The data was collected with a TMS320C40 digital signal processor using a 16-input–8-output analog to digital (A/D) and digital to analog (D/A) board. The same boards were used in the implementation of the compensator, at a sample rate of 5 kHz. A four-pole smoothing filter was implemented on the outputs of the D/A at a frequency of 2 kHz. Krohn-Hite model 7600 power amplifiers were used to drive the piezoceramic transducers. A custom designed circuit implemented the adaptive sensor/actuator to operate the center transducer simultaneously as a sensor and an actuator (Vipperman and Clark, 1999).

To develop a model of the radiation filters, the radiation modal expansion (RME) technique was applied (Gibbs *et al.*, 1998), and the plate was subdivided into 20 rectangular elements of equal surface area. The response between each actuator and the acceleration measured at the center-point of

each elemental area was obtained and included as part of the identified model. Through RME, a weighted sum of the velocity measured at each coordinate is developed for each radiation filter used. Thus, if four radiation filters are used, then the 20 velocity measurements can be reduced to four signals whose weighted sum is dependent upon the “shape” (applied through the weighted sum) of the radiation mode. This greatly simplifies the time required for system identification and provides a measure of the radiated sound power over the bandwidth of interest.

A loudspeaker was used to generate the disturbance, and it was placed on center with the plate at a 45-degree angle of incidence and 18 in. from the plate. The input was band-limited noise up to 2 kHz. A comparison of the sound power radiated to that predicted using a single radiation filter is presented in Fig. 5. Recording the frequency response between the disturbance and all 20 accelerometer positions and storing the data provides a mechanism for estimating the radiated power. The power computed from the discretized approximation of the Rayleigh integral is computed and la-

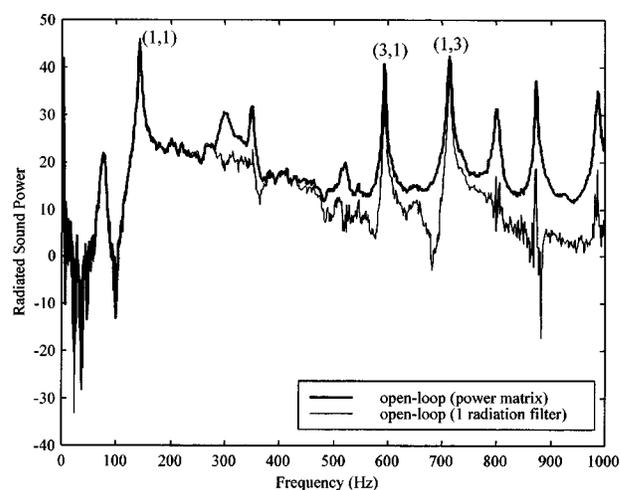


FIG. 5. Comparison of predicted sound power radiated using the power matrix and the first radiation mode.

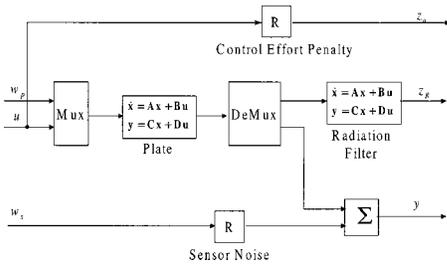


FIG. 6. Block diagram of control system design plant.

beled “power matrix” in Fig. 5. This measure serves as the reference (i.e., best that can be computed using all radiation filters). As illustrated, below 800 Hz, a performance measure based solely upon a single radiation filter is sufficient for designing compensators for dynamic control. While adding additional higher-order radiation filters leads to a better estimate of the radiated power at higher frequency, the single radiation filter, which requires only a single-state filter, sufficiently characterizes the desired performance of the volumetric radiators. Note also that peaks corresponding to resonances of the (1,2), (2,1), and (2,2) structural modes are notably absent from the measured estimate of sound power, owing to their poor radiation efficiency.

From these dynamic models, compensators were designed based upon H_2 -synthesis (Clark *et al.*, 1998). A block diagram of the model used for compensator design is presented in Fig. 6. As illustrated, radiation filters were applied to the output of the experimental model of the plate. A 28-state, reduced-order model of the plate was developed including a one-state model for the radiation filter. Performance variables were selected from z_u , a signal representing control effort penalty, and z_R , a signal representing the radiated power. Measured outputs, y , consisted of transducer 3 as a sensor for both control cases examined. However, control inputs, u , were selected from either transducer 3 for selection metric of Eq. (13) or transducer 2 for selection metric of Eq. (16). The process noise, w_p , was constructed by driving the modeled dynamics through transducer 3 and transducer 2 with independent, spectrally white inputs, and the sensor noise, w_s , was constructed with a spectrally white input applied to the sensors. A comparison of the measured output, y , for a given process noise and sensor noise are plotted for the identified model of the structure in Fig. 7. Since the same sensor was used in each control system design case, the process noise and sensor noise weightings were held at consistent levels in the design process.

IV. RESULTS

Applying H_2 -synthesis, a compensator was designed for each single-input, single-output (SISO) transducer pair selected by the two alternative performance metrics. The loop-gains for each are presented in Fig. 8. In calculating the loop-gain a full order evaluation model of the system was used, which contained the effects of structural modes outside of the desired bandwidth for control. As illustrated, the transducer pair (sensor 3–actuator 3) selected by the performance metric defined in Eq. (13) results in loop-gains near unity at a frequency of approximately 1838 Hz. Given that the dy-

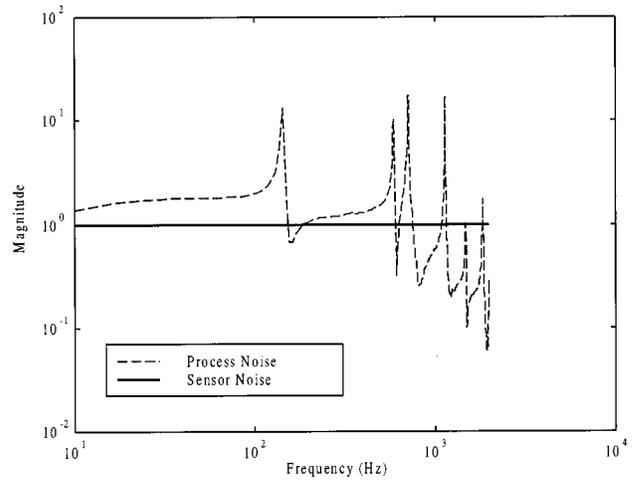


FIG. 7. Sensor output due to process and sensor noise inputs, as used for both control system designs.

namics of the system are ignored at this frequency in the design model, one would expect possible stability problems. However, for the transducer pair (sensor 3–actuator 2) selected by the performance metric defined in Eq. (16), the loop-gain is much less than unity at higher frequencies due to the inherent “roll-off” resulting from the spatial filtering. The in-bandwidth loop-gain (<1 kHz) is roughly the same at all of the resonance frequencies desired in the performance path as illustrated in Fig. 8.

The predicted performance based upon the full-order, experimentally identified dynamics of the system are presented in Fig. 9. The sound power radiated by the (1,1) mode (142 Hz), (3,1) mode (593 Hz), and (1,3) mode (712 Hz) are all attenuated significantly for both transducer pair options. However, at approximately 1838 Hz, an increase in acoustic response is predicted for the colocated transducer pair (sensor 3–actuator 3). In fact, the predicted closed-loop response was on the border of instability, consistent with expectation based upon the loop-gains of Fig. 8. The predicted instability for the experimental system was also consistent with that

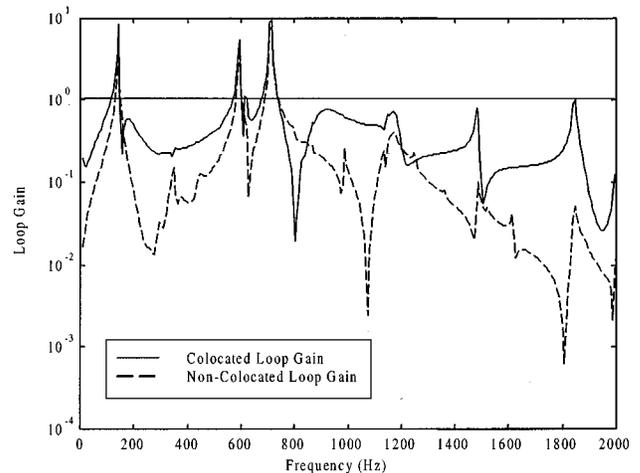


FIG. 8. Predicted loop-gain for alternative transducer pairs given the same design parameters (control effort penalty of $1.5e-5$ and sensor noise of 1).

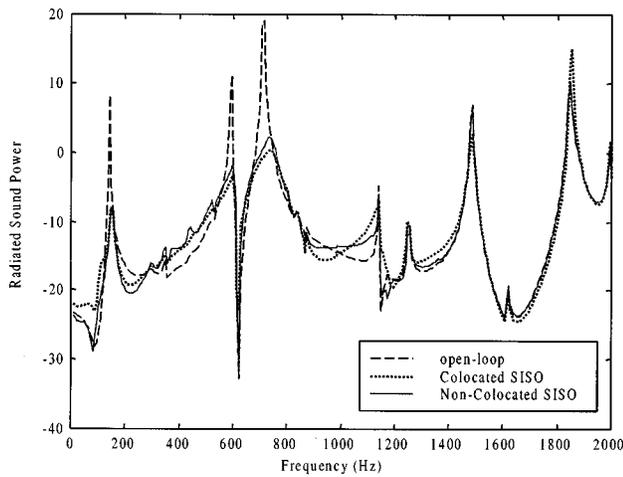


FIG. 9. Predicted closed-loop performance for SISO transducer pair alternatives given a disturbance introduced through the two piezoceramic transducers and a performance measure based upon the first radiation mode.

predicted for the analytical model considered previously by Clark and Cox (1998).

Upon implementing the two compensators on the experimental test-rig, the closed-loop response presented in Fig. 10 was obtained. The frequency responses between the disturbance applied through the loudspeaker and the acceleration measured at the 20 discrete points on the structure required to estimate the sound power radiated were obtained. Applying the full radiation matrix to these measured frequency response functions provides a means of estimating the total sound power radiated. Notice that the performance of the two SISO controllers is nearly identical below 1000 Hz as illustrated in Fig. 10. Also note that the only structural modes controlled are those present in the performance path associated with the first radiation mode. The acoustic power at structural resonances greater than 800 Hz are not modeled and are thus ignored in the dynamic compensator. Including more radiation modes would serve to enhance performance at higher frequencies.

However, upon considering the frequency responses above 1000 Hz (presented in Fig. 11), one observes that the

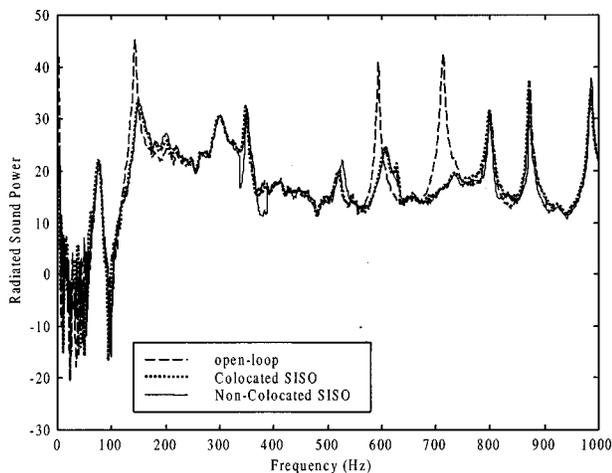


FIG. 10. Measured estimate of open- and closed-loop sound power radiated over the bandwidth targeted for control.

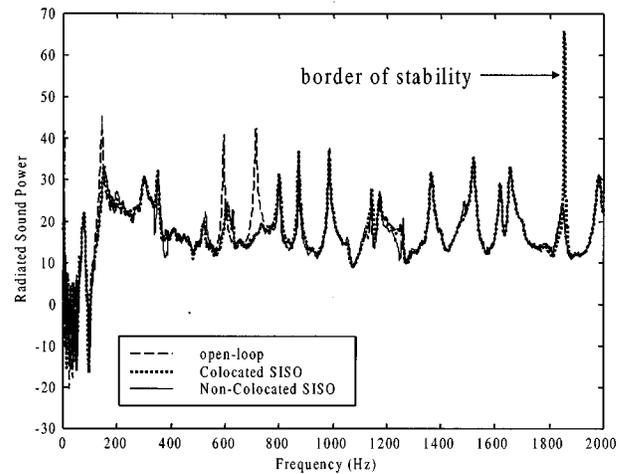


FIG. 11. Measured estimate of open- and closed-loop sound power radiated up to 2 kHz.

controller implemented with the colocated transducer pair (sensor 3–actuator 3) results in significant spillover at approximately 1838 Hz, as was predicted. In fact, the controller for this SISO transducer pair is on the border of instability for the given design parameters (a control effort penalty of $1.5e-5$ and a sensor noise of 1). Thus, through appropriate transducer selection, one can obtain the same level of performance within a prescribed bandwidth while minimizing the effects of spillover outside of the identified bandwidth, as demonstrated by the closed-loop response of the non-colocated transducer pair (sensor 3–actuator 2).

V. CONCLUSIONS

A band-limited method of selecting actuators and sensors for structural acoustic control is reviewed, and experimental results are presented to demonstrate the approach. By applying a selection metric which emphasizes in-bandwidth coupling to structural modes present in the performance path and de-emphasizes out-of-bandwidth coupling to structural modes in the control path, a crude method of “spatial loop-shaping” is developed. The results presented show that this technique can be used to define transducer locations which reduce the effect of spillover in feedback controllers. The technique serves to solidify methods applied in structural acoustic control and adaptive structures in a synergistic design philosophy which incorporates the selection of the transducers as part of the integrated control system design process.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge the support of the Structural Acoustics Branch of the NASA Langley Research Center under Grant No. NCC-1-250, Dr. Richard Silcox, Contract Monitor, and the AFOSR PECASE under Grant No. F49620-98-1-0383, Major Brian Sanders, Contract Monitor.

Baumann, W. T., Saunders, W. R., and Robertshaw, H. H. (1991). “Active suppression of acoustic radiation from impulsively excited structures.” *J. Acoust. Soc. Am.* **90**, 3202–3208.

- Charette, F., Berry, A., and Guigou, C. (1998). "Active control of sound radiation from a plate using a polyvinylidene fluoride volume displacement sensor," *J. Acoust. Soc. Am.* **103**, 1493–1503.
- Clark, R. L., and Cox, D. E. (1998). "Band-Limited Actuator and Sensor Selection of Disturbance Rejection: Application to Structural Acoustic Control," submitted to *AIAA Journal of Guidance, Control, and Dynamics*.
- Clark, R. L., and Frampton, K. D. (1997). "Aeroelastic structural acoustic coupling: Implications on the control of turbulent boundary layer noise transmission," *J. Acoust. Soc. Am.* **102**, 1639–1647.
- Clark, R. L., and Frampton, K. D. (1999). "Aeroelastic structural acoustic control," *J. Acoust. Soc. Am.* **105**, 743–754.
- Clark, R. L., and Fuller, C. R. (1991). "Control of sound radiation with adaptive structures," *J. Intell. Mater. Syst. Struct.* **2**, 431–452.
- Clark, R. L., and Fuller, C. R. (1992a). "Optimal placement of piezoelectric actuators and polyvinylidene fluoride error sensors in active structural acoustic control approaches," *J. Acoust. Soc. Am.* **92**, 1521–1533.
- Clark, R. L., and Fuller, C. R. (1992b). "Modal sensing of efficient acoustic radiators with PVDF distributed sensors in active structural acoustic approaches," *J. Acoust. Soc. Am.* **91**, 3321–3329.
- Clark, R. L., Saunders, W. R., and Gibbs, G. P. (1998). *Adaptive Structures: Dynamics and Control* (Wiley, New York).
- Clark, R. L., Burdisso, R. A., and Fuller, C. R. (1993). "Design approaches for shaping polyvinylidene fluoride sensors in active structural acoustic control (ASAC)," *J. Intell. Mater. Syst. Struct.* **4**, 354–365.
- Cunnefare, K. A. (1991). "The minimum multimodal radiation efficiency of baffled finite beams," *J. Acoust. Soc. Am.* **90**, 2521–2529.
- Frampton, K. D., and Clark, R. L. (1997). "Control of sound transmission through a convected fluid loaded plate with piezoelectric sensor/actuators," *J. Intell. Mater. Syst. Struct.* **8**(8), 686–696.
- Fuller, C. R. (1998). "Analysis of Active Control of Sound Radiation from Elastic Plates by Force Inputs," in *Proceedings of Inter-Noise '88*, Avignon, Vol. 2, pp. 1061–1064.
- Fuller, C. R., and Burdisso, R. A. (1991). "A wave number domain approach to active control of structure-borne sound," *J. Sound Vib.* **148**, 355–360.
- Fuller, C. R., Elliott, S. J., and Nelson, P. A. (1996). *Active Control of Vibration* (Academic, London).
- Gibbs, G. P., Clark, R. L., Cox, D. E., and Viperman, J. S. (1999). "Radiation modal expansion: Application to active structural acoustic control," *J. Acoust. Soc. Am.* (to be published).
- Gu, Y., Clark, R. L., Fuller, C. R., and Zander, A. C. (1994). "Experiments on active control of plate vibration using piezoelectric actuators and polyvinylidene fluoride PVDF modal sensors," *Trans. ASME, J. Vib. Acoust.* **116**, 303–308.
- Heck, L. P., and Naghshineh, K. (1994). "Large-Scale, Broadband Actuator Selection for Active Noise Control," in *Proceedings of Noise-Con 94*, pp. 291–296.
- Johnson, M. E., and Elliott, S. J. (1995). "Active control of sound radiation using volume velocity cancellation," *J. Acoust. Soc. Am.* **98**, 2174–2186.
- Juang, J.-N., and Pappa, R. S. (1985). "An eigensystem realization algorithm for modal parameter identification and model reduction," *J. Guid. Control. Dyn.* **8**(5), 620–627.
- Lim, K. B., and Gawronski, W. (1996). "Hankel singular values of flexible structures in discrete time," *J. Guid. Control. Dyn.* **19**(6), 1370–1377.
- Lim, K. B. (1997). "Disturbance rejection approach to actuator and sensor placement," *J. Guid. Control. Dyn.* **20**(1), 202–204.
- Lomas, N. S., and Hayek, S. I. (1977). "Vibration and acoustic radiation of elastically supported rectangular plates," *J. Sound Vib.* **52**(1), 1–25.
- Maillard, J. P., and Fuller, C. R. (1998). "Comparison of two structural sensing approaches for active structural acoustic control," *J. Acoust. Soc. Am.* **103**, 396–400.
- Ruckman, C. E., and Fuller, C. R. (1995). "Optimizing actuator locations in active noise control systems using subset selection," *J. Sound Vib.* **186**(3), 395–406.
- Smith, G. C., and Clark, R. L. (1999). "Optimal Transducer Placement for Output Feedback Control of Broadband Structural Acoustic Radiation," *Journal of Intelligent Material Systems and Structures* (to be published).
- Smith, G. C., Clark, R. L., and Frampton, K. D. (1998). "The Effect of Convected Fluid Loading on the Optimal Transducer Placement for Active Control of Sound Transmission Through an Aeroelastic Plate," *AIAA Paper 98-1981*.
- Snyder, S. D., and Hansen, C. H. (1990). "Using multiple regression to optimize active noise control system design," *J. Sound Vib.* **148**(3), 537–542.
- Snyder, S. D., Hansen, C. H., and Tanaka, N. (1993). "Shaped vibration sensors for feedforward control of structural radiation," in *Proceedings of the Second Conference on Recent Advances in Active Control of Sound and Vibration*, 28–30 April 1993, Blacksburg, VA.
- Thomas, D. R., and Nelson, P. A. (1995). "Feedback control of sound radiation from a plate excited by a turbulent boundary layer," *J. Acoust. Soc. Am.* **98**, 2651–2662.
- Vipperman, J. S., and Clark, R. L. (1999). "Implications of using collocated strain-based transducers for active structural acoustic control," *J. Acoust. Soc. Am.* **106**, 1392–1399.
- Wallace, C. E. (1972a). "Radiation resistance of a baffled beam," *J. Acoust. Soc. Am.* **51**, 936–945.
- Wallace, C. E. (1972b). "Radiation resistance of a rectangular panel," *J. Acoust. Soc. Am.* **51**, 946–952.
- Wang, B. T., Burdisso, R. A., and Fuller, C. R. (1991). "Optimal Placement of Piezoelectric Actuators for Active Control of Sound Radiation from Elastic Plates," in *Noise-Con*, pp. 267–274.

Statistical structures of indoor traffic noise in a high-rise city

S. K. Tang and W. H. Au

Department of Building Services Engineering, The Hong Kong Polytechnic University, Hong Kong, China

(Received 3 February 1998; revised 1 April 1999; accepted 1 August 1999)

Noise measurements were conducted in seven residential flats close to the traffic in Hong Kong. Noise-level time variations and their statistics are discussed. The results of the overall noise level statistics reveal that Gaussian noise-level distribution can only be found under the free traffic-flow condition while some characteristics of gamma noise-level distributions are observed when the flow is of the interrupted type. Short-duration noise-level time variation statistics reveal a well-defined relationship between skewness and kurtosis for each traffic-flow pattern identified. Results also suggest that the Pearson type I and IV distributions are useful for describing the short-duration noise-level distributions. © 1999 Acoustical Society of America. [S0001-4966(99)00512-3]

PACS numbers: 43.50.Lj, 43.50.Jh [MRS]

INTRODUCTION

In a densely populated high-rise city such as Hong Kong, the acoustical environment under the influence of traffic noise may differ from those of the smaller cities. This is largely due to the proximity of trunk roads to large buildings^{1,2} and the effect of city reverberation.³

The importance of traffic noise in environmental control has attracted the attention of many acousticians and scientists in the past few decades. Griffiths *et al.*⁴ studied the subjective effects of traffic-noise exposure on human beings in the London area. A similar study focused on the behavior of human beings exposed to traffic noise was carried out in two French cities by Lambert *et al.*⁵ Traffic-noise criteria for assessing the annoyance caused by traffic noise have also been studied. Scholes⁶ found that the A-weighted equivalent sound-pressure level L_{Aeq} and the traffic noise index (TNI) calculated from the A-weighted percentile levels L_{A10} and L_{A90} correlate satisfactorily with human annoyance. The use of percentile levels as the noise criteria was further investigated by Langdon and Griffiths.⁷ L_{A10} is widely used in traffic noise-control practices nowadays.

Physical noise measurements were also conducted by researchers all over the world. Examples are the works of Ko¹ in Hong Kong, Ishiyama *et al.*² in Tokyo, Cannelli⁸ in Rome, and more recently, Chakrabarty *et al.*⁹ in Calcutta. Scale-model studies have also been done by Liu¹⁰ and Yamashita and Yamamoto.¹¹ However, this list is by no means exhaustive. Empirical formulas for the estimation of L_{A10} under known traffic volume have also been developed.¹²

Statistical modeling of traffic noise is also a research topic. Foxon and Pearson¹³ concluded from the results of a site measurement that the distribution of traffic-noise level is Gaussian-like. However, Kurze¹⁴ has showed both mathematically and experimentally that the noise-intensity fluctuation due to free-flowing traffic is gamma-distributed. It should be noted that Foxon and Pearson¹³ and Kurze¹⁴ were using different quantities for describing the noise statistics. Foxon and Pearson¹³ used the noise level in dB while Kurze¹⁴ discussed the noise intensity. The noise level L is related to the noise intensity I by the well-known logarithmic relationship

$$L = 10 \log_0(I/I_{ref}),$$

where I_{ref} denotes the reference noise intensity. The more recent results of Don and Rees¹⁵ suggest complicated noise-level distributions depending on the compositions of road vehicles, and they have developed a model based on the superposition of Gaussian distributions for the estimation of traffic-noise probability distribution. However, their proposed model has not been widely adopted in the traffic-noise prediction practice. Despite the efforts of previous researchers, a commonly agreed-upon distribution function for traffic noise has not been sought, nor its shape under different traffic-flow patterns.

Most of the measurements mentioned above were not conducted at the facades of buildings. In a high-rise and densely populated city, the transmission of traffic noise the indoor built environment through the building facades is a big problem. Owing to the large variation of the indoor acoustical conditions, results of indoor traffic noise in existing literature are very limited. One example is due to Ko,¹⁶ who showed the existence of a fairly good correlation between indoor and outdoor traffic-noise levels in the open window case. The present study is an attempt to find out whether a fixed traffic-flow pattern will lead to a certain general shape of the indoor traffic noise-level distribution. It is hoped that the present results can provide information for further development of the traffic-noise prediction method.

I. SITE MEASUREMENTS

A. Methodology

In the present study, seven site measurements were carried out within the urban area of Hong Kong where the major source of noise was traffic. Simultaneous recording of outdoor and indoor $L_{Aeq, T=2s}$ was done using two Metrosonics dB-3100C integrating sound-level meters. They also measured noise-level probability distributions, the overall L_{A10} and L_{A90} . The outdoor noise levels were measured 1 m from the building facades while the indoor ones were recorded close to the centers of the indoor environments. Duration of each measurement was at least 25 min. A video recording of the traffic flow was carried out in parallel with each noise

TABLE I. General features of measurement sites.

Site	A	B	C	D	E	F	G
Road type	Main trunk	Main trunk	Main trunk	Distributor	Distributor	Distributor	Distributor
Speed limit (km/h)	70	70	70	50	50	50	50
Anticipated traffic flow	Free	Free	Free	Free	Interrupted	Interrupted	Interrupted
Number of lanes	6 (3 in each direction)	6 (3 in each direction)	6 (3 in each direction)	4 (2 in each direction)	6 (3 in each direction)	4 (2 in each direction)	4 (2 in each direction)
Distance from traffic light (km)	>1	>1	>1	~0.5	0.05	0.07	0.05
Traffic light duration ^a (s)	60G+60R both sides	30G+80R near side 20G+90R far side	37G+60R near side 20G+60R far side
37G+60G near side 20G+60R far side							
General building clearance (m)	>10	>10	>10	~5	generally no clearance	<5	>10
General building height (m)	80	80	65	35	34	40	65
Indoor reverberation time	0.8	0.8	0.4	0.5	0.4	0.4	0.8

^aG: green light; R: red light.

measurement for later traffic count and other analyses. Mean speeds of the vehicles were also estimated from their passage time between two lampposts with a known separation. In addition, reverberation measurement was carried out in each indoor environment using the Brüel & Kjær 2144 frequency analyzer in the multispectrum mode. Open window condition was adopted in the present study.

B. General site information

The seven sites selected in the present study are either near a main trunk road or a road junction with buildings on both sides of the associated roads. Thus, the results are mainly related to two types of traffic-flow patterns—the free flow¹⁴ and the interrupted flow.¹⁷ The latter flow pattern describes the situation in which the vehicles decelerate, stop, and accelerate again.¹⁷ This phenomenon is common for road junctions. These two mentioned traffic-flow patterns are also typical in high-rise and densely populated cities. It is not intended to subdivide the interrupted flow pattern into banked and pulsed flow patterns as in Don and Rees.¹⁵ This will be discussed later.

Table I summarizes the general features of these sites and Fig. 1(a) to (c) show the survey maps for sites C, D, and E which illustrate all the essential features of the present surveyed sites. Effect of traffic interruption by traffic light is important at sites E to G, which are close to road junctions. However, the traffic pattern for site D was unavoidably affected by a traffic light located about 500 m away from the measurement point, though the vehicles could move with speeds similar to those in sites A to C. Sites A to C are close to main trunk roads where there is no traffic light within 1 km from measurement points. It should also be noted that the buildings surrounding site E are packed very closely to each other so that the clearance between them is very small. In addition, the results from the reverberation time measurements reveal that a great variety of indoor environments has been covered in the present study. The indoor environments

associated with sites A, B, and G are unfurnished. The general heights of the buildings shown in Table I reflect a feature of a high-rise city. The road junctions studied by Chakrabarty *et al.*⁹ are surrounded by low-rise buildings.

C. Traffic-flow data

Traffic count and vehicle-speed estimation were done after the site measurements by playing back the video recordings. Traffic volume, the decomposition of traffic, and average vehicle speed for each site are tabulated in Table II. The average vehicle speeds for the sites close to road junctions are obtained during the green traffic light condition. Also, a majority of the vehicles fell into the category “passenger car and taxis.” In general, the average vehicle speed under the free-flow condition is higher than that under the interrupted flow. However, as stated before, the traffic flow at site D was affected by a traffic light 500 m away from the noise measurement point so that the vehicles may not have attained their normal speeds in the free-flow situation. It should also be noted that the speed limit enforced by the local government will limit the speeds of the vehicles and since the road at site D is not a main trunk road, the lower traffic volume will lead to a partially continuous flow pattern. This type of traffic flow is referred to as the partially continuous free flow in the foregoing discussions. It will be shown later that the results obtained at site D show the features of both the free and interrupted flows.

II. RESULTS AND DISCUSSION

Both the equivalent sound-pressure levels at 1 m from the building facade and in the indoor environments were recorded every 2 s in the present study. In the following sections, their time variations and statistical characteristics are discussed. In the present investigation, the time variations of the indoor noise levels follow closely those of the

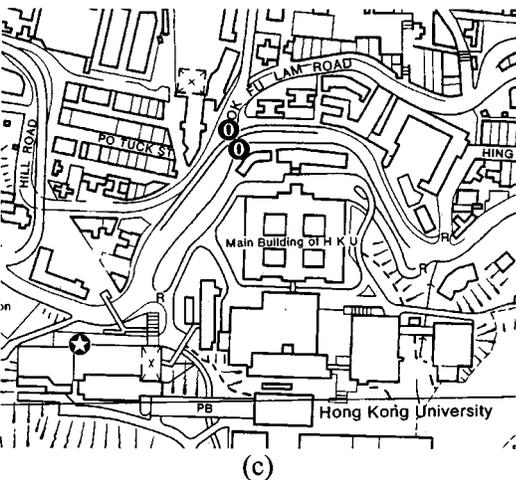
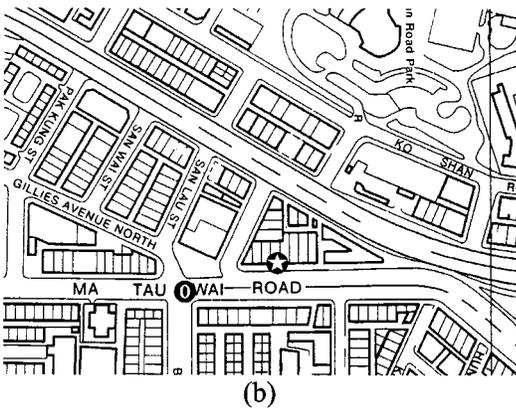
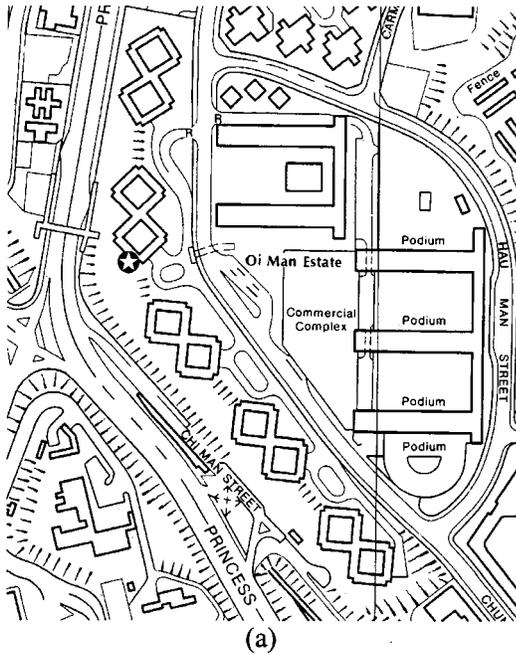


FIG. 1. Examples of measurement site locations. (a) site C; (b) site E; (c) site D. ⚡: Measurement location; ●: traffic light.

outdoors, as discussed in Ko.¹⁶ Also, it is found that the indoor reverberation characteristics do not affect this relationship, though it does affect the difference between L_{Aeq} in the indoor environment and at the building facade (not

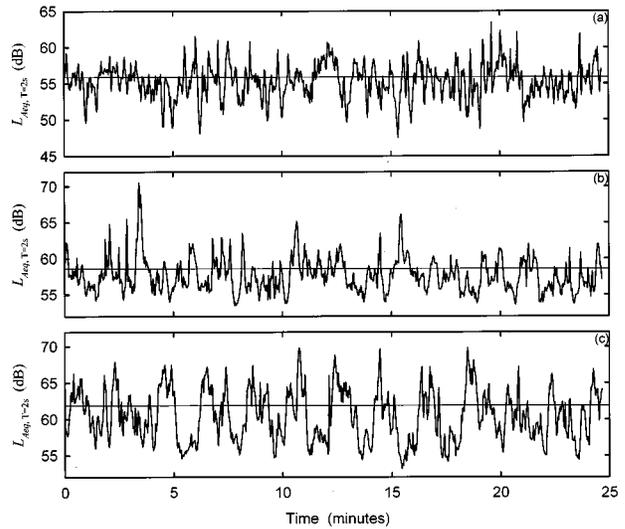


FIG. 2. Time variations of $L_{Aeq,T=2s}$. (a) site C ($L_{Aeq,T=25\text{ min}}=55.9\text{ dB}$); (b) site F ($L_{Aeq,T=25\text{ min}}=58.6\text{ dB}$); (c) site E ($L_{Aeq,T=25\text{ min}}=61.9\text{ dB}$). Horizontal straight line: $L_{Aeq,T=25\text{ min}}$.

shown here). In order to simplify the following discussion, the results presented hereinafter, unless otherwise stated, are those obtained indoors.

A. Noise-level variations and their overall statistics

Occasional lulls in the traffic flow result in reduced levels, which appear as downward-pointing spikes in the time trace of noise-level variations. A typical example of this is that obtained at site C (free flow) shown in Fig. 2(a). Bursts of accelerating or decelerating traffic, especially prevalent near traffic lights, produce the spikes of increased noise observed in Fig. 2(b) (site F, interrupted flow). Such spikes can have a magnitude as high as 15 dB. Figure 2(c) illustrates the noise-level time variation obtained at site E, which is affected by a traffic signal with equal duration of red and green light. The noise level varies periodically with nearly constant amplitude, resulting in a relatively uncommon statistical distribution. It will be discussed later.

The partially continuous free-flow traffic pattern at site D contains both lulls and bursts resulting from the traffic light located far away from the measurement point and the lower traffic volume (not shown here). A similar phenomenon is observed in an additional measurement conducted at site A when the traffic volume was only about 3300 vehicles per h and the nominal vehicle speed was 65 km/h. Such excursions from the average level significantly affect the shape, and thus influence the higher moments of the traffic noise-level distribution. Skewness, s , and kurtosis, k , as defined by Stuart and Ord¹⁸

$$s = \frac{\sum f(x - \bar{x})^3}{\sigma^3} \quad \text{and} \quad k = \frac{\sum f(x - \bar{x})^4}{\sigma^4} - 3, \quad (1)$$

where f is the probability distribution, $\{x\}$ the sample, σ the standard deviation of $\{x\}$, and \bar{x} the sample mean, have been calculated for the free-flowing (sites A, B, and C) and the interrupted flow (sites E, F, and G) traffic situations as well as for site D, which appears to be a mixture of both flow conditions. Results are presented in Table III. The skewness

TABLE II. Traffic-flow data summary.

Site	Number of vehicles/hour ^a									Mean speed (km/h)
	V1	V2	V3	V4	V5	V6	V7	V8	Total	
A	95	2304	816	0	96	960	517	242	5030	64
B	52	1620	1188	0	84	1080	480	186	4690	63
C	29	1992	84	0	0	36	7	0	2148	72
D	58	636	216	312	120	84	30	0	1456	45
E	8	780	0	144	204	36	34	4	1210	29
F	7	660	132	144	72	60	19	0	1094	30
G	2	336	48	12	48	24	4	0	472	23

^aV1: Motorcycle; V2: passenger car and taxis; V3: light truck or van of unladen weight <2.8 tons; V4: light bus; V5: double deck bus; V6: median truck of unladen weight <5.5 tons; V7: heavy truck of unladen weight >5.5 tons excluding container vehicle; V8: container vehicle.

for the interrupted traffic-flow cases is always positive. Figure 3 illustrates some typical examples of the indoor noise-level distributions for the three types of traffic conditions at 1-dB bin range. The corresponding distributions for the outdoor noise levels are very similar to those shown in Fig. 3 and are not presented.

For a gamma-distributed noise-intensity time record, the resulting noise-level distribution will have a negative skewness s . Therefore, Fig. 3 shows that the noise-intensity fluctuation may be gamma-distributed only for the case with a continuous free-traffic flow. However, the skewness of the present distributions is very much higher than that of the noise-level distributions having the same kurtosis but with gamma-distributed noise-intensity fluctuations as shown in Table III. This suggests that the present noise intensity does not follow the gamma distribution proposed theoretically by Kurze.¹⁴ Also, the present distributions differ from those presented by Ko,¹ which show significant positive skewness. However, it should be noted that although there is no agreement with the gamma distribution in these overall statistics, the noise-level distributions within short time intervals do follow some common probability density distributions. This will be discussed in the next section.

The corresponding noise-level distributions for the partially continuous free-traffic flow cases are Gaussian-like with small skewness and kurtosis (Table III). This is probably due to the randomness in the arrival time and spatial distribution of vehicles so that the results tend to agree with those of Johnson and Saunders.¹⁹ The Gaussian-like traffic

noise-level distribution of Foxon and Pearson¹³ appears to be due to this partially continuous flow pattern. Their measurement site was located in an area with both free flow and interrupted traffic—a situation similar to site D.

The higher the discontinuity of the traffic flow, the more positive the noise-level distribution skewness will be. In general, the shapes of the noise-level distributions do not really agree with those suggested by Don and Rees.¹⁵ Though the noise-level distribution for site E has two peaks, the present data were recorded in the afternoon instead of in the evening as in Don and Rees.¹⁵ Such discrepancy may be due to the fact that the results of Don and Rees¹⁵ are noise-level statistics within 300 s so that they are sensitive to the intermittent changes in the traffic-flow pattern. However, it will be shown in the next section that their proposed noise-level distribution shapes can occur in the short-time statistics. Noise-level distributions at site G and, in particular, site F have skewness close to unity and approximate more closely a gamma distribution (for a gamma distribution with mean $p=4.9$, $s=0.90$, and $k=1.22$). The positive skewness results from the high-magnitude upward spikes in the $L_{Aeq, T=2\text{ s}}$ time variation such that the overall equivalent sound-pressure level L_{Aeq} is higher than the mode of the level distribution [e.g., Fig. 2(b)].

B. Short-time statistics

Since the noise levels keep on changing throughout the whole measurement period, it is worthwhile to investigate

TABLE III. Skewness and kurtosis in overall indoor noise-level statistics.

Site	Skewness s^a	Kurtosis k	Standard deviation	
			σ (dB)	$L_{Aeq, T=25\text{ min}}$ (dB)
A	-0.45 (-0.77)	0.23	1.49	73.1
B	-0.38 (-0.65)	0.06	1.92	51.6
C	-0.12 (-0.78)	0.39	2.38	55.9
D	0.06 (-0.51)	-0.39	2.27	73.7
E	0.25 (-0.44)	-0.79	3.55	61.9
F	0.88 (no solution)	1.35	2.41	58.6
G	0.99 (-0.86)	1.02	4.13	65.0
Additional measurement	0.02 (-0.71)	0.44	1.89	71.3

^aData in parentheses denote skewness of noise-level distribution with same k but obtained from gamma-distributed noise-intensity time fluctuation.

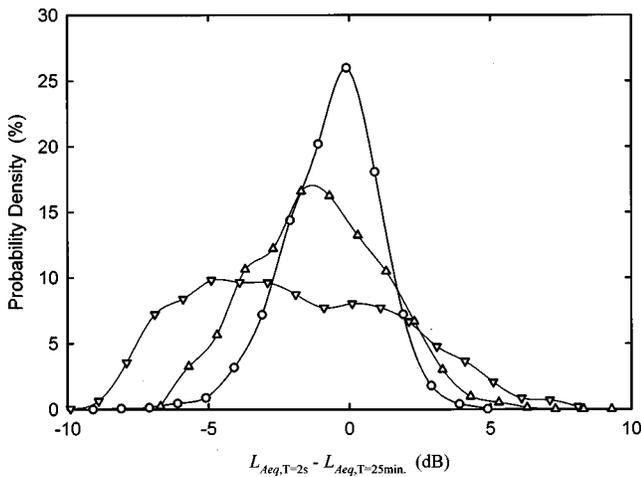


FIG. 3. Examples of probability density distributions from overall statistics of 25-min averages. ○: site A; △: site D; ▽: site E.

how the statistical properties of the noise-level variations will change with time for a deeper understanding on the nature of traffic-noise fluctuations. This section is an attempt to look into the variations of such properties and to look for their common characteristics. Statistical parameters, such as the standard deviation, skewness, and kurtosis, are calculated within any 5-min consecutive intervals of the recorded indoor $L_{Aeq, T=2s}$ time variations. It is not believed that shorter calculation intervals are suitable to ensure reliable statistics because transient noises and pulses will then contaminate the results so obtained. Five-minute intervals have been adopted for the measurement of unsteady noise inside office buildings²⁰ as well as in the traffic-noise study.¹⁵

The standard deviations of the noise levels in the present study are lower than those of Don and Rees¹⁵ (Fig. 4). It can also be observed from Fig. 4 as well as from Table III that higher σ s are found at sites E and G where the durations of the red traffic light are relatively short (Table I) and the nominal speeds of the vehicles are low (Table II). Deviation of the present short-time noise-level statistics from the Gaussian distribution is observed where the noise climate

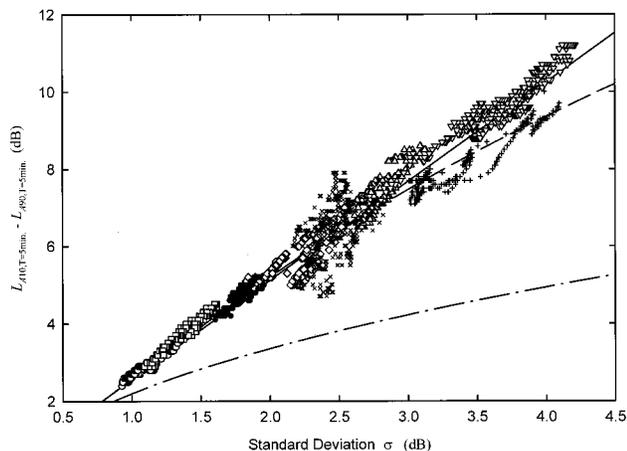


FIG. 4. Variation of 5-min noise climate with noise-level standard deviation. ○: site A; □: site B; ×: site C; △: site D; ▽: site E; ◇: site F; +: site G; ●: additional measurement at site A; —: Gaussian; - - -: Weibull (Ref. 22); - · - ·: gamma.

$L_{A10, T=5 \text{ min}} - L_{A90, T=5 \text{ min}}$ is concerned (Fig. 4). The noise climate can be assumed to be proportional to σ only when the latter is small ($\sigma \leq 1.5$ dB). This is different from the results of Don and Rees.¹⁵ Gamma noise-level distribution is not valid in the present case. Relatively large scattering is observed at $\sigma \approx 2.4$, where the variation range of the noise climate is around 4 dB. This results from the free-flow traffic at site C. This will be further discussed later.

Examples of the 5-min noise-level distributions, in 0.5-dB bins, indicate substantial differences with time at the same site [Fig. 5(a) and (b)]. An approximately uniform distribution with a small U-shape obtained from an interrupted traffic flow at site E [Fig. 5(e)] differs from profiles reported elsewhere (for instance, Don and Rees¹⁵). It is found from a replay of the video recording that the vehicles stopped and started periodically with a distinctive frequency because of the traffic signal (Table I). The time variation of the noise levels therefore appears approximately sinusoidal with nearly constant amplitude [Fig. 2(c)]. Since a sinusoidal time variation of noise levels will tend to produce a U-shaped distribution, the occurrence of the distribution observed in Fig. 5(e) is explained.

Since the short-time noise-level distributions vary substantially with time, it is not possible to analyze each of them individually. In order to characterize these distributions, the method of Pearson is employed. The type of distribution function for describing a set of sampled data is determined by analyzing the skewness and kurtosis of the sample on a skewness–kurtosis chart called the (β_1, β_2) chart of Pearson,¹⁸ where $\beta_1 = s^2$ and $\beta_2 = k$. Details of the procedure and the importance of the relationship between skewness and kurtosis can be found in standard textbooks on advanced statistics (for instance, Elderton and Johnson²¹). The present analysis is started off by comparing the (s, k) relationships of the present data with those of the common distributions.

Though many different shapes of the noise-level distributions can be observed and they seem to appear randomly, their skewness s and kurtosis k are bounded and, to a large extent, related for each type of traffic flow identified in the present study. Figure 6(a) is a plot of k against s for the interrupted traffic cases. Essential features of this plot are that s is always positive and k can be negative or positive but it increases and shows a definite relationship with s . The noise-level distributions are neither Gaussian nor gamma, although some of the overall statistics tend to reveal the characteristics of these two well-known distributions (Fig. 3 and Table III). The (s, k) relationships of some common distribution functions¹⁸ are also included in Fig. 6(a) but none of them agrees with the present results. Though some agreements with the noise-level distribution with a gamma-distributed noise-intensity time variation are observed for small positive s , this type of level distribution is in general not really valid in the positive s range. It should be noted that a similar (s, k) relationship is also observed for the outdoor measurements and thus they are not presented. It can also be noted from Fig. 6(a) that the data for site E show negative k with small s , showing that the associated noise-level distributions are quite uniform.

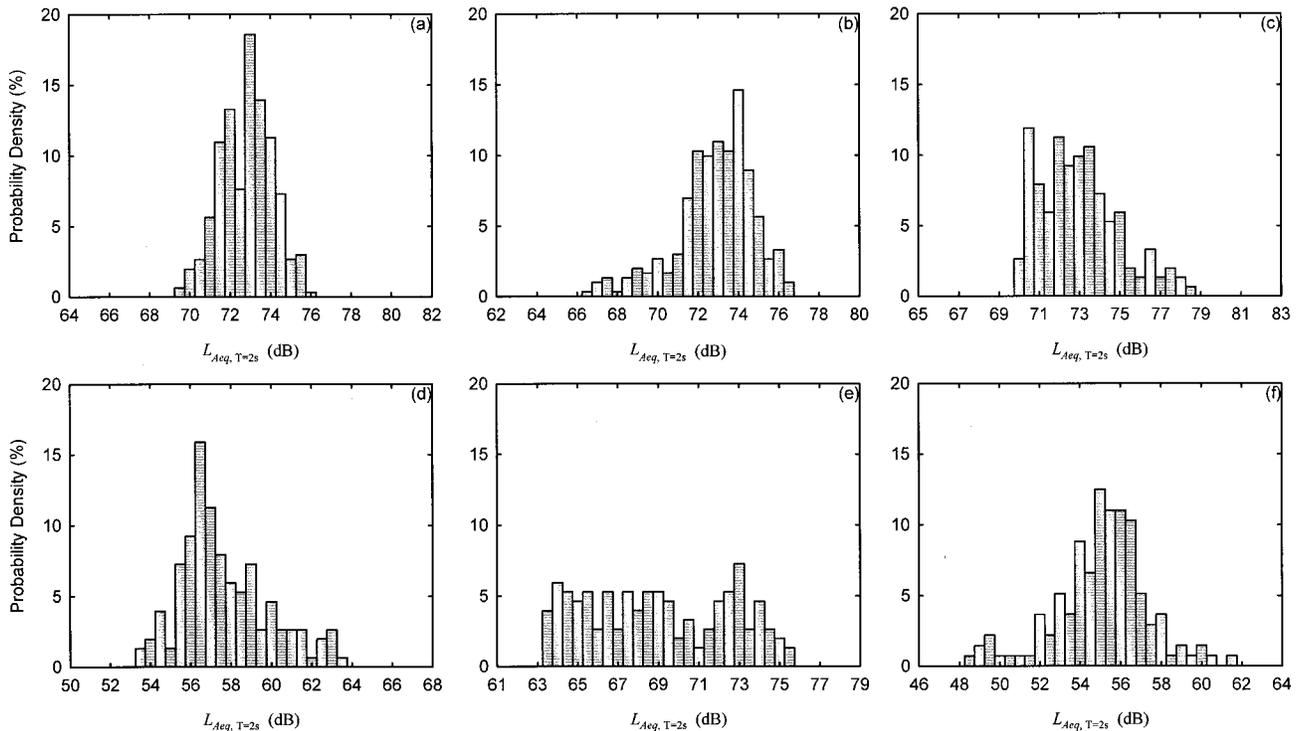


FIG. 5. Examples of 5-min noise-level distributions. (a) site A; (b) site A; (c) site D; (d) site F; (e) site E; (f) site C.

Figure 6(b) shows the (s,k) relationship for the free-traffic-flow cases, which is less ordered than that shown in Fig. 6(a). Unlike the case for interrupted traffic, skewness s in this case can be positive or negative and $-1 \leq s \leq 1$. This coexistence of positive and negative s , together with a relatively large variation in the kurtosis k and a less-ordered (s,k) relationship in the 5-min statistics will result in a large variation in the shape of the noise-level probability density and cumulative distributions. These two distributions have a significant effect on the noise climate, resulting in the large scattering of data associated with site C observed in Fig. 4 ($\sigma \approx 2.4$).

Skewness s in this free-flow case is not as large as that for the interrupted cases. The variation of k with s tends to be bounded by the prediction of gamma-distributed noise-intensity fluctuations for negative s within reasonable tolerance. For positive s , no collapse of data is possible in reality, as the gamma-distributed noise-intensity time fluctuation will result in a negatively skewed noise-level distribution. This is because the noise intensity and noise level are related by a logarithmic relationship. Some data tend to agree with that of a Weibull-distributed noise-level time variation²² and quite a number of the distributions have s and k close to zero, suggesting the presence of Gaussian noise-level distributions in the short-time statistics under the free-flow condition.

No agreement with the gamma-distributed noise-level prediction can again be observed under the free-flow condition. A Gaussian-distributed noise-intensity time fluctuation gives rise to a noise-level distribution with negative s and large k —a situation not observed in the present study. Thus, the associated (s,k) relationship is not included in the discussions. Also, a Weibull-distributed noise-intensity time fluctuation

gives $s = -1.14$ and $k = 5.40$, which are outside the range of the experimental results.

The (s,k) relationship for site D, which represents the case of a partially continuous free flow affected by a distant traffic light, is shown in Fig. 6(c). That for the additional measurement at site A agrees with Fig. 6(c) for small positive $s < 1$ (not shown here). Skewness s for this additional measurement, which was done close to a highway without flow restriction of any form, is seldom larger than 1. The corresponding traffic volume was low. Though the noise in this case appears intermittent, the continuous generation of noise by the vehicles does not in general produce very positively skewed noise-level distribution, as L_{Aeq} in this case does not depend very much on the magnitudes of several specific spikes in the noise-level time variation. This is not the case for a traffic flow with serious interruption.

Though both s and k vary with time, the present observed (s,k) relationship, at least for a particular type of traffic flow, tends to suggest that both the indoor and outdoor noise-level distributions are not as random as one expects. The present (s,k) relationships appear much more organized than those of Don and Rees.¹⁵ Also, the above (s,k) relationships are again found when the time interval for calculation is reduced to 4 or even 3 min, though a higher order of scattering can be anticipated (not shown here), suggesting the robustness of these relationships. For the shorter time-interval cases, the ranges of s and k become larger, in general.

It is observed that the short-time statistical structure of the noise-level time variation differs from the overall statistics quite significantly. It is expected, especially for the interrupted traffic-flow cases, because the short-time statistical

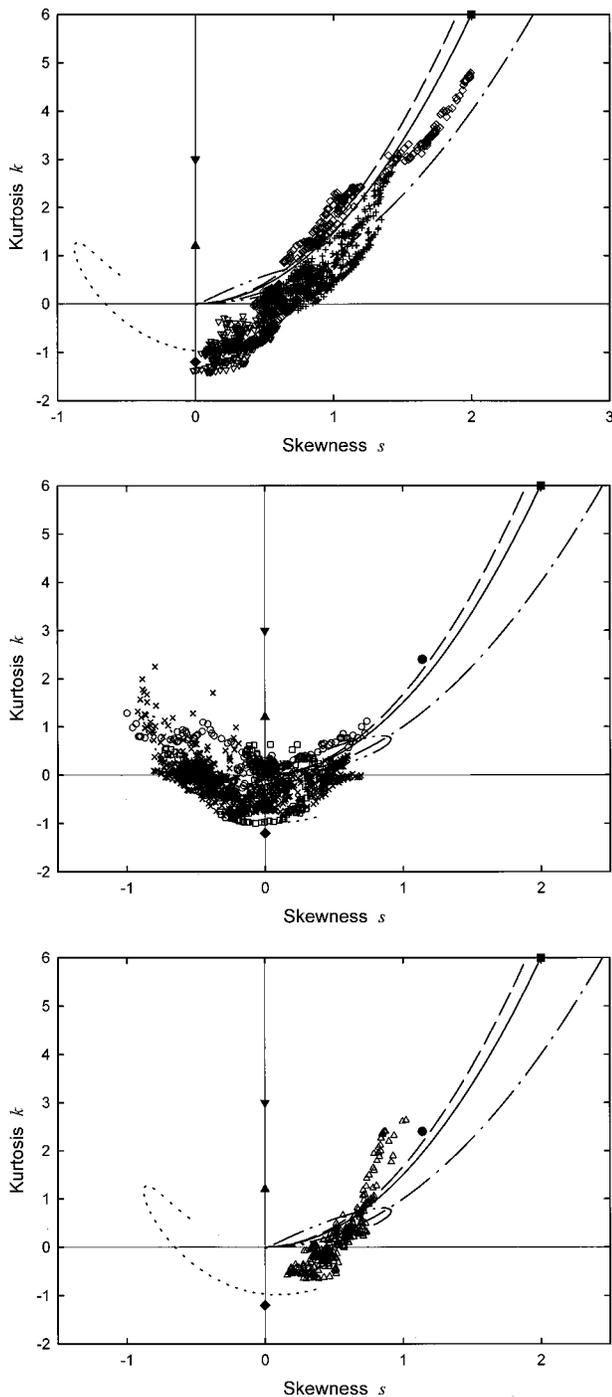


FIG. 6. Relationships between skewness and kurtosis. (a) interrupted traffic flow; (b) free flow; (c) partially continuous traffic flow. \circ : site A; \square : site B; \times : site C; \triangle : site D; ∇ : site E; \diamond : site F; $+$: site G; \bullet : Extreme value; \blacksquare : exponential; \blacklozenge : rectangular; \blacktriangle : logistic; \blacktriangledown : Laplace; --- : gamma-distributed noise level; $\text{-}\cdot\text{-}$: Poisson-distributed noise level; $\text{--}\cdot\text{--}$: inverse Gaussian-distributed noise level; $\text{-}\cdot\cdot\text{-}$: Weibull-distributed noise level; \cdots : gamma-distributed noise-intensity time fluctuation.

parameters are sensitive to relatively short-duration high-energy-level acoustical signals and the intermittent changes in the traffic condition. The question of whether there is any relationship between these two sets of statistical results is left to further investigation.

In order to find out the type of distribution that may be able to describe the general forms of present noise-level distributions, the present (s, k) relationships are compared with

those of the Pearson's family on the (β_1, β_2) chart.¹⁸ Fine details of the present short-time noise level distributions are ignored. Though there are some other families of frequency curves,²¹ the Pearson's family appears to be the most fundamental and is thus adopted in the present investigation.

Figure 7(a), (b), and (c) show the (β_1, β_2) chart¹⁸ and the (s, k) relationships of the interrupted flow, free-flow, and partially continuous flow, respectively. These (s, k) relationships are obtained by regression. The 95% confidence boundaries are also included here to provide for the data scattering. The Roman letters in Fig. 7 represent the types of Pearson distributions and the alphabets in parentheses denote the general shapes of the distributions described in Elderton and Johnson.²¹ The thin solid lines are the boundaries separating different regions of Pearson distributions. For further details of the (β_1, β_2) chart and the properties of the Pearson distributions, please see Refs. 18 and 21.

Figure 7(a) suggests that a majority of the noise-level distributions obtained under the interrupted traffic-flow condition can be approximated by the Pearson type I curve which takes the form

$$f(x) \sim \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2}, \quad (2)$$

where $m_i s$ and $a_i s$ are constants to be evaluated from the sample $\{x\}$. The shape of such distribution depends on the values of $m_i s$ and $a_i s$. For $s^2 \geq 2$, which is only found in the interrupted flow cases, the measured noise-level distribution remains close to the Pearson type I curve but there is a higher chance to have a J-shaped one (not shown here). For very small s , the distribution may also take the form of a U-shaped Pearson type I curve. When $s=0$, it becomes a Pearson type II distribution [$m_1 = m_2$ and $a_1 = a_2$ in Eq. (2)]. An example of a U-shaped noise-level distribution (which appears quite rectangular) has been shown in Fig. 5(e). It is obtained at site E with an interrupted traffic-flow condition. However, this region is very narrow and thus is not discussed further.

Figure 7(b) suggests that while it is possible to find noise-level distributions which take the forms of the Pearson type III, IV, V, and VI under the free-flow condition, the probability of having a bell-shaped Pearson type I distribution is still higher than those of the others. Some symmetrical Pearson type VII noise-level distributions can also be expected for small s . Unlike the case with interrupted traffic, a U-shaped Pearson type II distribution is not likely to be found.

The increase of k with s is relatively rapid under the partially continuous-flow condition, as shown in Fig. 7(c). Under this type of mixed-flow condition, no dominating distribution type exists and the noise-level distribution may take the form of Pearson type I, II, III, IV, V, or VI, depending on s . For $s^2 > 0.7$, Fig. 7(c) suggests the use of the Pearson type IV curve

$$f(x) \sim \left(1 + \frac{x^2}{a_1^2}\right)^{-m_1} e^{-m_2 \tan^{-1}(x/a_1)}, \quad (3)$$

to approximate the noise-level distribution, while for $s^2 < 0.3$, the bell-shaped Pearson type I curve becomes more

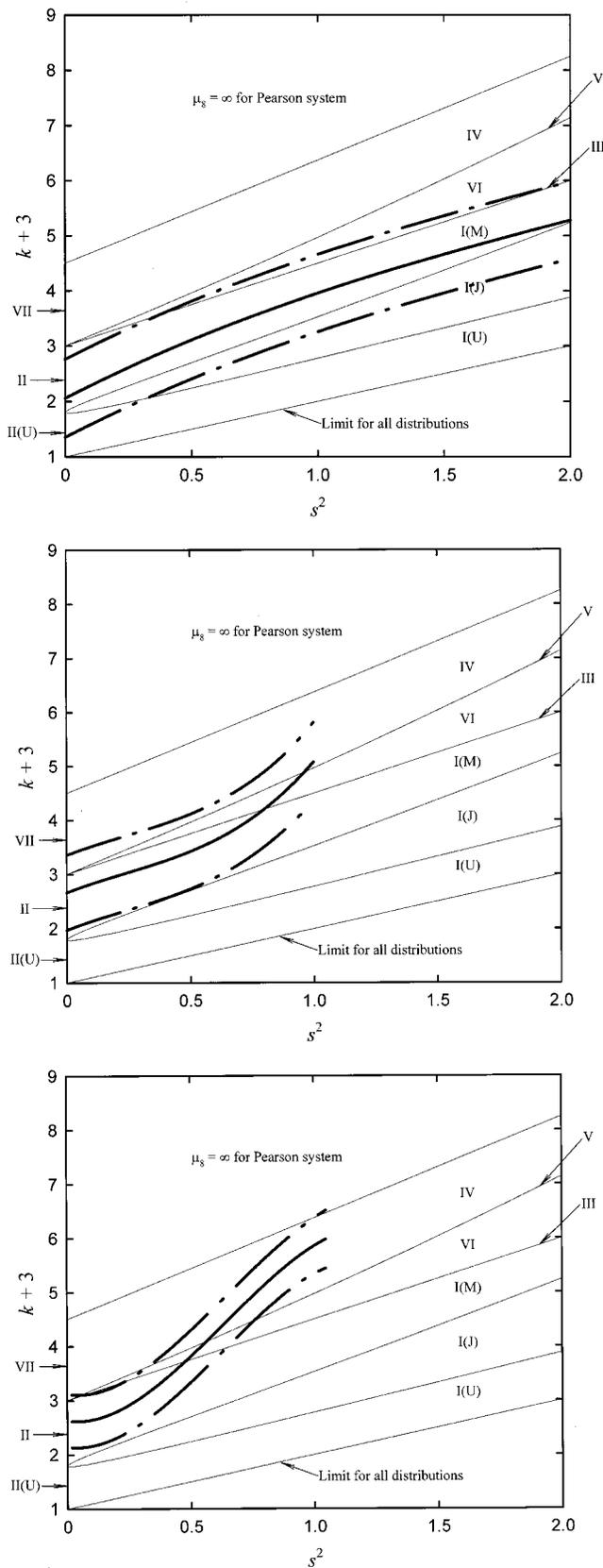


FIG. 7. Comparison between short-time level distributions and Pearson distributions. (a) interrupted traffic flow; (b) free flow; (c) partially continuous traffic flow. —: Regression line; - - -: 95% confidence boundaries. M: bell-shaped; U: U-shaped; J: J-shaped; Roman letters: types of Pearson distributions; thin solid lines: boundaries separating regions of Pearson distributions (Ref. 19).

appropriate. The noise-level distributions obtained under this traffic condition are usually bell-shaped (a property of the relevant Pearson curves²¹). The Pearson type VII and the U-shaped type II distributions do not appear to be relevant.

III. CONCLUSIONS

Simultaneous indoor and outdoor noise measurements with video recording were carried out at seven residential flats affected by traffic noise in the present study. The traffic-flow patterns on the main noise-producing roads can basically be divided into the free-flow and interrupted-flow types. The time variations of noise levels, their overall statistics, and the short-duration statistics are discussed. The short-duration statistics are also compared with those of the existing frequency distributions.

Upward noise-level spikes appear frequently in the interrupted-flow cases, giving rise to positive skewness in the noise-level statistics. For the free-flow cases, the noise level fluctuates about the equivalent sound-pressure level with occasional downward spikes due to short-duration lulls in the traffic flow. The present overall statistical results suggest that Gaussian noise-level distribution can only be found when the traffic volume is low and the flow is not interrupted very much. For the interrupted-flow case, the overall noise-level distributions show some characteristics of the gamma distribution.

Five-minute short-duration noise-level statistics are studied in the present investigation for a deeper understanding of the nature of the traffic-noise fluctuations. The relationships between skewness and kurtosis of the measured short-duration noise-level distributions are far from the Gaussian predictions. The results show some possibilities of having a gamma-distributed noise intensity or a Weibull-distributed noise-level time fluctuation under the free-flow condition. Skewness of the noise-level distribution is predominantly positive for the interrupted traffic-flow cases. No commonly used statistical functions can describe the corresponding relationship between skewness and kurtosis, though a gamma-like overall noise-level distribution is observed.

The skewness–kurtosis relationships of the present short-duration noise-level distributions suggest that a majority of these distributions, especially those for the interrupted flow, can be approximated by the Pearson type I curve instead of the Pearson type III or Gaussian distribution suggested in the existing literature. Gaussian noise-level distributions are likely to be found only under the free-flow traffic condition.

ACKNOWLEDGMENTS

This work was partially supported by a research grant from the Research Committee, The Hong Kong Polytechnic University. The constructive efforts and suggestions from the reviewers are greatly acknowledged.

- ¹N. W. M. Ko, "Traffic noise in a high-rise city," *Appl. Acoust.* **11**, 225–239 (1978).
- ²T. Ishiyama, K. Tateishi, and T. Arai, "An analysis of traffic noise propagation around main roads in Tokyo," *Noise Control Eng. J.* **36**, 65–72 (1991).
- ³F. M. Wiener, C. I. Malme, and C. M. Gogos, "Sound propagation in urban areas," *J. Acoust. Soc. Am.* **37**, 738–747 (1965).
- ⁴I. D. Griffiths, F. J. Langdon, and M. A. Swan, "Subjective effects of traffic noise exposure: reliability and seasonal effects," *J. Sound Vib.* **71**, 227–240 (1980).
- ⁵J. Lambert, F. Simonnet, and M. Vallet, "Patterns of behaviour in dwellings exposed to road traffic noise," *J. Sound Vib.* **92**, 159–172 (1984).
- ⁶W. E. Scholes, "Traffic noise criteria," *Appl. Acoust.* **3**, 1–21 (1970).
- ⁷F. J. Langdon and I. D. Griffiths, "Subjective effects of traffic noise exposure, II: Comparisons of noise indices, response scales, and the effects of changes in noise levels," *J. Sound Vib.* **83**, 171–180 (1982).
- ⁸G. B. Cannelli, "Traffic noise pollution in Rome," *Appl. Acoust.* **7**, 103–115 (1974).
- ⁹D. Chakrabarty, S. C. Santra, A. Mukherjee, B. Roy, and P. Das, "Status of road traffic noise in Calcutta metropolis, India," *J. Acoust. Soc. Am.* **101**, 943–949 (1997).
- ¹⁰X. Liu, "Analysis of the acoustical environment of urban dwellings," *Appl. Acoust.* **29**, 273–287 (1990).
- ¹¹M. Tamashita and K. Yamamoto, "Scale model experiments for the prediction of road traffic noise and design of noise control facilities," *Appl. Acoust.* **31**, 185–196 (1990).
- ¹²Department of Transport Welsh Office, *Calculation of Road Traffic Noise* (Her Majesty's Stationary Office, London, 1988).
- ¹³J. Foxon and F. J. Pearson, "A statistical model of traffic noise," *Appl. Acoust.* **1**, 175–188 (1968).
- ¹⁴U. J. Kurze, "Statistics of road traffic noise," *J. Sound Vib.* **18**, 171–195 (1971).
- ¹⁵C. G. Don and I. G. Rees, "Road traffic sound level distributions," *J. Sound Vib.* **100**, 41–53 (1985).
- ¹⁶N. W. M. Ko, "Indoor traffic noise in a high-rise city," *J. Sound Vib.* **47**, 599–601 (1976).
- ¹⁷P. Kokowski and R. Makarewicz, "Interrupted traffic noise," *J. Acoust. Soc. Am.* **101**, 360–371 (1997).
- ¹⁸A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics. Vol.1: Distribution Theory* (Charles Griffin, London, 1987).
- ¹⁹D. R. Johnson and E. G. Saunders, "The evaluation of noise from freely flowing road traffic," *J. Sound Vib.* **7**, 287–309 (1968).
- ²⁰R. Baldwin, *An Environmental Assessment for Existing Office Buildings* (Building Research Establishment, Watford, 1993).
- ²¹W. P. Elderton and N. L. Johnson, *Systems of Frequency Curves* (Cambridge University Press, Cambridge, 1969).
- ²²W. Weibull, "A statistical distribution function of wide applicability," *Trans. ASME, J. Appl. Mech.* **18**, 293–297 (1951).

A pulsed phase-sensitive technique for acoustical measurements

Vladimir F. Kozhevnikov

University of Utah, Salt Lake City, Utah 84112

Dmitrii I. Arnold

Russian Research Center "Kurchatov Institute," Moscow 123182, Russia

Matthew E. Briggs

University of Utah, Salt Lake City, Utah 84112

Salim P. Naurzakov

Russian Research Center, "Kurchatov Institute," Moscow 123182, Russia

John M. Viner and P. Craig Taylor

University of Utah, Salt Lake City, Utah 84112

(Received 5 January 1999; accepted for publication 8 September 1999)

The pulsed phase-sensitive (PPS) technique for measurements of sound velocity and attenuation in fluids and solids is reviewed. With this technique, which uses a cycle-overlap principle, a time delay is measured between any two acoustical pulses transmitted through a sample or reflected from its boundaries. A current realization of the technique allows one to resolve the time-delay variation down to 0.1 ns. Thus at relative measurements with the PPS technique, precise sound velocity data can be obtained for samples of small thickness (about 1 mm). The technique is versatile and can also be used for accurate absolute measurements. The technique is most advantageous for studies of samples with high attenuation, in particular near phase transitions and at high temperatures. The technique also allows one to measure sound attenuation with moderate accuracy. An application of the technique that employs narrow-band radio frequency (rf) bursts for relative measurements of the phase sound velocity is considered in detail. The technique is applied to studies of liquid alkali metals and mercury at temperatures up to 2100 K and pressures up to 200 MPa. As a verification of the capabilities of the technique, new results are presented on sound attenuation in mercury in the metal–nonmetal transition region. A table for sound velocity in mercury at temperatures from 550 to 1900 K and pressures up to 190 MPa is presented in the Appendix. © 1999 Acoustical Society of America. [S0001-4966(99)06412-7]

PACS numbers: 43.58.Vb, 43.35.Yb, 43.35.Bf [SLE]

INTRODUCTION

The well-known power of physical acoustics to study the properties of matter has resulted in the development of numerous techniques for measurements of sound velocity and attenuation in fluids and solids. Some of the techniques have found important industrial applications. The appropriate theory, principal methods, and corresponding experimental details can be found in comprehensive reviews and monographs, such as McSkimin,¹ Truell *et al.*,² Papadakis,³ Kolesnikov,⁴ Trusler,⁵ and Cantrell and Yost.⁶

The pulsed phase-sensitive (PPS) technique was originally developed for accurate measurements of sound velocity in liquid metals at high temperatures ($T \leq 2500$ K) and high pressures ($P \leq 300$ MPa).⁷ It will be seen, however, that the capabilities of the PPS technique are not restricted to this field.

The main difficulties with experiments at such conditions are well known. They are caused by small sample sizes, utilization of rather long and stationary buffer rods, and large sound attenuation in the sample and buffer rods. Although some techniques, such as pulse-echo overlap,^{8,9} pulse superposition,¹⁰ and phase comparison,¹¹ allow very high accuracy for the time-delay measurements, and some of them are adaptable for use with the buffer rods,^{9,12,13} the high-

temperature and high-pressure conditions are not favorable for the application of these techniques. Other accurate techniques, such as the pulse-echo^{2,14} and pulse interferometric techniques,^{1,15,16} are also inappropriate for measurements at simultaneously high temperatures and pressures. (Recently, M. Hensel¹⁷ conducted measurements of sound velocity in expanded mercury with the pulse-echo technique. Reliable data were obtained at temperatures below 1500 K.)

Only a pulse transmission-echo technique¹⁸ facilitates the measurements at both high temperatures and high pressures. This technique accommodates the buffer rods and utilizes a pulse transmitted through the sample in a single pass. Therefore this technique allows one to reach the highest temperatures where most samples exhibit the greatest attenuation. This pulse transmission–echo technique was used by Suzuki *et al.*¹⁹ in the pioneering high-temperature, high-pressure study of sound velocity in mercury.

A schematic diagram of the cell used in this technique is shown in Fig. 1. The sample space is the gap between the buffer rods. Three time delay intervals are measured: the delay for the echoes in each of the rods and the delay for the signal transmitted through the entire cell. The difference between the time required for the signal to travel through the cell, and the average of the echo times, represents the time required for the signal to traverse the sample. Similarly, the

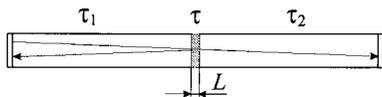


FIG. 1. Schematic diagram of a cell for acoustical measurements at high temperatures.

difference between the relative levels in decibels of the transmitted signal and the average of the echo signals yields attenuation due to the sample, provided that corrections are made to account for the sample–buffer rod interfaces. The velocity and coefficient of attenuation are then easily computed when the sample length is known.

A disadvantage of the original version of the transmission–echo technique is that the accuracy of the time-delay measurements is determined by the duration of the leading edge of the rectified rf pulses. The leading edge always occurs over a number of rf periods depending on the resonance properties of the transducer with corresponding electrical circuit, and the acoustical characteristics of the buffer rod.⁵ An uncertainty of 10%–20% is a reasonable estimate of the error in the speed measurements with this technique for samples of about 1-mm length.¹⁸ The PPS technique combines the advantages of the transmission–echo method with the high accuracy inherent in the phase comparison techniques.

The principles of the PPS technique and corresponding experimental apparatus are described in the next section. Applications of the technique are presented in Sec. II. Section III describes possible cell designs for acoustical measurements in liquids at high temperatures, and Sec. IV presents the summary. An Appendix contains tabulated data obtained for sound velocity in mercury at sub- and supercritical temperatures and pressures.

I. THE TECHNIQUE

Below we describe the technique for using the cell schematically shown in Fig. 1 with moderately narrow-band rf bursts, i.e., for phase sound velocity measurements. In such a case, the use of transducers of different resonance frequencies allows one to study the frequency dependence of the acoustical properties. However, these conditions are not mandatory. The technique is applicable without buffer rods, as well as with either one or two transducers. In addition, broadband pulses can be employed to measure the group velocity of sound.

A. Basic principles of the time-delay measurements

For high-temperature measurements, an ultrasonic probe pulse is introduced into the sample through a buffer rod. The sound propagation times in the buffer rods (including the small delays in the transducer–rod bonds⁷), τ_1 and τ_2 , are unknown. However, these times can be eliminated provided that two time delays, Δt_1 and Δt_2 , between a transmitted and reflected signal are measured, by applying the probe pulse first to one and then to the other transducer.

$$\Delta t_1 = \tau_1 + \tau + \tau_2 - 2\tau_1, \quad (1)$$

$$\Delta t_2 = \tau_2 + \tau + \tau_1 - 2\tau_2. \quad (2)$$

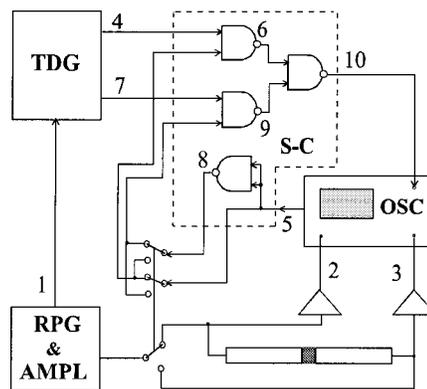


FIG. 2. Block diagram of the PPS apparatus at the University of Utah. TDG is the time delay generator; RPG is the generator of the coherent radio-frequency pulses; S-C is the synchronizer-commutator; OSC is the oscilloscope. The numbers correspond to the timing diagrams of Fig. 3.

The sound propagation time τ through the sample and the velocity c are

$$\tau = (\Delta t_1 + \Delta t_2)/2, \quad (3)$$

$$c = L/\tau, \quad (4)$$

where L is the length of the sample.

The intervals Δt are measured by delaying the sweep of the oscilloscope with respect to a basic synchronization pulse that periodically initiates the coherent rf probe pulses. The choice of the pulse repetition rate is determined by the time required for the echoes in the buffer rods to be attenuated before the next pulse is applied. The sweep of the oscilloscope is delayed for different times for alternate pulses applied to the two oscilloscope channels. (For example, if the two delays are $2\tau_1$ and $\tau_1 + \tau + \tau_2$, then the reflected and transmitted pulses are displayed, respectively.) In this way one may achieve a coincidence of the phases of the corresponding rf cycles for these two pulses (the cycle overlap or the cycle-for-cycle matching). To avoid undesirable interference effects, the time delay should be set using the rf cycles close to the onset of the first transmitted and reflected sound pulses where the magnitudes of the oscillations are nearly constant and there is no overlap of the waves reflected from the sample boundaries.

The block and time diagrams for the apparatus at the University of Utah are shown in Figs. 2 and 3, respectively. A MATEC 5100 gating modulator with plug-in is used to produce coherent rf pulses synchronously with the basic synchronization pulses shown on line 1 of Fig. 3. A Wavetek 3006 generator with frequency instability less than 0.2 ppm/hr is used as a continuous wave (cw) source for the gating modulator. (For accurate measurements of the frequency dependence of the acoustical properties, a counter to control the frequency of the source is added.) The reflected (line 2) and transmitted (line 3) rf pulses are amplified by the broadband MATEC 625 receivers, and then applied to the inputs of a dual-channel one-beam oscilloscope (Tektronix-475A), working in the alternating signal mode with external triggering.

A calibrated time-delay generator (EG&G Instruments, Inc., Model 9650A) with accuracy $0.3 \text{ ns} + 3 \cdot 10^{-7}$

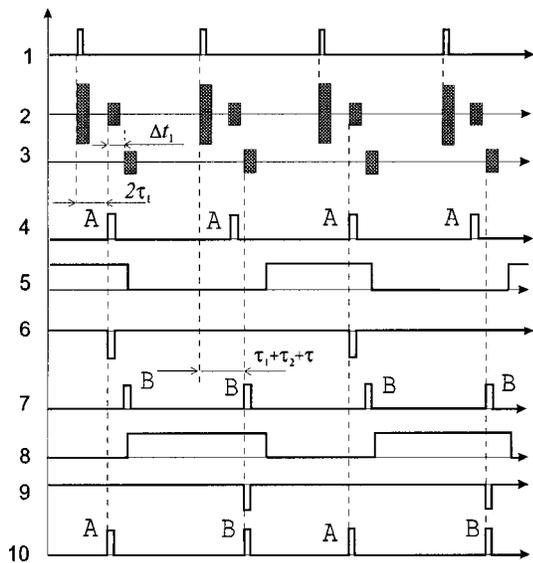


FIG. 3. Timing diagrams for the apparatus shown in Figs. 2 and 4. See text for details.

\times (delay), and jitter $50 \text{ ps} + 10^{-8} \times (\text{delay})$ produces synchronization pulses *A* (line 4) and *B* (line 7) delayed with respect to the pulse (line 1) by times close to $2\tau_1$ and $\tau_1 + \tau + \tau_2$, respectively. These synchronization pulses drive a home-made synchronizer-commutator (constructed from NAND elements from microcircuits 74LS00N) which produces a sequence of pairs of synchronization pulses *A* and *B* (line 10) that trigger the oscilloscope sweep. These pulses are unambiguously synchronized with the corresponding oscilloscope channels. Specifically, pulse *A* appears only when the first channel is active, and *B*—only when the second channel is active. The synchronization is achieved with the aid of a gate pulse (line 5). This gate pulse is an internal pulse from the oscilloscope, which switches its display between the channels at the end of each sweep in the alternating mode.

Thus on the oscilloscope screen an operator sees two pulses, each of which can be independently shifted on the horizontal axis (by variation of the corresponding delay of the delay generator), and amplified or attenuated on the vertical axis. This feature of the PPS technique distinguishes it from the other methods for time-delay measurements. The sweep rates are limited only by the oscilloscope. When the two traces are matched cycle-for-cycle, the difference between the time delays of the two channels (between *A* and *B* pulses in Fig. 3) of the delay generator yields the interval Δt (see Sec. IC for additional elucidation).

The PPS technique can be realized with various schemes. A principal element of all these schemes is the synchronizer-commutator that provides the sequence of pulses, shown on line 10 of Fig. 3, to synchronize the corresponding channels of the oscilloscope.

A diagram of the apparatus at the Kurchatov Institute is shown in Fig. 4. In this apparatus two one-channel delay generators, И1–8 (Russian standard), with accuracy $0.5 \text{ ns} + 0.5 \cdot 10^{-7} \times (\text{delay})$, and jitter $0.4 \text{ ns} + 10^{-8} \times (\text{delay})$ are used. These are connected in series. The delayed output of the first generator TDG1 drives the second generator TDG2, which allows TDG2 to read Δt directly. The generator for

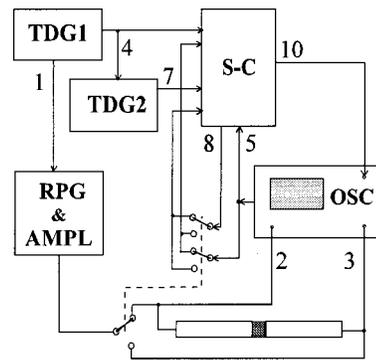


FIG. 4. Block diagram of the PPS apparatus at the Kurchatov Institute. The labels are defined in the caption to Fig. 2. The numbers correspond to the timing diagram of Fig. 3.

the coherent rf pulses is home made (courtesy of E. Grodzinskii). An oscilloscope C 1–70 with a dual-channel amplifier Я40–1101 is used. The timing diagrams for this apparatus are the same as in Fig. 3.

A third diagram of the PPS apparatus, in which the two time delays are provided by a single one-channel delay generator with the use of its internal clock signal, is described in Ref. 7.

The process of measuring sound velocity is partly automated. The monitoring of the time-delay generator, the recording of original data, and the analysis of these data are performed using a computer. However, the superposition of the rf pulses on the oscilloscope screen is performed by hand. A realization of the PPS technique that employs a digital oscilloscope (or other digital device to control the cycle matching) should allow one to completely automate the process and should provide an additional increase in the sensitivity of the apparatus.

B. Sensitivity and accuracy of the technique

The sensitivity of the technique, or the resolution of the smallest phase difference between the cycles of the transmitted and reflected rf pulses (or other pulses chosen for the measurements) is limited by the jitter in the delay generator and the scope trigger. These jitters result in an apparent trace thickness on the oscilloscope screen. The apparatus in Salt Lake City and Moscow allow one to distinguish the variation of Δt down to 0.1 and 0.5 ns, respectively.

The sensitivity represents a lower bound on the uncertainty of the time measurements, which also depends on other factors of a particular experiment. After necessary corrections of systematic errors (such as, for example, a phase shift caused by reflections from the two surfaces of the bond between the sample and the transducer for the measurements in solids,²⁰ and a possible difference in time delays for the different oscilloscope channels), the accuracy is determined from the dispersion of the individual results obtained at a given experimental point. (At high temperatures the dispersion is mostly determined by the temperature instability of the sample.) For example, in the experiments with mercury^{21–24} where the probe pulse was switched several times between the transducers for each experimental point, the dispersion of the measured Δt was $\pm 1 \text{ ns}$; this value

represents the uncertainty of the time measurements in this work. Drift of the delay pulses, which should also be taken into account to estimate uncertainty in the data, is much smaller than 1 ns.

The high sensitivity in the time-delay measurements allows accurate relative data on sound velocity to be obtained with samples of short lengths. The lower limit for the sample length L is determined by the duration of the leading edge of the rf pulses and by the sound velocity c in the sample: $L \geq NTc/2$, where N is number of rf periods T in the leading edge. Currently the minimum length of the samples used for the measurements with the PPS technique is about 0.7 mm.²⁴

C. Relative measurements of sound velocity

In principle, Eqs. (1)–(4) describe the procedure for making the absolute measurements of sound velocity. However, problems associated with selection of the corresponding cycles for the chosen rf pulses, as well as delays in the electrical circuits, prevent a simple realization of this approach. (Correct selection of the corresponding rf cycles can be accomplished by employing different path lengths, i.e., by measurements with different echoes,^{1,3,20} or with samples of different lengths.^{15,25}) These problems are obviated by making relative measurements. For relative measurements a change of the propagation time $\Delta\tau = \tau - \tau_r$ is measured as a function of variation in the environment with respect to a reference point. The parameter τ_r is an effective time delay measured at this point. The sound velocity c_r , and therefore the true value of the time delay in the reference point $\tau_r^0 = L_r/c_r$ (L_r is the length of the sample at the temperature of the reference point), must be known from the literature or from another experiment. Note that the absolute measurements at the reference point can also be performed with the PPS apparatus. This point is usually chosen at more or less normal conditions, when sound attenuation is not large and the high uniformity of a temperature field is readily achievable over relatively large samples (with length on the order of 1 cm, for example). Probably the best way to make absolute measurements at the reference point is to use the PPS apparatus in one of the classical modes, which have been specifically developed for precise measurements under such conditions (see Sec. 1F below).

Representative oscilloscope pictures with reflected and transmitted signals for the reference point (which is often the starting point of the experiment) are shown in Fig. 5. There are no delays at either of the delay generator channels for the signals shown in Fig. 5(a) (zero delay is set for A and B synchronization pulses in the diagram of Fig. 3). In Fig. 5(b) the time delays of the synchronization pulses are set so as to achieve an approximate alignment of the transmitted and reflected rf pulses. Finally, one chooses the rf cycles which will be used in all the further measurements [19th cycles in Fig. 5(b)] and displays the region around them on an expanded scale ($\times 10$). Measurements are made using this expanded scale. One corrects the delay of the B synchronization pulse to achieve the cycle-for-cycle matching with desirable resolution [approximately 0.1 ns for the signals in Fig. 5(c)], and the interval Δt_1 is determined as described in

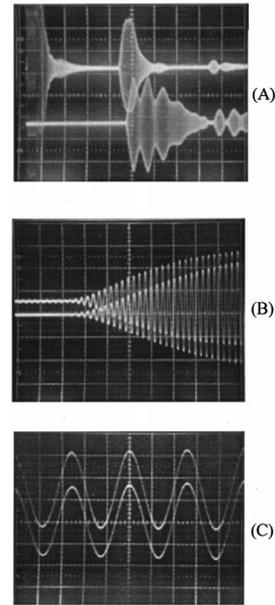


FIG. 5. Oscilloscope traces of the ultrasonic pulses for a brass cell filled with water. The length of each rod was 85 mm and the diameter was 14 mm. The length of the sample was about 2 mm. The frequency of the sound was 7.7 MHz. The upper traces represent the reflected signal, and the lower traces, the transmitted one. No matching network was used. The sweep rates are 10, 0.5, and 0.05 $\mu\text{s}/\text{div}$ for the traces in (a), (b), and (c), respectively. The sound reflections within the sample are seen in the lower trace of (a).

Sec. 1A. The interval Δt_2 is measured in the same way, and $\tau = \tau_r$ is computed using Eq. (3). The measured time τ_r differs from the true sound propagation time $\tau_r^0 = L_r/c_r$. This difference $d \equiv \tau_r^0 - \tau_r = \delta t + nT/2$, where δt represents a time delay in the electrical circuits; T is the period of sound oscillation, and n is an integer, which counts the mismatch of the rf cycles chosen for the measurements; and the coefficient $1/2$ comes from Eq. (3).

All further measurements must be conducted using the same rf cycles. That is, the experiment must be monitored in such a way, which allows one to identify these cycles on the oscilloscope screen. In practice, this identification is readily achieved if the variation in the sound propagation time between neighboring experimental points is limited to two or three periods of the rf oscillations. In regions where the sound velocity varies rapidly, such as near a critical point, it is better to keep the time variation to within one period. While running the experiment, in regions with weak sound dispersion one can check for a lost cycle by turning the magnifier off and looking at the onsets of the pulses [Fig. 5(b)]. If there are no phase errors, then the coincidence of the leading edges remains nearly the same as it was in the beginning. It should be stressed that for relative measurements the strict correspondence of the cycles (or exact alignment of the leading edges) is not of primary importance. It is important to perform all measurements with the same cycles chosen at the reference point.

The sound velocity is calculated from the formula

$$c = L/(\tau_r^0 + \Delta\tau) = L_0[1 + \alpha(T - T_0)]/(d + \tau), \quad (5)$$

where L_0 is the sample length at temperature T_0 (ordinarily T_0 is room temperature); α is the linear thermal expansion

coefficient of the cell material relative to T_0 ; T is the temperature of the sample. The time τ is the effective propagation time computed using Eq. (3). For determination of the thermal expansion of the sample, the principal parts of the cell (see Sec. ID below) must be made from the same material.

In first-order phase transitions, like the liquid–gas transition, the speed of sound changes discontinuously, and it is essentially impossible to follow the phases of the signals. After such a transition a new coincidence of the phases must be established. The denominator in Eq. (5), $(\tau_r^0 + \Delta\tau)$, will now differ from the true value of the propagation time by $mT/2$, where m is another integer. The matching can be corrected in different ways. One way^{21,22} is to obtain data by passing around the critical point. Using this method one finds^{21,22} that m is 0, or ± 1 .

D. Sample length

Correct measurement of the sample length is obviously an important component of the acoustical measurements. While it has been discussed for solids (see Ref. 2, for example), we will briefly consider this issue for fluids.

Commonly, a ringlike spacer with flat and parallel ends confines the sample length. An elastic element of the cell (spring) presses the rod(s) to the spacer, insuring that the spacer and sample lengths are identical (see Ref. 12, for example). The thermal expansion of the sample is then determined by knowledge of the thermal expansion of the spacer.

In the experiments at high temperature and high pressure, the buffer rods are usually hermetically sealed in the cell body. Such construction significantly complicates the use of an elastic element in the cell. The use of a spacer is possible provided that the cell components (such as the cell body, the buffer rods, and the spacer) that contribute to the thermal expansion are made strictly from the same material (including the crystalline orientation if single crystal buffer rods are employed). However, the spacer is not recommended for highly attenuating samples, because parasitic signals through the spacer are not negligible any more. In such a case the requirement of the identical materials allows one to remove the spacer and to calculate the sample thermal expansion as it is done in Eq. (5). The length L_0 can be determined from calibration measurements of the sound velocity versus temperature with a sample for which the temperature dependence of the sound velocity is well known; or from direct measurements on the outer ends of the buffer rods.

For example, our experiments with mercury^{21–24} have been carried out with “spacer-less” cells (the cell design is described in Sec. III below). The change of position of the inserted buffer rod with respect to the state when it is in contact with the other rod was measured with a micrometer. In different experiments the length L_0 varied from 0.7 to 2.6 mm. The uncertainty in the length data was within $\pm(5–7)$ μm . (The sources of this uncertainty are nonparallelism in the ends of the buffer rods and in the axes of the inserted buffer rod and the cell body.) The sound velocity data obtained with the different sample lengths and the cells made

of different materials (molybdenum and niobium) agree well with one another as well as with the data on mercury of Spetzler and Meyer¹² (obtained by the pulse superposition technique with a spacer of 8 mm-length). A value of 10 μm was taken as an upper limit on the uncertainty for the sample length at room temperature. Such an uncertainty in the sample length was sufficient for the purposes of the experiments;^{21–24} however, it can be improved if necessary.

E. Uncertainties in the sound velocity data

Equation (5) yields the following formula for an estimate of the fractional uncertainty δc of the sound velocity data at the relative measurements,

$$\begin{aligned} \delta c \equiv \Delta c/c \approx & \{ [\delta c_r (1 + \Delta\tau/\tau_r)^{-1}]^2 + (\delta\tau)^2 \\ & + [\delta L_0 (1 + \tau_r/\Delta\tau)^{-1}]^2 + [\alpha(T - T_0)\delta\alpha]^2 \\ & + [\alpha T \delta T]^2 \}^{1/2}, \end{aligned} \quad (6)$$

where Δc is the absolute error of the velocity c , δc_r , is the uncertainty in the reference point, $\delta\tau = \Delta(\Delta\tau)/(\tau_r + \Delta\tau)$ is the uncertainty in the time-delay measurements $[\Delta(\Delta\tau)$ is the absolute error of the $\Delta\tau$ measurements], δL_0 is the uncertainty in the sample length at room temperature, and $\delta\alpha$ is the uncertainty of the literature data on the thermal expansion coefficient of the cell material, and δT is the uncertainty in the sample temperature T . The last two terms in Eq. (6) are ordinarily significantly smaller than the first three and can be omitted.^{26b}

For example, in the experiments with mercury,²² the data on sound velocity at the reference point were taken from the literature. The uncertainty in this value is on the order of 10^{-3} . The uncertainty of the sample length measured at room temperature is also of the same order. (The length $L \approx 2$ mm.) The error $\Delta(\Delta\tau)$ is 1 ns, and the $\delta\tau$ term in Eq. (6) is on the order of 10^{-4} , which is much smaller than the other main contributions. [The last two terms in Eq. (6) are on the order of 10^{-5} .] Therefore, the uncertainty in the sound velocity δc at low temperatures (not far from the reference point) is dominated by the uncertainty of the literature data for δc_r . On the other hand, at high temperatures ($T > 1500$ K) the uncertainty is dominated by the uncertainty in the length of the sample. Thus the uncertainty in the time measurements essentially does not contribute to the error in the sound velocity data,²² which is less than 0.4%. Tabulated data on sound velocity in mercury obtained with the PPS technique are presented in the Appendix.

F. A test

The technique was first tested with water. An essential advantage of water as a testing substance is a relatively flat maximum in sound velocity near 74 °C. This feature allows one to make accurate measurements of sound velocity near the maximum even with a relatively large uncertainty in the sample temperature.

The sound velocity in pure water at standard pressure is known with very high precision; disagreement of the data of different groups^{27–29} is about 0.5 m/s or 0.03%. Reproducibility of the data obtained in each of these works is much

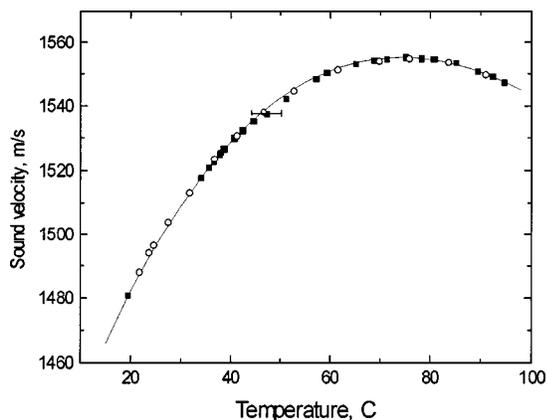


FIG. 6. Sound velocity in water. Solid curve is depicted on data of Del Grosso and Mader (Ref. 28); squares and circles are experimental points measured in the runs with the PPS technique.

better. What is more important for the relative measurements is that the change of sound velocity in these works, for example between points at 74 and 20 °C, is in fact identical: 72.8 m/s in Refs. 27 and 28, and 72.7 m/s in Ref. 29.

Experiments with water using the PPS technique have shown explicitly the principal importance of the use of the same material for all the parts of the cell. Application of nonidentical materials resulted in significant disagreement between the measured and literature data even over less than a 100-deg temperature interval.³⁰

Results of two test runs with distilled water are shown in Fig. 6. (In Refs. 30 and 31 one can find results of other test experiments with water performed with cells made of different materials and with spacers of different length.) A solid curve represents the data of Ref. 28. The relative measurements were performed with respect to a reference point at room temperature, in which sound velocity was taken from Ref. 28. In this work the measurements were performed using an acoustic interferometer with about a 90-mm change of path length. The authors' estimate of the relative accuracy is on the order of 10^{-5} . In our tests we used a horizontally oriented stainless steel cell consisting of a cell body, two inserted buffer rods, and a spacer of 5 mm length. The uncertainty of the spacer length was $\pm 2 \mu\text{m}$. A tungsten/rhenium thermocouple was installed in a groove of the cell body near the middle of the sample space. A total length of the cell assembly was 365 mm. A cylindrical heater of about 80-mm length surrounded the cell; it was centered with respect to the sample space.

To insure that the sample was under steady-state conditions, the delays Δt_1 and Δt_2 were measured several times for each experimental point. The temperature was assumed to be stable, if successive time delays were reproduced with an error not greater than 1 ns. This value therefore represents the maximum error in the time measurements $\Delta(\Delta\tau)$.

In Fig. 6 a horizontal bar denotes the uncertainty in the temperature measurements ($\pm 3^\circ\text{C}$). The uncertainty in the time in Eq. (6) ($\delta\tau \approx 3 \cdot 10^{-4}$) is the principal contribution to the uncertainty of the data on sound velocity. Thus δc in these measurements is about or less than 0.03%.

The data obtained agree well with precise literature data.

The precision of the sound velocity measurements can be estimated from comparison of the results obtained near the maximum with literature data used for the reference point. The maximum sound velocity in Ref. 28 equals 1555.147 m/s; we obtained 1555.0 and 1555.1 m/s. Thus the discrepancy between the reference data and the data obtained in these tests with the PPS technique is close to the discrepancy between the precise literature data from different groups^{28,29} obtained with significantly larger samples.

It should be pointed out, however, that precise absolute measurements are very complicated, and it is a great advantage to make them with long samples.

G. PPS and other techniques

The ability to manipulate the signals independently in both the vertical and horizontal axes of the oscilloscope screen makes the PPS technique very versatile. In fact, many of the known pulse techniques can be realized as particular modes of operation of the PPS apparatus. Several examples are given below.

The pulse superposition technique¹ can be realized if the repetition rate of the probe pulse is increased to achieve an interference of the echoes in the sample.

By removing the buffer rods, one can realize a quasi phase-comparison technique¹¹ (without real interference of the sound waves) by setting the delay of channel for the reflected pulse to zero (line 4 in Fig. 3). A transmitted pulse (this can be a reflected pulse as well) is then delayed to overlap with the probe pulse on the oscilloscope screen. Note that this mode is virtually the same as a transmission variant of the pulse echo-overlap technique.⁸

The double pass technique²⁵ provides one of the aforementioned ways to select the corresponding cycles and exclude δt , and therefore obtain absolute data on sound velocity. The probe pulse is split for two samples of different lengths L_1 and L_2 . The delays are measured with the PPS apparatus for three sample lengths: L_1 , L_2 , and $L_1 - L_2$.

The echo-overlap (PEO) technique^{3,8,9,20} can be realized by switching the oscilloscope to one channel operation mode and removing the inverter from the synchronizer-commutator [NAND element (8) in Fig. 2]. The delays of the synchronization pulses *A* and *B* (lines 4 and 7 of Fig. 3, respectively) are adjusted so as to achieve the rf cycle matching between any two echoes from the sequence of the pulses reflected from the boundaries of the sample. In order to display both echoes, the sweep duration must be less than the interval between the *B* and *A* pulses. As was mentioned, this mode of operation of the PPS apparatus also allows one to perform absolute measurements of sound velocity, and can be used, in particular, to obtain the data for the reference point. It should be noted that in making absolute measurements, especially with small samples, one must be careful with possible delays in the electrical circuits. In particular, a difference in the lengths of the cables connecting the outputs of the delay generator with synchronizer-commutator may result in a systematic error.

The PEO and PPS techniques are similar in accuracy because they are both based on the rf-cycle overlap principle.

However, the PPS technique is easier to perform and is applicable to a wider variety of physical and technical problems.

H. Sound attenuation

Simultaneously with the phase measurements, one can measure the amplitude of the rf pulses, and thus obtain data on sound attenuation.¹⁸ For completeness, we describe this procedure. If the amplitude of the probe signal is A_0 , in the plane-wave approximation the amplitudes of the signals reflected at the first (A_1) and at the second (A_2) buffer rods and that transmitted through the cell (A_t) can be written as

$$\begin{aligned} A_t &= K_t A_0 t_1 T_{1S} t_s T_{S2} t_2, \\ A_1 &= K_1 A_0 t_1^2 R_{1S}, \\ A_2 &= K_2 A_0 t_2^2 R_{2S}, \end{aligned} \quad (7)$$

where t_s , t_1 , t_2 are attenuations of the sound waves in the sample, the first buffer rod, and the second buffer rod, respectively. T_{1S} and T_{S2} are the acoustical transmission coefficients through the corresponding rod/sample boundaries. R_{1S} and R_{2S} are the acoustical reflection coefficients. The transmission and reflection coefficients are determined by the acoustical impedances of the sample and the buffer rods according to well-known formulas.¹ The coefficients K_t , K_1 , and K_2 depend on the geometrical factors of the cell, such as the nonparallel configurations of the rod ends, diffraction of the sound waves, and losses in the buffer-rod/transducer boundaries. These coefficients also depend on the temperature distribution in the buffer rods.

From Eq. (7), the sound attenuation t_s can be written

$$t_s = K A_t (R_{1S} R_{2S})^{1/2} (T_{1S} T_{S2})^{-1} (A_1 A_2)^{-1/2}, \quad (8)$$

where K is a combination of K_t , K_1 , and K_2 . The coefficient of sound attenuation α is found from

$$\alpha = -\ln t_s / L. \quad (9)$$

The sound attenuation for short uniform samples [$L < (D^2/8\lambda)$, where D is the sample diameter, and λ is the wavelength of sound³²] is determined mainly by absorption of sound.

The temperature dependence of the coefficients K in Eqs. (7) and (8) does not allow one to eliminate them completely, even for the relative measurements ($\alpha - \alpha_r$), where α_r is the coefficient of attenuation at a reference point. Other obstacles preventing accurate measurements of the attenuation are the impedances of the sample and the hot ends of the buffer rods. To calculate these impedances, data on density are required, which are not always available. In these acoustical experiments the average sound velocity in the buffer rods can be readily measured, but special efforts are required to determine the sound velocity in the rods near the rod/sample interfaces. These numbers are needed to compute the impedances. An additional problem for the absorption measurements concerns diffraction effects in the buffer rods, which must be carefully taken into account when the plane-wave approximation is not valid.^{1,2,32} For these reasons it is difficult to estimate a priori the accuracy of the sound attenu-

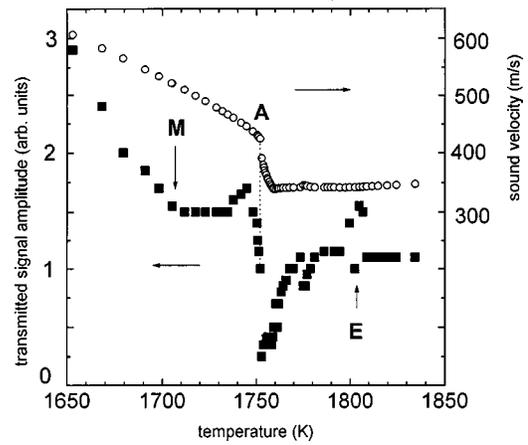


FIG. 7. Sound velocity and transmitted signal amplitude for mercury at 162 MPa, as measured with the PPS technique. Point A corresponds to the liquid-gas transition, E to the prewetting transition, and M to the metal-nonmetal transition. See text for details.

ation data measured in this approach. Nevertheless, even qualitative data on sound absorption are very important, and this method provides the easiest way to obtain such information for fluids at high temperatures and high pressures. In particular, the amplitude measurements have played a crucial role in the interpretation of the acoustical anomalies in mercury^{23,24} and have yielded the most accurate estimate of the critical parameters of this metal.²² Examples of sound attenuation computed using Eqs. (7)–(9) are available in Refs. 18 and 33, and in the next section.

I. Limitations of the technique

A necessary condition for the PPS technique to work is the existence of at least one signal transmitted through the sample. Because the technique allows one to carry out the measurements with very short samples, this condition puts lower limits on the transmission coefficients T_{rs} and T_{sr} [see Eq. (7)], which are determined by the ratio of the sample and the buffer rod impedances.¹ Since the sound velocities for all substances do not vary appreciably in the solid, liquid, and gas phases,³⁴ the aforementioned condition leads to a limitation on the mass density of the sample. To date, the lower limit on density in the experiments that employ the PPS technique with metallic (mostly molybdenum) cells is about 0.7 g/cm³. This value corresponds to approximately double the critical density for cesium²⁶ and to one-tenth of the critical density for mercury.^{26b}

II. EXPERIMENTAL EXAMPLES

The PPS technique has so far been applied mostly to liquid metals, such as cesium²⁶ and mercury,^{21–24} at high temperatures (up to 2100 K) and high pressures (up to 190 MPa).

Figure 7 presents experimental data for the sound velocity in mercury and data for the transmitted signal amplitude measured at a frequency of 10 MHz with a molybdenum cell at a pressure of 162 MPa.

The maximum uncertainty of the results on sound velocity measurements is $\pm 0.4\%$. The temperature was measured

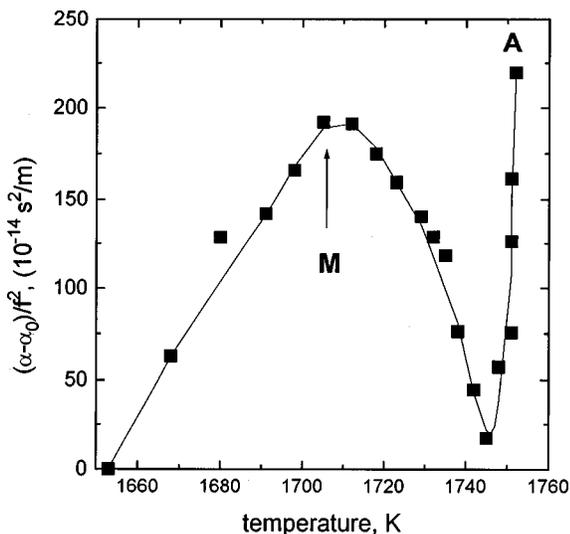


FIG. 8. Sound attenuation in liquid mercury at a pressure 162 MPa. The points labeled A and M are defined in the caption to Fig. 7.

by two tungsten–rhenium thermocouples manufactured from a single pair of wire coils, which were calibrated according to the international standard. The thermocouples were installed near the upper and lower edges of the sample space.²² Absolute error of the temperature measurements varies from 5 at 1000 K to 10 at 2100 K. Discrepancies in the readings of the thermocouples made from different parts of the coils lay within 1 K. Pressure was measured by a Heise manometer with maximum error ± 0.2 MPa.

In Fig. 7 point A marks the liquid–gas transition; E denotes a prewetting transition;²³ and M corresponds to a density close to 9 g/cm^3 , where, according to data on the electronic properties (see Ref. 24 for references), mercury undergoes a metal–nonmetal transition. (A phase diagram for mercury, which was obtained from acoustical measurements using the PPS technique, is available in Refs. 23 and 24.) A similar minimum in the amplitude of the transmitted signal was observed for other isobars at their intersection with the isochore of 9 g/cm^3 (see, for example, Figs. 3 and 4 in Ref. 24). The variation of the attenuation coefficient, as estimated from Eqs. (7)–(9), is shown in Fig. 8. The flat, poorly shaped minimum in the transmitted signal amplitude of Fig. 7 corresponds to a well-formed maximum in the attenuation curve of Fig. 8. This maximum coincides well with the metal–nonmetal transition. An estimate of the error of the data on sound attenuation depends on assumptions made for the temperature distribution along the buffer rods, but for any reasonable assumptions this is less than 30%.

III. THE CELL FOR HIGH-TEMPERATURE MEASUREMENTS

In static experiments at high temperatures and high pressures, a cell with a heater is placed in a high-pressure vessel that is filled with an inert pressure-transferring medium.⁷ Under such conditions, two kinds of cells are currently used in acoustical experiments. These cells contain either two buffer rods inserted in the cell body^{7,17,19,33} or just one inserted rod. In the latter case, the second rod is a part of the cell

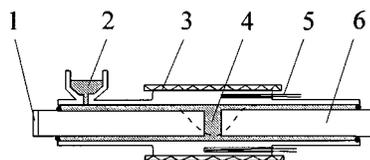


FIG. 9. Schematic diagram of a cell with two inserted buffer rods. 1 is the transducer; 2 is the sample reservoir; 3 is the heater; 4 is the sample space; 5 is the thermocouple; 6 is the buffer rod. A dashed line depicts an approximate meniscus shape at boiling temperature. See text for details.

body.^{21,22} These designs are schematically shown in Figs. 8 and 9. A cell with an inserted, or adjustable, rod always has a finite space between the rod and the cell body. (In Figs. 8 and 9 this space is magnified for clarity.) The cells with two adjustable rods, or “two-rod” cell (Fig. 8), are used both in horizontal and vertical positions. The “one-rod” cells (Fig. 9) are used vertically. We compare the advantages and disadvantages of these two cells.

The approximate symmetry of the temperature distribution along a horizontal two-rod cell is advantageous in some situations, but such symmetry does not play a serious role for acoustical measurements because the contributions of the rods either cancel each other or can be subtracted independent of the temperature distribution in the rods. A disadvantage of the horizontal orientation is an inevitable vertical temperature gradient over the sample (at normal gravity), which becomes crucially important as one approaches the liquid–gas coexistence curve. The dashed line in Fig. 8 shows an approximate meniscus shape at the boiling point for a symmetrical longitudinal temperature distribution. The extreme instability of such a meniscus and the corresponding strong convection flows lead to poor results in this region. In particular, this geometry leads to a significant increase in the transition width (the temperature interval between the last point in the liquid phase and the first point in the gas phase). With the PPS technique we failed to perform measurements within an interval of about 50°C near the boiling point because of a smeared picture on the oscilloscope screen. It is possible that the oscillations in the sound velocity observed in gaseous mercury near saturation by Yao *et al.*³³ have the same origin.

The vertical orientation of the cell with two adjustable buffer rods provides a radial symmetry of the temperature distribution over the sample, but it does not improve the

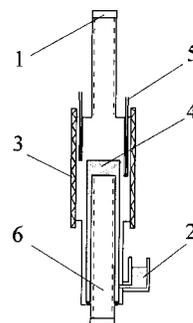


FIG. 10. Schematic diagram of a cell with one inserted buffer rod. The numerical labels are defined in the caption to Fig. 9. The dashed lines on the rods denote a thread. See text for details.

situation near the liquid–gas transition. In this regime, strong convection flows arise because a second meniscus is formed above the sample space. In our experiments the transition width with such a cell was about the same as for the horizontal orientation.

One can take data closer to the coexistence curve using the vertically oriented one-rod cell (Fig. 10). In this case the transition width is determined by a temperature distribution over just the sample space. That is, in contrast to the above-mentioned situations, the peripheral regions do not contribute. For this geometry one can obtain transition width as small as 2 K.²² Note that such a cell allows one to use the buffer rods with a thread on the side surface (see Fig. 10) without influencing the convection flows near the liquid–gas transition. (In a cell with two inserted rods the threaded configuration leads to expansion of the temperature interval over which convection flows exist in the sample space.) The thread surface is used to suppress diffraction effects in the rods.^{1,35} The capability to work near the coexistence of two phases is important in many respects, in particular for comparison with other experimental data.

IV. SUMMARY

The PPS technique combines the advantages of the transmission/echo technique and the phase-comparison methods for sound velocity measurements. The sensitivity of the time-delay measurements is approximately 0.1 ns. Thus using the PPS technique, precise relative data can be ob-

tained over a broad temperature range for samples of short length (currently about 1 mm), and therefore for samples with large attenuation. The PPS technique also allows simultaneous measurements of sound attenuation. The use of narrow-band rf pulses provides measurements of the phase velocity of sound, which allows one to investigate the frequency dependence of the acoustical properties. The technique is versatile; many of the well-known pulse methods can be realized with an apparatus designed for the PPS technique. The PPS technique significantly expands the applicability of precise phase-comparison methods and opens new opportunities for studies of the physical properties of matter, particularly near phase transitions and at high temperatures. A natural limitation of the technique is poor impedance matching for the sound waves at the interfaces between the sample and the buffer rods. A further improvement in the sensitivity of the technique can be achieved by digitizing the process of matching the rf cycles.

Note added in proof: The anomaly of sound absorption in mercury presented in Sec. II has been recently confirmed by H. Kohno and M. Yao, *J. Phys.: Condens. Matter* **11**, 5399 (1999).

ACKNOWLEDGMENTS

The help of Sonny Ko in the construction of the apparatus at the University of Utah is greatly appreciated. This work was partly supported by a grant from the Research

TABLE AI. Sound velocity in mercury.

<i>T</i> , K	<i>P</i> _{sat} , MPa	Sound velocity (m/s) at pressures (MPa)									
		sat	30	60	80	100	120	140	160	170	190
550		1335	1342	1349	1354	1360	1363	1370	1374	1378	1382
600		1309	1316	1324	1329	1335	1340	1346	1350	1352	1358
650	0.1	1283	1291	1299	1304	1310	1316	1322	1326	1329	1335
700	0.3	1257	1265	1275	1280	1285	1291	1288	1303	1306	1311
750	0.6	1229	1238	1249	1254	1260	1267	1274	1280	1283	1287
800	1.0	1200	1211	1223	1229	1236	1243	1250	1257	1260	1263
850	1.9	1170	1183	1195	1202	1211	1218	1225	1233	1236	1239
900	3.0	1139	1153	1167	1175	1183	1191	1199	1208	1212	1215
950	4.6	1108	1122	1136	1146	1155	1163	1171	1181	1186	1189
1000	6.6	1075	1090	1105	1115	1125	1134	1143	1153	1159	1162
1050	9.4	1041	1055	1073	1084	1094	1104	1114	1124	1130	1134
1100	12.8	1006	1020	1039	1051	1062	1072	1082	1093	1099	1106
1150	17.1	970	982	1002	1016	1028	1038	1050	1061	1067	1076
1200	22.2	933	941	963	979	992	1003	1015	1027	1034	1044
1250	28.3	893	894	923	939	953	966	980	993	1000	1011
1300	35.3	854	285	879	897	913	927	944	957	965	977
1350	43.4	812	291	830	852	869	885	903	920	927	940
1400	52.6	768	296	777	802	822	842	860	878	887	901
1450	63.0	723	302	297	747	771	794	813	833	844	860
1500	74.7	676	307	303	686	714	741	763	786	797	813
1550	87.8	627	312	310	309	648	680	707	734	746	762
1600	102.7	577	317	317	315		610	644	675	689	709
1650	119.5	527		323	323	320	526	568	607	624	649
1700	138.7	476		329	329	328	326	475	528	547	580
1750	160.6	426		335	336	335	334	333		460	502
1800				341	342	342	342	342	342		420
1850				346	347	349	349	350	350	352	
1900				352	353	355	356	357	358	359	

APPENDIX

Smoothed data for sound velocity in mercury measured with the PPS technique are presented in Table AI. (See Ref. 22 for the details.) The uncertainty in these data is less than 0.4% (see Sec. IE). The data from a handbook³⁴ are used for computing the saturated pressure P_{sat} below 30 MPa. Critical parameters for mercury (1764 K and 167 MPa) are also determined from the acoustical data. Values of the sound velocity in the abnormal regions of subcritical isobars and near the critical isochore, where the sample was not uniform because of wetting,^{23,24} are excluded from the table. It is interesting to note that, for mercury vapor, the derivative $\partial c/\partial P$ at constant temperature changes sign at supercritical temperatures.

- ¹H. J. McSkimin, in *Physical Acoustics. Principles and Methods*, edited by W. P. Mason (Academic, New York, 1964), Vol. I/A, Chap. 4, pp. 217–332.
- ²R. Truell, C. Elbaum, and B. B. Chick, *Ultrasonic Methods in Solid State Physics* (Academic, New York, 1969).
- ³E. P. Papadakis, in *Physical Acoustics. Principles and Methods*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1976), Vol. XII, Chap. 5, pp. 277–374.
- ⁴A. E. Kolesnikov, *Ultrasonic Measurements* (Standard, Moscow, 1970), in Russian.
- ⁵J. P. M. Trusler, *Physical Acoustics and Metrology of Fluids* (IOP, Adam Hilger, Bristol, 1991).
- ⁶J. H. Cantrell and W. T. Yost, in *Handbook of Acoustics*, edited by M. J. Crocker (Wiley, New York, 1998), Chap. 41, pp. 483–495.
- ⁷D. I. Arnold, A. M. Gordeenko, P. N. Ermilov, V. F. Kozhevnikov, and S. P. Naurzakov, *Prib. Tekh. Eksp.* **5**, 167–170 (1985) [English transl.: *Instrum. Exp. Tech.* **28**, 1173–1177 (1985)].
- ⁸J. E. May, *IRE Natl. Conv. Rec.* **6**, 134–142 (1958).
- ⁹E. Papadakis, *J. Appl. Phys.* **35**, 1474–1482 (1964).
- ¹⁰H. J. McSkimin, *J. Acoust. Soc. Am.* **33**, 12–16 (1961).
- ¹¹B. M. Abraham, Y. Eckstein, J. B. Ketterson, M. Kuchnir, and P. R. Roach, *Phys. Rev. A* **1**, 250–257 (1970).
- ¹²H. A. Spetzler, M. D. Neyer, and Tin Chan, *High Temp.-High Press.* **7**, 481–496 (1975).
- ¹³J. P. Petitet, R. Tufeu, and B. Le Neindre, *Int. J. Thermophys.* **4**, 35–50 (1983).
- ¹⁴B. Chick, G. Anderson, and R. Truell, *J. Acoust. Soc. Am.* **32**, 186–193 (1960).
- ¹⁵Yu. S. Trelin, I. N. Vasil'ev, V. P. Proskurin, and T. A. Tsyganova, *Teplofiz. Vys. Temp.* **4**, 364–368 (1966) [English transl.: *High Temp.* **4**, 352–355 (1966)].
- ¹⁶I. I. Novikov, Yu. S. Trelin, and T. A. Tsyganova, *Teplofiz. Vys. Temp.* **7**, 1220 (1969) [English transl.: *High Temp.* **7**, 1140–1141 (1969)].
- ¹⁷M. Hensel, Thesis, Marburg University, 1993.
- ¹⁸A. W. Nolle and P. W. Sieck, *J. Appl. Phys.* **23**, 888–894 (1952).
- ¹⁹K. Suzuki, M. Inutake, and S. Fujiwaka, IPPJ-310, Institute of Plasma Physics, Nagoya (1977); M. Inutake, K. Suzuki, and S. Fujiwaka, *J. Phys. (Paris)* **40**, C7-685–C7-686 (1979).
- ²⁰E. P. Papadakis, *J. Acoust. Soc. Am.* **42**, 1045–1051 (1967).
- ²¹V. F. Kozhevnikov, D. I. Arnold, and S. P. Naurzakov, *J. Phys.: Condens. Matter* **6**, A249–A254 (1994).
- ²²V. F. Kozhevnikov, D. I. Arnold, and S. P. Naurzakov, *Fluid Phase Equilibria* **125**, 149–157 (1996).
- ²³V. F. Kozhevnikov, D. I. Arnold, S. P. Naurzakov, and E. M. Fisher, *Phys. Rev. Lett.* **78**, 1735–1738 (1997).
- ²⁴V. F. Kozhevnikov, D. I. Arnold, S. P. Naurzakov, and E. M. Fisher, *Fluid Phase Equilibria* **150–151**, 625–632 (1998).
- ²⁵J. R. Cunningham and D. G. Ivey, *J. Appl. Phys.* **27**, 967–974 (1956).
- ²⁶(a) V. F. Kozhevnikov, *Zh. Eksp. Teor. Fiz.* **97**, 541–558 (1989) [English transl.: *Sov. Phys. JETP* **70**, 298–310 (1990)]; (b) V. F. Kozhevnikov, S. P. Naurzakov, and A. P. Senchenkov, *J. Mosc. Phys. Soc.* **1**, 171–197 (1991).
- ²⁷M. Greenspan and C. E. Tschiegg, *J. Acoust. Soc. Am.* **31**, 75–76 (1959).
- ²⁸V. A. Del Grosso and C. W. Mader, *J. Acoust. Soc. Am.* **52**, 1442–1446 (1972).
- ²⁹A. A. Alexandrov and M. S. Trakhtenherz, *Teploenergetika* No. **4**, 57–60 (1973) (in Russian).
- ³⁰O. I. Girnii, Thesis, Moscow Aviation Institute, Moscow, 1988.
- ³¹V. F. Kozhevnikov, D.Sc. dissertation, Kurchatov Institute of Atomic Energy, 1990.
- ³²W. P. Mason, in *Physical Acoustics. Principles and Methods*, edited by W. P. Mason (Academic, New York, 1964), Vol. I/A, Chap. 7, pp. 485–500.
- ³³M. Yao, K. Okada, T. Aoli, and H. Endo, *J. Non-Cryst. Solids* **205–207**, 274–277 (1996).
- ³⁴*Handbook of Physical Quantities*, edited by I. G. Grigoriev and E. Z. Meilikhov (CRC, New York, 1997).
- ³⁵H. J. McSkimin and E. S. Fisher, *J. Appl. Phys.* **31**, 1627–1639 (1960).

Passive wideband cross correlation with differential Doppler compensation using the continuous wavelet transform

Brian G. Ferguson^{a)} and Kam W. Lo

Defence Science and Technology Organisation, P.O. Box 44, Pyrmont NSW, Australia

(Received 16 May 1997; revised 7 July 1999; accepted 30 July 1999)

An estimate of the time difference for the signal emitted by a stationary source to arrive at two spatially separated sensors is given by the time displacement that maximizes the cross-correlation function. For a fast moving source, however, this estimate is found to be in error because the time scales of the received signals are different for the two sensors. The correct time delay can be extracted by evaluating the continuous wavelet transform, which has the same functional form as the wideband cross-ambiguity function. When the signal-to-noise ratio is high, the coordinates of the ambiguity surface's global maximum provide reliable estimates of both the differential time of arrival (or time delay) and the ratio of the time scales of the signals received by the two sensors. The continuous wavelet transform is computed using the one-step chirp z -transform method, the cross-wavelet transform method, and the two-step methods where multirate sampled replicas of the sensor waveforms are cross correlated, or else the sensor waveforms are interpolated using the discrete Fourier transform prior to cross correlation. The latter method is applied to real acoustic data recorded from an orthogonal configuration of three microphones during the low-altitude transit of a jet aircraft. The resulting time delay estimates are used to calculate the variation with time of the azimuth and elevation angles of the aircraft during the transit. © 1999 Acoustical Society of America. [S0001-4966(99)03511-0]

PACS numbers: 43.60.Gk, 43.28.Tc [JCB]

INTRODUCTION

The location of an acoustic source can be estimated from the differences in the arrival times of its emitted signal at spatially separated sensors. Dommermuth and Schiller¹ described the estimation of the trajectory of a constant-velocity aircraft using an array of microphones. When the moving signal source can be considered "quasi-stationary" over short observation time intervals, conventional cross-correlation techniques²⁻⁴ designed for stationary sources can be used to obtain good estimates of the differential time of arrivals (often simply referred to as the time delays). However, if the "quasi-stationary" assumption for the moving source is not valid over the observation period, a relative time-scaling effect (often referred to as differential Doppler) will occur between the signals received by a pair of sensors. This results in a loss of signal coherence⁵⁻⁸ which, in turn, degrades the performance of conventional cross-correlation methods.

Over a short observation interval, the source signal at the output of one sensor can be modeled as a *constant* time-scaled and time-delayed version of the other, which is equivalent to assuming that the time delay of one signal relative to the other varies linearly with time over the short observation interval. If $\tilde{s}(t)$ denotes the source signal at the output of one of the sensors, then the signal at the output of the other sensor is $\tilde{s}[(t-\beta)/\alpha] \equiv \tilde{s}[t-D(t)]$, where α and β are the *constant* relative time scale and time delay between the two signals, respectively, and $D(t)$ is the equivalent *time-varying* relative time delay given by $D(t) = at + b$,

where $a = (\alpha - 1)/\alpha$ and $b = \beta/\alpha$. Thus the estimation of α and β is equivalent to the estimation of a and b . In practice, α and β are estimated simultaneously, unless α is known *a priori* in which case it can be compensated for in the processing. Knapp and Carter⁹ developed an optimal maximum likelihood (ML) correlator that cross correlates the output of one sensor with a time-scaled and time-delayed version of the other sensor output (with appropriate prefiltering). The ML estimates of α and β are obtained by maximizing the cross-correlation output with respect to the time-scale and time-delay variables. This procedure is equivalent to finding the peak of the wideband cross-ambiguity function of the two sensor outputs. The main difficulty with this method is the practical implementation of the time-scaling operation. Nevertheless, a number of techniques for implementing the time-scaling operation have been described in the literature.¹⁰⁻¹⁴

Several suboptimal techniques that bypass the time-scaling operation have been proposed.¹⁵⁻¹⁷ These methods use a piecewise constant approximation to the linear time delay $D(t)$. Specifically, the observation time interval is divided into N small time segments centred at t_n , $1 \leq n \leq N$; $D(t)$ is assumed to be constant at $at_n + b$ over the n th time segment. The technique described in Ref. 15 computes the short-time correlogram over each time segment to obtain a noisy estimate of the "constant" time delay in the corresponding segment. Estimates of a and b are then extracted by exploiting the time evolution of successive time delay measurements using a least squares method. This incoherent technique is more susceptible to large errors from ambiguous peaks than is the ML correlator. In a coherent technique^{16,17} referred to as the deskewed short-time (DST) correlator, the

^{a)}Electronic mail: brian.ferguson@dsto.defence.gov.au

n th correlogram is deskewed (time shifted) by $\hat{a}t_n$ for $1 \leq n \leq N$. The N deskewed correlograms are then summed to form the DST correlator output. The value \hat{a} that maximizes the peak of the DST correlator output provides an estimate of a , and the corresponding position of the peak on the time lag axis gives an estimate of b .

While these suboptimal techniques work effectively when $D(t)$ varies slowly (and linearly) with time and may be more computationally efficient than the optimal ML method, their performance is seriously degraded in the case of a rapid variation with time of $D(t)$ because each time segment becomes too short to generate a useful short-time correlogram for which the time delay information is not smeared. As long as $D(t)$ can be assumed linear, the optimal ML correlator is more robust than the suboptimal techniques. The design of the prefilters for the optimal ML correlator requires *a priori* knowledge of the signal and noise spectra which are often not available, and, without prefiltering, the performance of the cross correlator will be suboptimal.

This paper considers the use of the wideband cross-ambiguity function (or equivalently, with a change of variables, the continuous wavelet transform) for the joint estimation of the relative time scale and time delay, with application to the angular location of aircraft using widely separated passive acoustic sensors. For fast moving sources like jet aircraft flying at low altitudes in proximity to the acoustic sensors, the effect of relative time scaling increases as the separation between the sensors widens which requires the processor to compensate for the differential Doppler effect prior to cross correlating the sensor outputs. Despite the increase in computation time, using a large sensor spacing has the advantages of improved estimation accuracy of the source position⁴ and better spatial resolution of two closely spaced sources.¹⁸

I. SIGNAL MODEL

Figure 1 shows a general geometric configuration for the source and the sensors. The passive array consists of N sensors (microphones) with sensor n located at (x_n, y_n, z_n) , $1 \leq n \leq N$. At time t , a moving source is at a distance $R_n(t)$ from sensor n ; the range, azimuth angle, and elevation angle of the source (relative to the origin) are given by $R(t)$, $\theta(t)$, and $\phi(t)$, respectively. The source emits continuously a broadband acoustic signal $s(t)$.

Due to the propagation delay, the signal received by sensor n at time t was actually emitted from the source at an earlier time $t - D_n(t)$, where $D_n(t)$ is the propagation delay which itself changes with time. Therefore, the signal received by sensor n at *sensor* (or *receiving*) time t can be written as

$$x_n(t) = \rho_n[t - D_n(t)]s[t - D_n(t)], \quad (1)$$

where ρ_n is an attenuation factor which is assumed to vary slowly with time. Since the source is at a distance $R_n[t - D_n(t)]$ from sensor n at time $t - D_n(t)$, the propagation delay $D_n(t)$ is given by

$$D_n(t) = \frac{1}{c}R_n[t - D_n(t)], \quad (2)$$

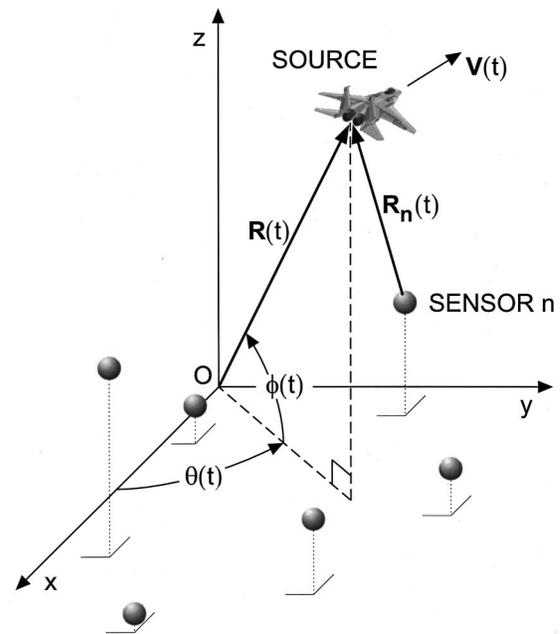


FIG. 1. General configuration of sensors and source.

where c is the (assumed constant) speed of sound in air. Similarly, the signal emitted at *source* (or *emission*) time t_o will be received by sensor n at a later time

$$t_{no} = t_o + R_n(t_o)/c, \quad (3)$$

where the second term is the propagation delay $D_n(t_{no}) = R_n(t_o)/c$.

An approximate expression for the signal received by sensor n over a sufficiently small time interval centred at $t = t_{no}$ can be obtained by using the first three terms of the Taylor series of $D_n(t)$:

$$D_n(t) \approx D_n(t_{no}) + D_n'(t_{no})(t - t_{no}) + \frac{1}{2}D_n''(t_{no})(t - t_{no})^2 \quad \text{for } |t - t_{no}| < \delta, \quad (4)$$

where $D_n'(t_{no})$ and $D_n''(t_{no})$ denote the first and second derivatives of $D_n(t)$ evaluated at $t = t_{no}$, respectively. Differentiating Eq. (2) with respect to t gives

$$D_n'(t) = \frac{1}{c}R_n'[t - D_n(t)][1 - D_n'(t)], \quad (5)$$

$$D_n''(t) = \frac{1}{c}\{R_n''[t - D_n(t)][1 - D_n'(t)]^2 - R_n'[t - D_n(t)]D_n''(t)\}. \quad (6)$$

By setting $t = t_{no}$ in Eqs. (5) and (6) and then using Eq. (3), the following results are obtained after some algebraic manipulation:

$$D_n'(t_{no}) = \frac{R_n'(t_o)}{R_n'(t_o) + c}, \quad (7)$$

$$D_n''(t_{no}) = \frac{c^2 R_n''(t_o)}{[R_n'(t_o) + c]^3}. \quad (8)$$

Substituting Eq. (4) into Eq. (1) and then using Eqs. (7) and (8) and the approximation $\rho_n[t-D_n(t)] \approx \rho_n(t_o)$ for $|t-t_{no}| < \delta$ yields

$$x_n(t) \approx \rho_n(t_o) s[\alpha_n^{-1}(t-t_{no}) - \gamma_n(t-t_{no})^2 + t_o] \quad \text{for } |t-t_{no}| < \delta, \quad (9)$$

where

$$\alpha_n = 1 + \frac{R'_n(t_o)}{c}, \quad (10)$$

$$\gamma_n = \frac{c^2 R''_n(t_o)}{2[R'_n(t_o) + c]^3}. \quad (11)$$

By writing $R_n(t) = \sqrt{\mathbf{R}_n(t) \cdot \mathbf{R}_n(t)}$, where $\mathbf{R}_n(t)$ is the radial vector to the source from sensor n , and noting that $\mathbf{R}'_n(t)$ and $\mathbf{R}''_n(t)$ are the source velocity $\mathbf{V}(t)$ and acceleration $\mathbf{A}(t)$, respectively, it follows that

$$R'_n(t) = v_n(t), \quad (12)$$

$$R''_n(t) = a_n(t) + \frac{u_n^2(t)}{R_n(t)}, \quad (13)$$

where $v_n(t) = \mathbf{V}(t) \cdot \mathbf{R}_n(t) / R_n(t)$ is the radial component of $\mathbf{V}(t)$ along $\mathbf{R}_n(t)$, $u_n(t) = \sqrt{V^2(t) - v_n^2(t)}$ is the cross component of $\mathbf{V}(t)$ along $\mathbf{R}_n(t) \times (\mathbf{V}(t) \times \mathbf{R}_n(t))$ that is orthogonal to $\mathbf{R}_n(t)$, and $a_n(t) = \mathbf{A}(t) \cdot \mathbf{R}_n(t) / R_n(t)$ is the radial component of $\mathbf{A}(t)$ along $\mathbf{R}_n(t)$. Substituting Eqs. (12) and (13) into Eqs. (10) and (11) gives

$$\alpha_n = 1 + \frac{v_n(t_o)}{c}, \quad (14)$$

$$\gamma_n = \frac{a_n(t_o) + u_n^2(t_o) / R_n(t_o)}{2c[1 + v_n(t_o) / c]^3}. \quad (15)$$

By choosing the *sensor* time reference at t_{1o} , that is, replacing t by $t+t_{1o}$, the signals received by sensor 1 and sensor m ($2 \leq m \leq N$) can be written, respectively, as

$$x_1(t+t_{1o}) \approx \rho_1(t_o) s(\alpha_1^{-1}t - \gamma_1 t^2 + t_o) \quad \text{for } |t| < \delta, \quad (16)$$

$$x_m(t+t_{1o}) \approx \rho_m(t_o) s[\alpha_m^{-1}(t - \beta_{m1}) - \gamma_m(t - \beta_{m1})^2 + t_o] \quad \text{for } |t - \beta_{m1}| < \delta, \quad (17)$$

where t is the *sensor* time with reference to t_{1o} and $\beta_{m1} = t_{mo} - t_{1o}$ is the time difference between sensor 1 and sensor m in receiving the source signal $s(t_o)$. In most applications, the second-order term in the argument of s can be ignored over a small time interval. With this assumption, Eqs. (16) and (17) reduce to

$$x_1(t+t_{1o}) \approx \rho_1(t_o) s(\alpha_1^{-1}t + t_o) \quad \text{for } |t| < \delta, \quad (18)$$

$$x_m(t+t_{1o}) \approx \rho_m(t_o) s[\alpha_m^{-1}(t - \beta_{m1}) + t_o] \quad \text{for } |t - \beta_{m1}| < \delta. \quad (19)$$

Finally, by defining $\tilde{s}(t) = \rho_1(t_o) s(\alpha_1^{-1}t + t_o)$ and $\tilde{x}_n(t) = x_n(t+t_{no})$, Eqs. (18) and (19) can be written as

$$\tilde{x}_1(t) \approx \tilde{s}(t) \quad \text{for } |t| < \delta, \quad (20)$$

$$\tilde{x}_m(t) \approx \rho_{m1} \tilde{s}\left(\frac{t - \beta_{m1}}{\alpha_{m1}}\right) \quad \text{for } |t - \beta_{m1}| < \delta, \quad (21)$$

where $\rho_{m1} = \rho_m(t_o) / \rho_1(t_o)$ and $\alpha_{m1} = \alpha_m / \alpha_1$. It can be seen that $\tilde{x}_m(t)$ is a time-scaled and time-delayed replica of $\tilde{x}_1(t)$. (The term time delay is used throughout this paper. When a time delay has a negative value, it actually represents a time advance.) For this reason, α_{m1} and β_{m1} are referred to as the relative time scale and time delay between sensor 1 and sensor m , respectively. Note that β_{m1} and α_{m1} can be expressed in terms of the source position or velocity at *source* time t_o as follows:

$$\beta_{m1} = [R_m(t_o) - R_1(t_o)] / c, \quad (22)$$

$$\alpha_{m1} = \frac{1 + v_m(t_o) / c}{1 + v_1(t_o) / c}. \quad (23)$$

In practice, the output of a sensor consists of the signal received from the source, plus a component due to noise. Using Eqs. (20) and (21), a simple mathematical model for the outputs of sensor 1 and sensor m is

$$y_1(t) = \tilde{s}(t) + n_1(t) \quad \text{for } |t| < \delta, \quad (24)$$

$$y_m(t) = \rho_{m1} \tilde{s}\left(\frac{t - \beta_{m1}}{\alpha_{m1}}\right) + n_m(t) \quad \text{for } |t - \beta_{m1}| < \delta, \quad (25)$$

where $n_1(t)$ and $n_m(t)$ are additive noises for the respective sensors. The source signal $s(t)$ and the noises $n_1(t)$ and $n_m(t)$ are assumed to be stationary and mutually uncorrelated. With these assumptions, $y_1(t)$ and $y_m(t)$ are stationary over the respective time intervals specified in Eqs. (24) and (25); however, they are jointly nonstationary.⁹ By suitably processing the sensor outputs, it is possible to estimate the relative time scale α_{m1} and time delay β_{m1} .

It should be emphasized that the variable t in Eqs. (24) and (25) represents the *sensor* time variable with reference to t_{1o} . Also, Eqs. (24) and (25) are valid for different time intervals which may or may not be (partially) overlapped depending on the actual source-sensor geometry. This has important implications on the processing requirement. When β_{m1} is small compared with δ , the joint estimation of α_{m1} and β_{m1} simply requires observing the outputs of sensor 1 and sensor m over the same period of time $[-T_1/2, T_1/2]$, where $\beta_{m1} \ll T_1/2 < \delta$. This is the case that has been considered in the open literature (for example, see Ref. 19 on the time-delay values used in time-delay estimation simulations for stationary sources). However, when the spacing between sensor 1 and sensor m is very large, it is most likely that $\beta_{m1} > \delta$, especially when the signal is emitted from the source near the end-fire direction. In this case, the joint estimation of α_{m1} and β_{m1} requires the observation time interval for sensor m , $[-T_m/2, T_m/2]$, to be sufficiently larger than that for sensor 1, $[-T_1/2, T_1/2]$, so that the signal given by Eq. (25) is included in the observation. The upper bound on the extra observation time required for sensor m is simply *twice* the maximum time delay between the two sensors, that is, $2 \max\{\beta_{m1}\} = 2d_{m1} / c$, where d_{m1} is the spacing between sensor 1 and sensor m .

II. WIDEBAND CORRELATION PROCESSING WITH DIFFERENTIAL DOPPLER COMPENSATION

When the effect of relative time scaling between two sensors is small (that is, α_{m1} is essentially unity), a good estimate of the time delay can be obtained by cross correlating the outputs of the two sensors.⁴ The accuracy of the time-delay estimate degrades once α_{m1} departs from unity,⁵⁻⁸ unless the mismatch between the time scales of the two signal waveforms is compensated for prior to cross correlating the sensor outputs.⁹ However, since the relative time scale is not known *a priori*, it must be estimated, along with the time delay. Here, α_{m1} and β_{m1} are estimated jointly by evaluating the wideband cross-ambiguity function defined by

$$A_{y_m y_1}(\tau, \sigma') = \sqrt{\sigma'} \int_{-\infty}^{\infty} y_m(t) y_1^*[\sigma'(t - \tau)] dt, \quad \sigma' > 0, \quad (26)$$

where * denotes complex conjugation. Under the change of variables $\sigma' \rightarrow 1/\sigma$, the wideband cross-ambiguity function is equivalent to the continuous wavelet transform¹⁰ of $y_m(t)$ with respect to the mother wavelet $y_1(t)$, that is

$$W_{y_1 y_m}(\tau, \sigma) = \frac{1}{\sqrt{\sigma}} \int_{-\infty}^{\infty} y_m(t) y_1^*\left(\frac{t - \tau}{\sigma}\right) dt, \quad \sigma > 0. \quad (27)$$

Note that continuous wavelet transform $W_{y_1 y_m}(\tau, \sigma)$ can also be interpreted as a two-dimensional cross-correlation function between $y_m(t)$ and $y_1(t)$ for various time delays τ and time scales σ . In other words, the time base of the sensor waveform $y_1(t)$ is scaled by a factor of σ , and the modified waveform is then cross correlated with the other sensor waveform $y_m(t)$.

The limits of the continuous wavelet transform's integral are from $-\infty$ to ∞ , but for the present application, $y_m(t)$ and $y_1(t)$ are essentially time limited to $[-T_m/2, T_m/2]$ and $[-T_1/2, T_1/2]$, respectively, where $T_m = T_1 + 2d_{m1}/c$. Thus, the practical implication of employing Eq. (27) is that it will be applied to two signals with different durations and needs to be computed only during the time period $[-\sigma T_1/2 + \tau, \sigma T_1/2 + \tau]$. The finite observation time T_1 is bounded above by 2δ , which ensures the signal model is valid. It is bounded below by the condition⁸ that the correlation times of signal and noise are much less than T_1 so that the time correlation $W_{y_1 y_m}(\tau, \sigma)/\sqrt{\sigma} T_1$ gives a useful approximation to the ensemble average over the noise and signal distributions when the time scale variable σ matches the true relative time scale α_{m1} .

Assuming $T_1 \leq 2\delta$ and then using Eqs. (24), (25), and (27), it can be shown that the ensemble average of $W_{y_1 y_m}(\tau, \sigma)$ attains its maximum at $\tau_{\max} = \beta_{m1}$ and $\sigma_{\max} = \alpha_{m1}$. Thus, β_{m1} and α_{m1} can be estimated by finding the values of τ and σ that maximize $W_{y_1 y_m}(\tau, \sigma)$, that is, the estimates of β_{m1} and α_{m1} are given by

$$(\hat{\beta}_{m1}, \hat{\alpha}_{m1}) = \arg\{\max_{\tau, \sigma} W_{y_1 y_m}(\tau, \sigma)\}. \quad (28)$$

The basic method of solving Eq. (28) is to perform a two-dimensional search over all possible values of τ and σ .

It is important to note that the effect of time-scaling changes and depends on the time that it is applied, and consequently, the time-scaling and time-delaying operations are not commutative.¹¹ In Eq. (27), the mother wavelet $y_1(t)$ is first compensated in time scale before correlating with $y_m(t)$. If the time-scaling and time-delaying operations are applied in reverse order, a different result will be obtained. The effect of the time-varying property of time scaling can be illustrated by using $y_m(t)$ instead of $y_1(t)$ as the mother wavelet. In this case, the continuous wavelet transform is given by

$$W_{y_m y_1}(\tau, \sigma) = \frac{1}{\sqrt{\sigma}} \int_{-\infty}^{\infty} y_1(t) y_m^*\left(\frac{t - \tau}{\sigma}\right) dt, \quad \sigma > 0. \quad (29)$$

Using Eqs. (24) and (25), it can be easily shown that the ensemble average of $W_{y_m y_1}(\tau, \sigma)$ attains its maximum at $\tau_{\max} = -\beta_{m1}/\alpha_{m1}$ and $\sigma_{\max} = 1/\alpha_{m1}$. This result is clearly different from that obtained using $y_1(t)$ as the mother wavelet.

III. EVALUATING THE CONTINUOUS WAVELET TRANSFORM

In practice, $W_{y_1 y_m}(\tau, \sigma)$ is evaluated at discrete values of τ and σ using samples of $y_1(t)$ and $y_m(t)$ as follows:

$$W_{y_1 y_m}(qT_s, \sigma_p) \approx \frac{1}{f_s \sqrt{\sigma_p}} \sum_k y_m(kT_s) y_1^*\left(\frac{k - q}{\sigma_p} T_s\right), \quad \sigma_p > 0, \quad (30)$$

where p, q, k are integers, f_s is the sampling frequency, and $T_s = 1/f_s$. Note that the discrete values of τ are uniformly spaced at time intervals of T_s and so the maximum quantization error in the time-delay estimate is $\pm T_s/2$. A major problem in computing Eq. (30) is to create the time-scaled replicas, $y_1(kT_s/\sigma_p)$, of the mother wavelet, $y_1(kT_s)$, which may be computationally intensive because the analytic expression for the received signal $y_1(t)$ is unknown. The basic method involves interpolation or decimation of $y_1(kT_s)$ by a factor of σ_p , depending on whether $\sigma_p > 1$ or $\sigma_p < 1$. In general, a low-pass filter is required in the decimation process to avoid aliasing. The amount of signal energy removed by the antialiasing filter depends on the sampling frequency f_s and decimation factor σ_p^{-1} . For the joint estimation of the relative time scale and time delay, the sampling frequency f_s is chosen to be significantly larger than the Nyquist sampling rate, which, in turn, is at least twice the signal bandwidth B_1 . This ensures most of the signal energy is preserved even with the maximum decimation factor (minimum σ_p). The resampling frequency in the case of maximum decimation is given by $f_s \min\{\sigma_p\}$ and, so, no antialiasing filter is required if $f_s \min\{\sigma_p\} \geq 2B_1$.

Four different approaches are now described for computing $W_{y_1 y_m}(qT_s, \sigma_p)$. The first two are based on interpolation prior to cross correlation, while the other two are based on the chirp z -transform and the cross-wavelet transform.

A. Multirate sampling

The time-scaled replicas $y_1(kT_s/\sigma_p)$ can be created from $y_1(kT_s)$ using the multirate sampling method.²⁰ The time scale value σ_p is written as I/D , where I and D are positive integers. To change the sampling rate of the signal from f_s to $\sigma_p f_s$, the sampling rate is first increased to $I f_s$ by adding $(I-1)$ zeros between each pair of signal samples $\{y_1(kT_s), y_1[(k+1)T_s]\}$, then the resultant signal is passed through a low-pass filter for antiimaging and antialiasing, and finally every D th sample of the filter output is selected. To improve the computational efficiency, multistaging²⁰ is often used. This requires decomposing I and D into products of prime numbers and using multiple lower-order low-pass filters.

For a given time-scale value σ_p , Eq. (30) indicates that $W_{y_1 y_m}(qT_s, \sigma_p)$ is a discrete (linear) correlation of $y_m(kT_s)$ and $y_1(kT_s/\sigma_p)$, which can be computed using fast correlation via the fast Fourier transform. Assume $y_1(kT_s)$ has a length of $L_1(-L_1/2 \leq k \leq L_1/2 - 1)$ and $y_m(kT_s)$ has a length of $L_m(-L_m/2 \leq k \leq L_m/2 - 1)$, where L_1, L_m are even and $L_1 \leq L_m$. The length of the fast Fourier transform, M , is chosen to avoid circular correlation, that is, M is a power of 2 such that $M \geq L_m + L_1 \max\{\sigma_p\} - 1$. The sequences $y_m(kT_s)$ and $y_1(kT_s/\sigma_p)$ are first padded to the same length M with the appropriate numbers of zeros. Then the fast Fourier transform of $y_m(kT_s)$ is multiplied by the complex conjugate of the fast Fourier transform of $y_1(kT_s/\sigma_p)$. Finally, the inverse fast Fourier transform of the resulting product is divided by $f_s \sqrt{\sigma_p}$ to give $W_{y_1 y_m}(qT_s, \sigma_p)$.

B. Discrete Fourier transform interpolation

A simpler way to create time-scaled replicas $y_1(kT_s/\sigma_p)$ is to interpolate $y_1(kT_s)$ using the *discrete* Fourier transform, which does not require a low-pass filter.²¹ In this method, the time scale values are limited to $\sigma_p = (L_1 + p)/L_1$, where L_1 is the length of the sequence $y_1(kT_s)$ and p is an integer multiple of 2. The time scale increment is p/L_1 and so the maximum quantization error in the relative time scale estimate is $\pm p/2L_1$. Let $Y_1(k)$ be the discrete Fourier transform of $y_1(kT_s)$, $-L_1/2 \leq k \leq L_1/2 - 1$.

1. Time base expansion (dilation)

For $\sigma_p > 1$ ($p > 0$), $p/2$ zeros are appended to each end of $Y_1(k)$, which effectively increases the sampling rate from f_s to $\sigma_p f_s$. Taking the inverse discrete Fourier transform of the zero-padded spectrum and multiplying the result by σ_p gives $y_1(kT_s/\sigma_p)$, $-(L_1 + p)/2 \leq k \leq (L_1 + p)/2 - 1$.

2. Time base contraction (compression)

For $\frac{1}{2} < \sigma_p < 1$ ($-L_1/2 < p < 0$), $L_1/2 + p$ zeros are appended to each end of $Y_1(k)$, which effectively increases the sampling rate to $2\sigma_p f_s$. Taking the inverse discrete Fourier transform of the zero-padded spectrum and multiplying the result by $2\sigma_p$ yields $y_1(kT_s/2\sigma_p)$, $-(L_1 + p) \leq k \leq (L_1 + p) - 1$. Retaining one for every two samples of $y_1(kT_s/2\sigma_p)$ gives $y_1(kT_s/\sigma_p)$, $-(L_1 + p)/2 \leq k \leq (L_1 + p)/2 - 1$, with the

sampling rate of $y_1(t)$ now effectively being $\sigma_p f_s$. To avoid aliasing, the new sampling frequency $\sigma_p f_s$ must be at least twice the signal bandwidth B_1 , that is, $\sigma_p f_s \geq 2B_1$, which places a lower bound on σ_p . For example, if $f_s = 4B_1$, the lower bound on σ_p will be $\frac{1}{2}$, which is small enough for most applications as the relative time scale is seldom less than $\frac{1}{2}$. The higher sampling rate also results in a smaller quantization error in the time delay estimate.

Having obtained the time-scaled replicas $y_1(kT_s/\sigma_p)$, $W_{y_1 y_m}(qT_s, \sigma_p)$ is computed for each value of σ_p using fast correlation via the fast Fourier transform.

C. Chirp z-transform

An alternative approach to computing $W_{y_1 y_m}(qT_s, \sigma_p)$ uses the chirp z-transform, which avoids generating the time-scaled replicas of the mother wavelet, $y_1(kT_s/\sigma_p)$. The time scale values σ_p of the chirp z-transform are chosen arbitrarily. Using Parseval's theorem, Eq. (27) can be written as

$$W_{y_1 y_m}(\tau, \sigma) = \sqrt{\sigma} \int_{-\infty}^{\infty} Y_m(f) Y_1^*(\sigma f) e^{j2\pi f \tau} df, \quad \sigma > 0, \quad (31)$$

where $Y_n(f)$ is the Fourier transform of $y_n(t)$, $n=1, m$. Evaluating Eq. (31) at discrete values of $\tau = qT_s$ and $\sigma = \sigma_p > 0$, and approximating the integral with a finite summation gives

$$W_{y_1 y_m}(qT_s, \sigma_p) \approx \frac{\sqrt{\sigma_p}}{f_s M} \sum_{k=-M/2}^{M/2-1} Y_{dm}(kf_s/M) Y_{d1}^*(\sigma_p kf_s/M) e^{j2\pi qk/M}, \quad (32)$$

where $Y_{dn}(kf_s/M)$ ($n=1, m$) are the samples of the discrete-time Fourier transform of $y_n(t)$ defined by

$$Y_{dn}(f) = \sum_k y_n(kT_s) e^{-j2\pi k f T_s}. \quad (33)$$

In Eq. (32), $M \geq L_m + L_1 \max\{\sigma_p\} - 1$ is chosen to implement the linear correlation indicated in Eq. (27) rather than a circular correlation; usually, M is chosen to be a power of 2. Given $Y_{dm}(kf_s/M)$ and $Y_{d1}^*(\sigma_p kf_s/M)$, Eq. (32) can be computed efficiently using an inverse M -point fast Fourier transform.

Here $Y_{dm}(kf_s/M)$ is obtained by padding each end of $y_m(kT_s)$ with $(M - L_m)/2$ zeros, followed by an M -point fast Fourier transform. The scaled wavelet spectrum, $Y_{d1}(\sigma_p kf_s/M)$, is computed using the chirp z-transform as suggested by Jones and Baraniuk.²² First, each end of $y_1(kT_s)$ is padded with $(M - L_1)/2$ zeros. Then using the definition of $Y_{d1}(f)$, it follows that

$$Y_{d1}(\sigma_p kf_s/M) = \sum_{n=-M/2}^{M/2-1} y_1(nT_s) e^{-j2\pi \sigma_p kn/M} \quad (34)$$

for $-M/2 \leq k \leq M/2 - 1$. By putting $m = n + M/2$ and $l = k + M/2$ ($0 \leq m, l \leq M - 1$), Eq. (34) can be written as

$$Y_{d1}(\sigma_p k f_s / M) = e^{j\pi\sigma_p k} \sum_{m=0}^{M-1} y_1 \left[\left(m - \frac{M}{2} \right) T_s \right] \times e^{-j2\pi\sigma_p l m / M} e^{j\pi\sigma_p m}. \quad (35)$$

The summation in Eq. (35) is recognized as the chirp z -transform of the sequence $y_1[(m - M/2)T_s]$ with parameters $W = e^{-j2\pi\sigma_p l / M}$ and $A = e^{-j\pi\sigma_p}$.²³ The computation time of the chirp z -transform is only a few times longer than that of an M -point fast Fourier transform. Further details regarding the implementation of the chirp z -transform are given in Ref. 23. Note that the case $\sigma_p > 1$ requires samples of $Y_{d1}(f)$ outside the interval $|f| \leq f_s/2$. Since these samples correspond to frequencies above the Nyquist frequency and assuming that the sampling rate of $y_1(t)$ satisfies the Nyquist criterion, then set $Y_{d1}(\sigma_p k f_s / M) = 0$ for $|k| > M/2\sigma_p$.

D. Cross-wavelet transform

Here $W_{y_1 y_m}(\tau, \sigma)$ can also be evaluated using the cross-wavelet transform as proposed by Young,¹¹ that is,

$$W_{y_1 y_m}(\tau, \sigma) = \frac{1}{c_g} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_g y_m(b, a) W_g^* y_1 \left(\frac{b - \tau}{\sigma}, \frac{a}{\sigma} \right) \frac{da db}{a^2}, \quad \sigma > 0, \quad (36)$$

where $W_g y_n(b, a)$ is the continuous wavelet transform of the signal $y_n(t)$ with respect to an arbitrarily chosen mother wavelet $g(t)$ ($n = 1, m$) and c_g is the admissibility constant. The continuous wavelet transform of $W_g y_n(b, a)$ is defined by

$$W_g y_n(b, a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} y_n(t) g^* \left(\frac{t - b}{a} \right) dt, \quad (37)$$

where a and b are the time-scale and time-delay variables, respectively, and both a and b can be positive or negative. Usually, the mother wavelet $g(t)$ is complex so that the resulting continuous wavelet transform, $W_g y_n(b, a)$, is also complex. If $g(t)$ has a real spectrum, then for real signals $y_1(t)$ and $y_m(t)$, it can be shown (see Appendix) that Eq. (36) can be written as

$$W_{y_1 y_m}(\tau, \sigma) = \frac{2}{c_g} \operatorname{Re} \left[\int_0^{\infty} \frac{da}{a^2} \int_{-\infty}^{\infty} W_g y_m(b, a) W_g^* y_1 \left(\frac{b - \tau}{\sigma}, \frac{a}{\sigma} \right) db \right], \quad \sigma > 0, \quad (38)$$

where $\operatorname{Re}[\]$ denotes the real part of the bracketed quantity. Equation (38) indicates that only positive scale values need to be considered in the computation.

Evaluating Eq. (38) at discrete values of $\tau = qT_s$ and $\sigma = \sigma_p > 0$, and approximating the double integral with a finite double summation gives

$$W_{y_1 y_m}(qT_s, \sigma_p) \approx \frac{2\Delta a}{c_g f_s} \operatorname{Re} \left[\sum_{i=0}^{N_a-1} \frac{1}{a_i^2} \sum_k W_g y_m(kT_s, a_i) \times W_g^* y_1 \left(\frac{k - q}{\sigma_p} T_s, \frac{a_i}{\sigma_p} \right) \right], \quad (39)$$

where $a_i = a_0 + i\Delta a$ ($0 \leq i \leq N_a - 1$) for some a_0 , $\Delta a > 0$. To obtain $W_{y_1 y_m}(qT_s, \sigma_p)$, first compute $W_g y_n(kT_s, a_i)$ for $n = 1, m$ using the discrete Fourier transform interpolation method or chirp z -transform method with $g(t)$ as the mother wavelet. It is important to take sufficient samples of $g(t)$ so that the discrete signal $g(kT_s)$ contains essentially all the energy in $g(t)$. Let L_g be the length of $g(kT_s)$. The length of the two-dimensional sequence $W_g y_n(kT_s, a_i)$ along the discrete time-delay axis, M_n , is the smallest integer that satisfies $M_n \geq L_n + L_g \max\{a_i\} - 1$. When the wavelet spectrum, $G(f)$, has a simple analytical expression (for example, the Morlet wavelet), it can be used directly in the chirp z -transform method to generate the scaled wavelet spectrum, thus saving the sampling of the mother wavelet and subsequent chirp z -transform operations. [The Morlet wavelet is defined as $g(t) = e^{-at^2} e^{j2\pi f_0 t}$. It has a real spectrum given by $G(f) = \sqrt{\pi/\alpha} e^{-\pi^2(f-f_0)^2/\alpha}$.] For a given σ_p and a_i , the inner summation in Eq. (39) is a discrete correlation of $W_g y_m(kT_s, a_i)$ and $W_g y_1(kT_s/\sigma_p, a_i/\sigma_p)$, and so it can be computed using the fast Fourier transform. The result of this correlation (obtained for each a_i) is then used to evaluate the outer summation to obtain $W_{y_1 y_m}(qT_s, \sigma_p)$ at the given σ_p for all qT_s . The creation of the two-dimensional sequence $W_g y_1(kT_s/\sigma_p, a_i/\sigma_p)$ for a given σ_p and the choice of N_a , together with the length of the fast Fourier transform, are discussed below.

Assume $W_g y_m(b, a)$ is negligibly small for $a > a_N$. Nonzero results for Eq. (39) require $a_i = a_0 + i\Delta a \leq a_N$. Therefore, N_a is chosen to be the smallest integer that satisfies $N_a \geq (a_N - a_0)/\Delta a$. To obtain $W_g y_1(kT_s, a_i/\sigma_p)$ for a given kT_s , the linear interpolation technique is applied to the real and imaginary parts of $\{W_g y_1(kT_s, a_{j-1}), W_g y_1(kT_s, a_j)\}$ where a_i/σ_p is contained in (a_{j-1}, a_j) . When using this simple interpolation technique, it is assumed Δa is sufficiently small and/or $W_g y_1(b, a)$ varies slowly with a . The length of the resulting two-dimensional sequence $W_g y_1(kT_s, a_i/\sigma_p)$ along the discrete time-delay axis is the same as that for $W_g y_n(kT_s, a_i)$, which is M_1 . For a given a_i , $W_g y_1(kT_s/\sigma_p, a_i/\sigma_p)$ can be created from $W_g y_1(kT_s, a_i/\sigma_p)$ using the discrete Fourier transform interpolation technique, provided the sampling frequency is sufficiently larger than the signal bandwidth. The time scale values are limited to $\sigma_p = (M_1 + p)/M_1$ and the resulting two-dimensional sequence $W_g y_1(kT_s/\sigma_p, a_i/\sigma_p)$ has a maximum length of $M_1 \max\{\sigma_p\}$ along the discrete time-delay axis. Using the fast Fourier transform to compute the discrete (linear) correlation of $W_g y_m(kT_s, a_i)$ and $W_g y_1(kT_s/\sigma_p, a_i/\sigma_p)$ requires the length of the fast Fourier transform, M , to be chosen so that it is a power of 2 and that $M \geq M_m + M_1 \max\{\sigma_p\} - 1$.

An alternative way of evaluating Eq. (38) is to use the chirp z -transform. Applying Parseval's theorem to the inner

integral of Eq. (38) and discretization of the resulting equation gives

$$W_{y_1 y_m}(qT_s, \sigma_p) = \frac{2\sigma_p \Delta a}{c_g f_s M} \operatorname{Re} \left[\sum_{i=0}^{N_a-1} \frac{1}{a_i^2} \sum_{k=-M/2}^{M/2-1} Y_{dm} \left(\frac{kf_s}{M}, a_i \right) \right. \\ \left. \times Y_{d1}^* \left(\frac{\sigma_p kf_s}{M}, \frac{a_i}{\sigma_p} \right) e^{j2\pi qk/M} \right], \quad (40)$$

where $Y_{dn}(kf_s/M, a_i)$ are the samples of the discrete-time Fourier transform of $W_{g y_n}(b, a_i)$ with respect to the variable b for a given a_i ($n=1, m$). If $Y_{dm}(kf_s/M, a_i)$ and $Y_{d1}(\sigma_p kf_s/M, a_i/\sigma_p)$ are known, the inner summation in Eq. (40) can be computed efficiently at the given σ_p and a_i for all q using an inverse M -point fast Fourier transform. For a given a_i , $Y_{dm}(kf_s/M, a_i)$ is simply the M -point fast Fourier transform of $W_{g y_m}(kT_s, a_i)$. For a given σ_p and a_i , $Y_{d1}(\sigma_p kf_s/M, a_i/\sigma_p)$ is obtained from $W_{g y_1}(kT_s, a_i/\sigma_p)$ using the chirp z -transform.

E. Comparison of different methods

In the multirate sampling method, the time scale values σ_p are of the form I/D . The crucial issue of this approach is the design of the multistage low-pass filters. Efficiency can be improved by optimizing the order of the filter at each stage. However, if a small time-scale increment is required, the interpolation factor I and decimation factor D will be quite large, and, consequently, many filters are needed. This is often impractical. In addition, creating each of the replicas $y_1(kT_s/\sigma_p)$ requires a new set of filters, thus increasing even further the number of filters.¹⁰ Consequently, this method is computationally intensive even for moderate time-scale increments. Another problem with this method is that the low-pass filters will introduce time delays to the time-scaled replicas $y_1(kT_s/\sigma_p)$, and these must be compensated for when estimating the differential time of arrival.

With the discrete Fourier transform interpolation method,²¹ the time-scale values are limited to $\sigma_p = (L_1 + p)/L_1 \geq 2B_1/f_s$. A smaller time-scale increment requires a longer sequence for $y_1(kT_s)$, that is, a larger L_1 . However, the largest value of L_1 is limited by the sampling frequency of the signal and the maximum observation time over which the signal model remains valid. Though this problem may be overcome by padding $y_1(kT_s)$ with zeros to increase the effective signal length, the cost of computing the discrete Fourier transform increases with L_1 . Therefore, this method becomes less efficient for very small time-scale increments ($\leq 10^{-4}$). Nevertheless, for time-scale increments on the order of 10^{-3} , or larger, this method is the most efficient.

The main advantage of the chirp z -transform method over the previous two methods is that the time scale values σ_p are arbitrary and so very small time-scale increments can be achieved without increasing the system complexity or the signal length. The efficiency of this method relies on the computational cost of the M -point chirp z -transform which is only a few times larger than that of an M -point fast Fourier transform. Consequently, for time-scale increments of the

order of 10^{-3} , or larger, this method is less efficient than the discrete Fourier interpolation method but it is more efficient than the others.

With the cross-wavelet transform method, the processing is performed in the wavelet domain. Although this method bypasses the need for multirate sampling, it is computationally very intensive as a two-dimensional interpolation is required for every time scale value σ_p .

To conclude, the discrete Fourier transform interpolation method is the most efficient method to compute the continuous wavelet transform for the present application of passive wideband cross correlation of acoustic signals with differential Doppler compensation unless the required time-scale increment is too small ($\leq 10^{-4}$), in which case the chirp z -transform method is preferred.

IV. COMPUTER SIMULATIONS

A series of computer simulations was performed to study the joint estimation of the relative time scale and differential time of arrival using Eq. (28). An idealized signal model was used in the simulations, that is, Eqs. (24) and (25) were assumed valid for all t . The outputs of two sensors (sensor 1 and sensor 2) with equal SNRs were generated as follows. A white Gaussian sequence was passed through a digital bandpass filter to produce a discrete random signal $s(kT_s)$ bandlimited between 100 and 300 Hz. The sampling frequency f_s was 6 kHz. Using the discrete Fourier transform interpolation technique, a time-scaled sequence $s(kT_s/\alpha_{21})$ was created, which was then delayed in time by $\beta_{21} = iT_s$ to produce $s[(k-i)T_s/\alpha_{21}]$. Independent white Gaussian noises of equal variances were finally added to $s(kT_s)$ and $s[(k-i)T_s/\alpha_{21}]$ to form the sensor outputs $y_1(kT_s)$ and $y_2(kT_s)$, respectively. The discrete Fourier transform interpolation method, the chirp z -transform method, and the cross-wavelet transform method were each used to compute the continuous wavelet transform for a variety of signal-to-noise ratio scenarios, for both small and large time delays, but for a fixed relative time scale. The results verify the effectiveness of using the wideband cross-ambiguity function in the joint estimation of the relative time scale and time delay. Also, the results obtained using the three different methods were the same in all cases. This is expected because these methods are merely different approaches to the evaluation of the continuous wavelet transform and there is no advantage in using one in preference to the others for noise suppression. Comparing the computation time for each method used in the simulation corroborates the statement that the discrete Fourier transform interpolation method is the most efficient method for the given time-scale increment, followed by the chirp z -transform method. Moreover, the cross-wavelet transform method was found to be so computationally intensive (its computation time is several order of magnitudes longer than those of the other two methods) that it was considered unsuitable for practical applications.

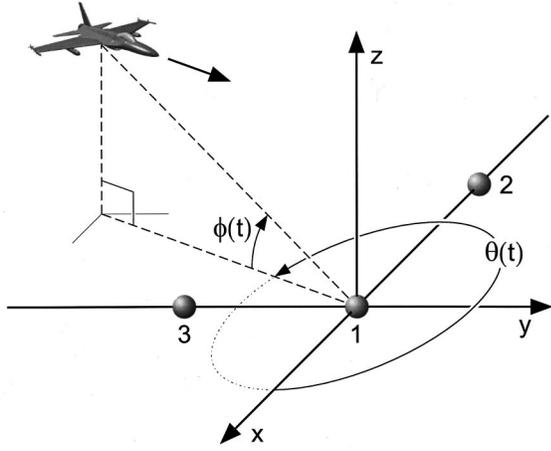


FIG. 2. The geometrical configuration of the three sensors and the source.

V. APPLICATION TO THE ANGULAR LOCATION OF JET AIRCRAFT

A. Estimating the azimuth and elevation angles of a far-field source using three orthogonal sensors

Consider a particular sensor configuration as shown in Fig. 2. Three sensors are located in the xy -plane with sensor 1 at the origin, sensor 2 at $(-d_{21}, 0, 0)$, and sensor 3 at $(0, -d_{31}, 0)$. Denote the unit vectors along the x , y , and z axes as \mathbf{x} , \mathbf{y} , and \mathbf{z} , respectively. The source position at time t can be described by one of the following position vectors:

$$\mathbf{R}_1(t) = \mathbf{R}(t) = R(t) [\cos \phi(t) \cos \theta(t) \mathbf{x} + \cos \phi(t) \sin \theta(t) \mathbf{y} + \sin \phi(t) \mathbf{z}], \quad (41)$$

$$\mathbf{R}_2(t) = \mathbf{R}_1(t) + d_{21} \mathbf{x}, \quad (42)$$

$$\mathbf{R}_3(t) = \mathbf{R}_1(t) + d_{31} \mathbf{y}, \quad (43)$$

where $R(t)$, $\phi(t)$, $\theta(t)$ are the range, elevation angle, and azimuth angle of the source at t respectively. By writing $R_m(t) = \sqrt{\mathbf{R}_m(t) \cdot \mathbf{R}_m(t)}$ for $m=2,3$ and using Eqs. (41)–(43), it is easy to show that

$$R_2(t) = [R^2(t) + d_{21}^2 + 2d_{21}R(t)\cos \phi(t)\cos \theta(t)]^{1/2}, \quad (44)$$

$$R_3(t) = [R^2(t) + d_{31}^2 + 2d_{31}R(t)\cos \phi(t)\sin \theta(t)]^{1/2}. \quad (45)$$

Using the far-field assumption ($R(t) \gg d_{21}, d_{31}$), Eqs. (44) and (45) can be approximated as

$$R_2(t) \approx R(t) + d_{21} \cos \phi(t) \cos \theta(t), \quad (46)$$

$$R_3(t) \approx R(t) + d_{31} \cos \phi(t) \sin \theta(t). \quad (47)$$

Setting $t = t_o$ in Eqs. (46) and (47) and then substituting the results into Eq. (22) gives

$$\beta_{21} \approx d_{21} \cos \phi(t_o) \cos \theta(t_o) / c, \quad (48)$$

$$\beta_{31} \approx d_{31} \cos \phi(t_o) \sin \theta(t_o) / c. \quad (49)$$

An estimate of the elevation and azimuth angles of the source at t_o can be obtained from Eqs. (48) and (49) using the estimated values for β_{21} and β_{31} , that is,

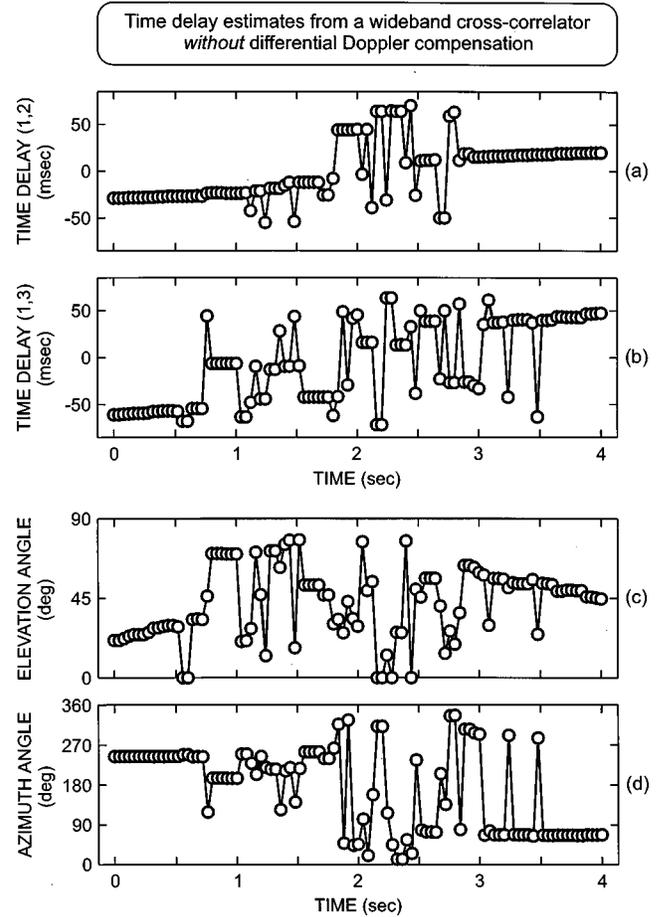


FIG. 3. Temporal variation of the time delay estimates obtained by cross correlating the outputs of (a) sensors 1 and 2 and (b) sensors 1 and 3. The variation with time of the elevation and azimuth angles of the jet aircraft during its transit are shown in (c) and (d), respectively.

$$\hat{\phi}(t_o) = \cos^{-1} [c \sqrt{(\hat{\beta}_{21}/d_{21})^2 + (\hat{\beta}_{31}/d_{31})^2}], \quad (50)$$

$$0 \leq \phi \leq \pi/2, \quad (50)$$

$$\hat{\theta}(t_o) = \tan^{-1} \left(\frac{\hat{\beta}_{31}/d_{31}}{\hat{\beta}_{21}/d_{21}} \right), \quad 0 < \theta \leq 2\pi. \quad (51)$$

B. Experimental results

The digital data from three microphones were recorded during the low-altitude transit of a jet aircraft. The sampling frequency f_s was 2 kHz, the sensor spacings d_{21} and d_{31} were 25 m, and the speed of sound in air c was 340 m/s. The data were divided into overlapped blocks and then processed so as to produce 25 observations per second, with the frequency range of interest being 50–400 Hz. The length of each data block (or observation time) was 0.34 s (T_1) for sensor 1 and 0.487 s (T_m) for other sensors. The jet was in level flight as it approached sensor 1 from a direction not far from the negative y -axis. The geometry is depicted in Fig. 2.

Figure 3(a) shows the variation with time of the time delay estimate derived by cross correlating the outputs of sensors 1 and 2, then finding the time displacement that corresponds to the maximum value of the cross-correlation function. Note that only the output of the standard cross-

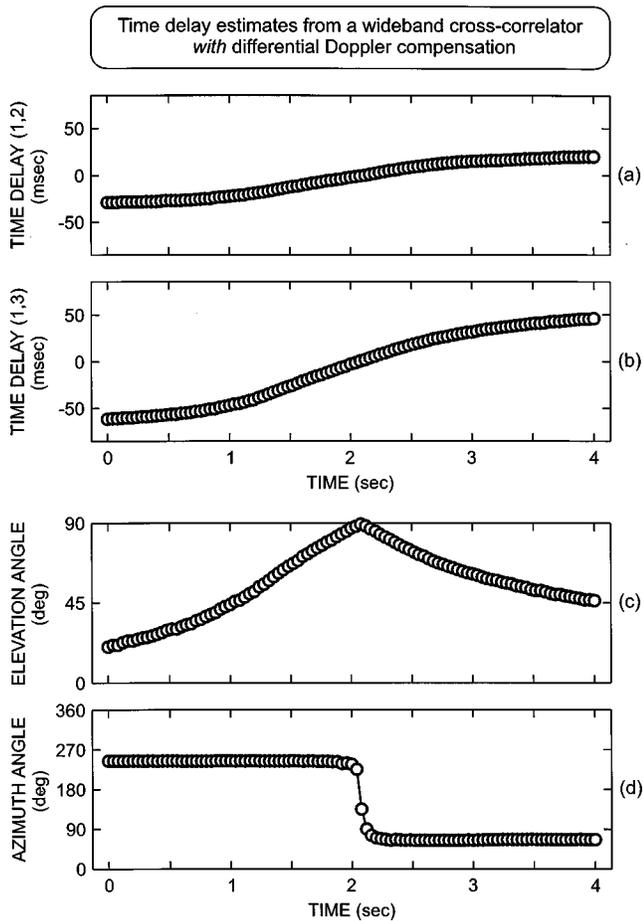


FIG. 4. Similar to Fig. 3 but the time delay estimates are derived by evaluating the passive wideband cross-correlation function with differential Doppler compensation.

correlation processor was used with no compensation for differential Doppler between the received signals and, as a result, the aircraft track is only observed at the beginning and end of the track (when the time bases of the received signals are essentially the same), with noisy time delay estimates being obtained around the closest point of approach (when the time bases of the received signals are significantly different). Similar observations can be made for the time delay estimates obtained by standard cross-correlation processing of the outputs of sensors 1 and 3 [see Fig. 3(b)].

The time delay estimates from each pair of the sensors, which are shown in Fig. 3(a) and (b), are substituted into Eqs. (50) and (51) to estimate the elevation and azimuth angles of the jet; the variation with time of these angles are shown in Fig. 3(c) and (d), respectively. (Note that the angle estimates are shown as a function of the sensor time rather than the unknown source time.) Failure of the standard cross-correlation processor to compensate for the differential Doppler prevents the observation of the aircraft's track in Fig. 3.

The same sensor data set was again processed, but using passive wideband cross-correlation processing with differential Doppler compensation, which is implemented by evaluating the continuous wavelet transform using the discrete Fourier transform interpolation method. The time-scale increment is about 0.0088. The results of the processing are shown in Fig. 4 where the improvement in performance is

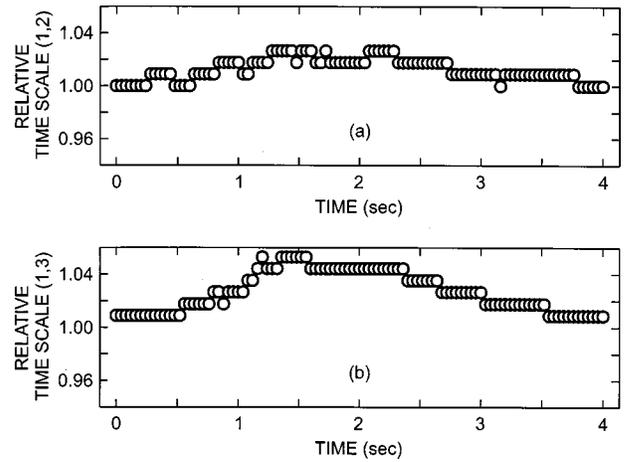


FIG. 5. Variation with time of the relative time scale estimates derived by evaluating the passive wideband cross-correlation function with differential Doppler compensation.

dramatic when compared with the results obtained from the standard cross-correlation processor (see Fig. 3). The calculated variation with (sensor) time of the aircraft's elevation and azimuth angles are shown in Fig. 4(c) and (d), respectively; the jet's track is clearly observed throughout the transit. The variation with time of the estimates of the relative

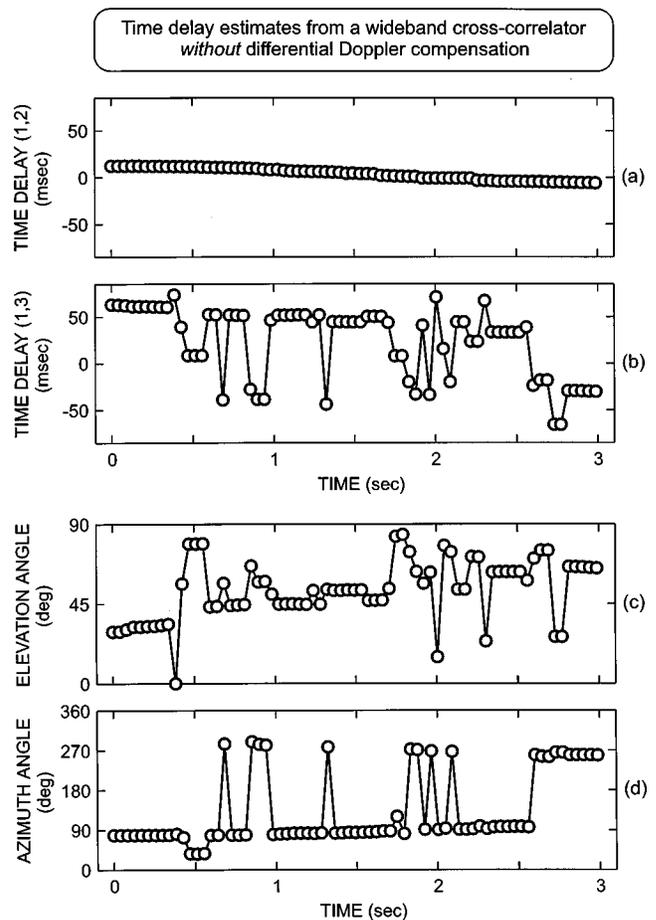


FIG. 6. Temporal variation of the time delay estimates obtained by cross correlating the outputs of (a) sensors 1 and 2 and (b) sensors 1 and 3. The variation with time of the elevation and azimuth angles of the jet aircraft during its transit are shown in (c) and (d), respectively.

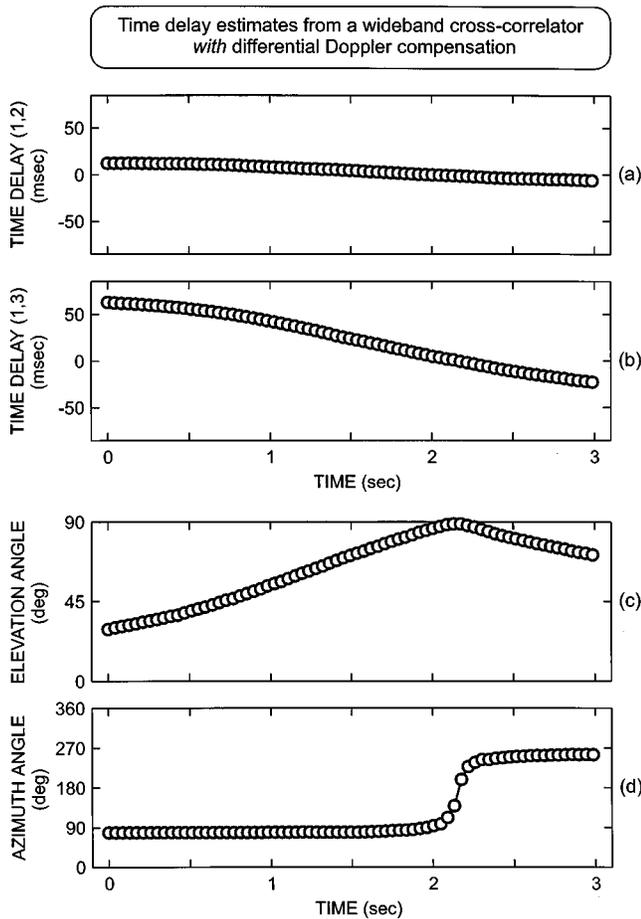


FIG. 7. Similar to Fig. 6 but the time delay estimates are derived by evaluating the passive wideband cross-correlation function with differential Doppler compensation.

time scales for the two sensor pairs are shown in Fig. 5. The data from another set of sensors collected on a different occasion during the low-altitude transit of another type of jet aircraft were also processed in a similar way using the standard cross-correlation processor (see Fig. 6) and the passive wideband cross correlator with differential Doppler compensation (see Fig. 7). Again, the improvement in performance as a result of differential Doppler compensation is

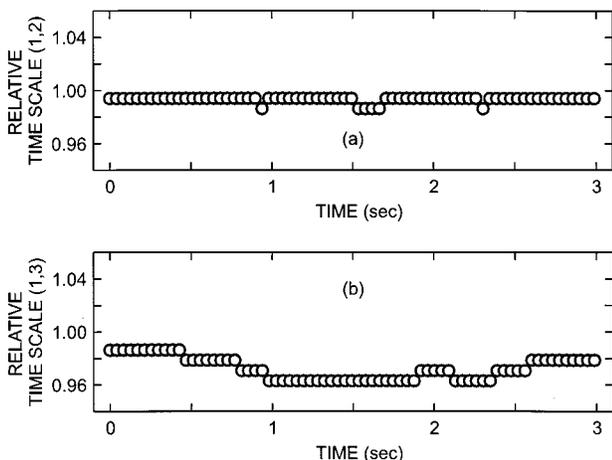


FIG. 8. Similar to Fig. 5 but for different aircraft.

dramatic. The difference in the time scales of the received signals is greater for the sensor pair (1,3) because the aircraft approached from a direction that was close to end-fire (direction of maximum relative time scale). However, for the orthogonal sensor pair (1,2) the difference in the time scales of the received signals is much smaller because the aircraft track is broadside to the axis of the sensor pair (direction of minimum relative time scale). The difference in the time delay estimates obtained using the two processors for this pair of sensors is small because the relative time-scale estimates are very close to unity (see Fig. 8).

VI. CONCLUSIONS

For fast moving sources and widely spaced sensors, the passive wideband cross-correlation method for time delay estimation requires compensation for the relative time scaling between the received signals, otherwise the time delay estimates are in error.

Passive wideband cross-correlation processing of two signals with differential Doppler compensation is equivalent to evaluating the continuous wavelet transform of one of the received signals using the other received signal as the mother wavelet. The most computationally efficient method to generate the time scale replicas of the mother wavelet is to interpolate the time series data using the discrete Fourier transform technique, prior to estimating the differential time of arrival using the cross-correlation method which, in turn, is efficiently implemented in the frequency domain using the fast Fourier transform.

The observation of jet aircraft transits requires processing the acoustic data using a passive wideband cross-correlation technique with differential Doppler compensation from which the time delay estimates can be used to calculate the variation with time of the azimuth and elevation angles of the jet during its transit.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the field trials support provided by Lieutenant Colonel Rik Modderman, Commanding Officer 1 Reconnaissance Intelligence Surveillance and Target Acquisition Regiment, and the artwork produced by Lionel Criswick for this paper.

APPENDIX: PROOF OF EQ. (38)

Using Parseval's theorem, Eq. (37) can be written as

$$W_g y_n(b, a) = \sqrt{|a|} \int_{-\infty}^{\infty} Y_n(f) G^*(af) e^{i2\pi fb} df, \quad (A1)$$

where $Y_n(f)$ and $G(f)$ are the Fourier transforms of $y_n(t)$ and $g(t)$, respectively. For a real signal $y_n(t)$, $Y_n(f) = Y_n^*(-f)$. If the wavelet spectrum $G(f)$ is also real, then by using Eq. (A1), it is clear that

$$W_g y_n(b, a) = W_g^* y_n(b, -a). \quad (A2)$$

By splitting the inner integral of Eq. (36) into two, i.e.,

$$\int_{-\infty}^{\infty} (\cdot) \frac{da}{a^2} = \int_{-\infty}^0 (\cdot) \frac{da}{a^2} + \int_0^{\infty} (\cdot) \frac{da}{a^2},$$

and then applying Eq. (A2) to the first integral for $n = 1, m$, it can be seen that

$$\begin{aligned} & \int_{-\infty}^0 W_g y_m(b, a) W_g^* y_1 \left(\frac{b-\tau}{\sigma}, \frac{a}{\sigma} \right) \frac{da}{a^2} \\ &= \int_0^{\infty} W_g y_m(b, -a) W_g^* y_1 \left(\frac{b-\tau}{\sigma}, \frac{-a}{\sigma} \right) \frac{da}{a^2} \\ &= \int_0^{\infty} W_g^* y_m(b, a) W_g y_1 \left(\frac{b-\tau}{\sigma}, \frac{a}{\sigma} \right) \frac{da}{a^2}. \end{aligned} \quad (\text{A3})$$

Combining this result with the second integral yields

$$\begin{aligned} & \int_{-\infty}^{\infty} W_g y_m(b, a) W_g^* y_1 \left(\frac{b-\tau}{\sigma}, \frac{a}{\sigma} \right) \frac{da}{a^2} \\ &= 2 \operatorname{Re} \left[\int_0^{\infty} W_g y_m(b, a) W_g^* y_1 \left(\frac{b-\tau}{\sigma}, \frac{a}{\sigma} \right) \frac{da}{a^2} \right]. \end{aligned} \quad (\text{A4})$$

Substituting Eq. (A4) into Eq. (36) gives Eq. (38).

¹F. Dommermuth and J. Schiller, "Estimating the trajectory of an accelerationless aircraft by means of a stationary acoustic sensor," *J. Acoust. Soc. Am.* **76**, 1114–1122 (1984).

²C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-24**(4), 320–327 (1976).

³J. C. Hassab and R. E. Boucher, "Optimum estimation of time delay by a generalized correlator," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-27**(4), 373–380 (1979).

⁴G. C. Carter, "Time-delay estimation for passive sonar processing," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-29**(3), 463–470 (1981).

⁵W. R. Remley, "Correlation of signals having a linear delay," *J. Acoust. Soc. Am.* **35**, 65–69 (1963).

⁶T. J. Patzewitsch, M. D. Srinath, and C. I. Black, "Nearfield performance of passive correlation processing sonars," *J. Acoust. Soc. Am.* **64**, 1412–1423 (1978).

⁷J. W. Betz, "Effects of uncompensated relative time companding on a broad-band cross correlator," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-33**(3), 505–510 (1985).

⁸W. B. Adams, J. P. Kuhn, and W. P. Whyland, "Correlator compensation requirements for passive time-delay estimation with moving source or receivers," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-28**(2), 158–168 (1980).

⁹C. H. Knapp and G. C. Carter, "Estimation of time delay in the presence of source or receiver motion," *J. Acoust. Soc. Am.* **61**, 1545–1549 (1977).

¹⁰L. G. Weiss, "Wavelets and wideband correlation processing," *IEEE Signal Process. Mag.* **11**(1), 13–32 (Jan. 1994).

¹¹R. K. Young, *Wavelet Theory and Its Applications* (Kluwer Academic, Boston, 1993).

¹²J. W. Betz, "Cross correlation-based time delay estimation," Ph.D. dissertation, Northeastern University, Boston, MA, 1984.

¹³K. Scarbrough, "Analysis of time delay estimator performance," Ph.D. dissertation, Kansas State University, Manhattan, KS, 1984. Published as a Tech. Rep. of the Naval Underwater Systems Center, New London, CT, No. TR 7203, June 1984.

¹⁴K. C. Ho and Y. T. Chan, "Optimum discrete wavelet scaling and its application to delay and doppler estimation," *IEEE Trans. Signal Process.* **SP-46**(9), 2285–2290 (1998).

¹⁵E. Weinstein and D. Kletter, "Delay and doppler estimation by time-space partition of the array data," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-31**(6), 1523–1535 (1983).

¹⁶G. W. Johnson, D. E. Ohlms, and M. L. Hampton, "Broadband correlation processing," *Proc. Int. Conf. Acoust., Speech Signal Process., ICASSP 83*, Boston, pp. 583–586 (1983).

¹⁷J. W. Betz, "Comparison of the deskewed short-time correlator and the maximum likelihood correlator," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-32**(2), 285–294 (1984).

¹⁸S. Stergiopoulos and A. T. Ashely, "An experimental evaluation of split-beam processing as a broadband bearing estimator for line array sonar systems," *J. Acoust. Soc. Am.* **102**, 3556–3563 (1997).

¹⁹R. L. Kirlin and J. N. Bradley, "Delay estimation simulations and a normalized comparison of published results," *IEEE Trans. Acoust., Speech, Signal Process.* **ASSP-30**(3), 508–511 (1982).

²⁰R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1983).

²¹B. G. Ferguson, "Time delay estimation techniques applied to the acoustic detection of jet aircraft transits," *J. Acoust. Soc. Am.* **106**, 255–264 (1999).

²²D. L. Jones and R. G. Baraniuk, "Efficient approximation of continuous wavelet transforms," *Electron. Lett.* **27**(9), 748–750 (1991).

²³L. R. Rabiner, R. W. Schafer, and C. M. Rader, "The chirp z -transform algorithm," *IEEE Trans. AU-17(2), 86–92 (1969).*

Optimal array element localization

Stan E. Dosso^{a)}

School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia V8W 3P6, Canada

Barbara J. Sotirin

Cold Regions Research and Engineering Laboratory, US Army Corps of Engineers, 72 Lyme Road, Hanover, New Hampshire 03775-1290

(Received 28 October 1998; revised 3 August 1999; accepted 30 August 1999)

Advanced array processing methods require accurate knowledge of the location of individual elements in a sensor array. Array element localization (AEL) methods are typically based on inverting acoustic travel-time measurements from a series of controlled sources at well-known positions to the sensors to be localized. An important issue in AEL is designing the configuration of source positions: a well-designed configuration can produce substantially better sensor localization than a poor configuration. In this paper, the effects of the source configuration and of errors in the data, source positions, and ocean sound speed are quantified using a sensor-position error measure based on the *a posteriori* uncertainty of a general formulation of the AEL inverse problem. Optimal AEL source configurations are determined by minimizing this error measure with respect to the source positions using an efficient hybrid optimization algorithm. This approach is highly flexible, and can be applied to any sensor configuration and combination of errors; it is also straightforward to apply constraints to the source positions, or to include the effects of data errors that vary with range. The ability to determine optimal source configurations as a function of the number of sources and of the errors in the data, source positions, and sound speed allows the effects of each of these factors to be examined quantitatively in a consistent manner. A modeling study considering these factors can guide in the design of AEL systems to meet specific objectives for sensor localization. © 1999 Acoustical Society of America. [S0001-4966(99)03912-0]

PACS numbers: 43.60.Pt, 43.30.Xm [DLB]

INTRODUCTION

Measuring ocean acoustic fields at an array of sensors allows the application of advanced signal processing methods such as beamforming¹ and matched-field inversion.²⁻⁵ However, these methods require accurate knowledge of the positions of individual sensors in the array. For example, a general rule to achieve a loss of less than 1 dB in array-processing gain requires the sensor positions be known to within $\lambda/10$, where λ is the wavelength at the frequency of interest.^{6,7} Deploying large hydrophone arrays at sea is not an exact procedure, and sufficiently accurate sensor positions are often not known after the fact. In many cases, an acoustic survey is required to localize the sensors more precisely after deployment, a procedure referred to as array element localization (AEL). These AEL surveys are generally based on transmitting acoustic signals from controlled sources at well-known positions to the sensors to be localized. Given knowledge of the ocean sound speed, measurements of the arrival times of these signals can then be inverted for estimates of the sensor locations.

Typical approaches to AEL differ for static and dynamic arrays. In the case of a dynamic array, such as a vertical line array (VLA) which moves with currents in the water column, AEL is required on an on-going basis. The AEL systems for VLA's are typically based on continually transmitting signals from a set of acoustic transponders positioned about the

array on the seafloor [e.g., Fig. 1(a)], allowing the sensor positions to be tracked with time.⁸⁻¹² In most cases, the transponders are triggered by a control unit that is synchronized with the VLA recording system, so that the instant of signal transmission (referred to as source instant) is known and the measured arrival times can be interpreted as absolute travel times. In the case of a static array, such as a horizontal line array (HLA) fixed to the seafloor, AEL is required only once. The AEL measurements for a static array can be obtained by deploying acoustic sources in the water column about the array site and recording the arrivals at the sensors¹³⁻¹⁵ [Fig. 1(b)]. The sources can be synchronized with the array recording system (e.g., by using a transponder system or by independently measuring the source instant) providing absolute travel-time measurements. Alternatively, the sources can be independent of the recording system (nonsynchronized); in this case the measured arrival-time data provide only relative travel times. Relative travel-time data provide less information than absolute measurements; however, the use of nonsynchronized sources, such as light bulbs imploded at depth in the water column, allows for simple, inexpensive AEL surveys.^{15,16}

The precision to which the array sensors can be localized in an AEL survey depends on a number of factors. Perhaps the most obvious is the uncertainty of the measured arrival-time data. However, uncertainties in the source positions, which are often considered "known" parameters, can also introduce substantial localization errors and often represent the limiting factor in AEL.^{11-13,15} In addition, errors in

^{a)}Electronic mail: sdosso@uvic.ca

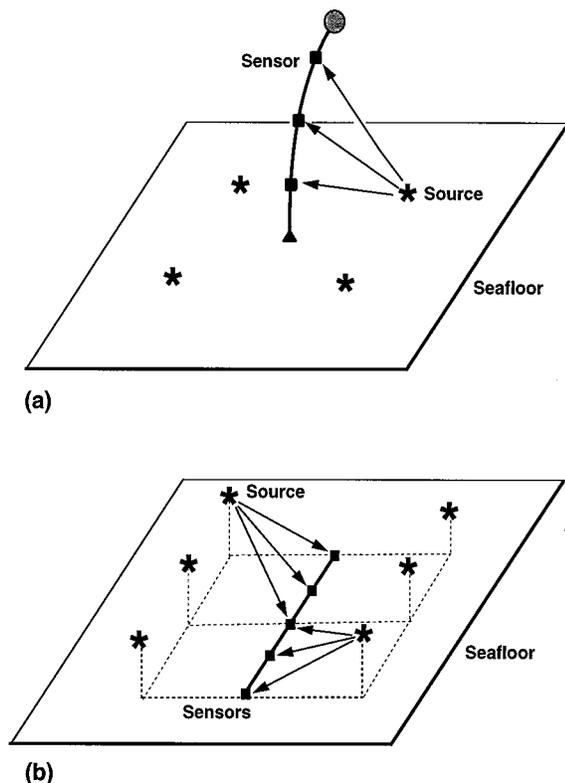


FIG. 1. Schematic diagrams of typical AEL surveys for (a) a vertical line array and (b) a bottomed horizontal line array.

the ocean sound speed can be significant. Although measured sound-speed profiles are usually accurate in a relative sense, bias errors of as much as 2 m/s are not uncommon.¹⁷ The translation of errors in the data, source positions, and sound speed into errors in the sensor positions is determined by the AEL inverse problem; simple estimates based on forward calculations are not correct (e.g., a 1-ms travel-time error in an ocean of sound-speed 1500 m/s does *not* imply a 1.5-m positioning error). A poorly conditioned inverse problem will magnify the effects of errors to a much greater extent than a well-conditioned problem.¹⁸ The conditioning of AEL inversion is determined by the source-sensor geometry. Hence, for a given sensor array, the configuration of source positions determines the conditioning of the inversion and strongly affects the accuracy of sensor localization. We emphasize that only by jointly considering the effects of both errors and the inversion can AEL uncertainties be studied in a meaningful way.

In this paper, the effects of the source configuration and of errors in the data, source positions, and sound speed are quantified using an error measure based on the *a posteriori* uncertainty of a general formulation of the AEL inverse problem. Several error measures are defined including the mean three-dimensional (3-D) error in the sensor positions and the maximum 3-D error. The optimal AEL source configuration for a particular scenario can be determined by minimizing the error measure with respect to the source positions. This is a challenging nonlinear optimization problem, and is solved here using an efficient hybrid optimization algorithm.^{18–20} The concept of designing optimal experi-

ments by minimizing a measure of the *a posteriori* error, resolution, or conditioning of the inverse problem has been previously applied to ocean acoustic tomography²¹ and to geophysical inverse problems.^{22–24} The algorithm developed here for AEL is very flexible, and can be used to determine the optimal source configuration for any sensor configuration, for relative or absolute travel-time data, and for any combination of errors in the data, source positions, and sound speed. It is also straightforward to apply physical constraints to the source configuration, or to include the effects of data errors that vary with range. In addition, the ability to determine optimal source configurations as a function of the type of data, the number of sources, and the errors in the data, source positions, and sound speed allows the effects of each of these factors to be examined quantitatively in a consistent manner. A modeling study considering these factors can guide in the design of AEL systems to meet specific objectives for sensor localization.

The following section of this paper develops the inversion and optimization methods that form the basis for optimal AEL survey design. In Sec. II we illustrate the application of these methods for both horizontal and vertical arrays, including determination of optimal source configurations and assessment of the effects due to various error sources. Finally, in Sec. III we summarize this paper.

I. THEORY

To develop an algorithm for optimal AEL survey design requires consideration of the AEL inverse problem. The AEL represents a nonlinear inverse problem; however, it is well suited to linearization and iterated linear inversion. Methods of linear inverse theory allow underdetermined inversions to be formulated by treating all parameters as unknowns with differing degrees of *a priori* knowledge, and provide estimates of the *a posteriori* uncertainties of the solution parameters.^{25–27} To include the effects of uncertainty in the source positions and sound speed, a simultaneous inversion is formulated here for sensor positions, source positions, sound-speed bias, and source instants.

The procedure for designing optimal AEL source configurations developed in this paper consists of (i) defining an AEL error measure based on the sensor-position uncertainties derived from the linearized solution, and (ii) determining the source positions that minimize this error measure by using an efficient hybrid optimization algorithm. The components of this scheme are described as follows. In Sec. IA we develop a general approach to linearized inversion for AEL. This development extends the inversion in Ref. 12 to include the sound-speed bias, and provides an analytic expression for the *a posteriori* uncertainty of the solution parameters. In Sec. IB we define several AEL error measures based on the *a posteriori* uncertainties in sensor positions, and in Sec. IC we describe the hybrid optimization algorithm used to minimize this error.

A. AEL inversion

The set of acoustic arrival times \mathbf{t} measured in an AEL survey can be written in general vector form as

$$\mathbf{t} = \mathbf{T}(\mathbf{m}) + \mathbf{n}. \quad (1)$$

In (1), the forward mapping \mathbf{T} represents the arrival times of the acoustic signals along direct ray paths between sources and receivers (given explicitly in the Appendix). The model \mathbf{m} of unknown parameters is taken to consist of 3-D position variables x, y, z for each sensor, position variables x', y', z' for each source, the sound-speed bias c_b , and, for nonsynchronized surveys, the source instant t^0 for each source. Finally, \mathbf{n} represents the data errors (noise).

As mentioned above, the AEL inverse problem of determining an estimate $\hat{\mathbf{m}}$ of \mathbf{m} is functionally nonlinear; however, a local linearization can be obtained by expanding $\mathbf{T}(\hat{\mathbf{m}}) = \mathbf{T}(\mathbf{m}_0 + \delta\mathbf{m})$ in a Taylor series to first order about an arbitrary starting model \mathbf{m}_0 to yield

$$\mathbf{t} = \mathbf{T}(\mathbf{m}_0) + \mathbf{J}(\mathbf{m}_0)[\hat{\mathbf{m}} - \mathbf{m}_0], \quad (2)$$

where \mathbf{J} represents the Jacobian matrix of partial derivatives $J_{ij} = \partial T_i / \partial m_j$ (given explicitly in the Appendix). Equation (2) can be written

$$\mathbf{J}\hat{\mathbf{m}} = \mathbf{t} - \mathbf{T}(\mathbf{m}_0) + \mathbf{J}\mathbf{m}_0 \equiv \mathbf{d}, \quad (3)$$

where the explicit dependence on \mathbf{m}_0 has been suppressed. Note that \mathbf{d} consists entirely of known or measured quantities, and may be considered modified data for the problem. Equation (3) represents a linear inverse problem for $\hat{\mathbf{m}}$, the solution of which is considered below. Since nonlinear terms have been neglected in (3), the linearized inversion may need to be repeated iteratively until the solution converges (i.e., update $\mathbf{m}_0 \leftarrow \hat{\mathbf{m}}$ and repeat the inversion until $\hat{\mathbf{m}} = \mathbf{m}_0$).¹²

To consider the linear inverse problem (3), assume that the noise on the data can be represented by a zero-mean, Gaussian-distributed random process with covariance matrix $\mathbf{C}_n = \langle \mathbf{n}\mathbf{n}^T \rangle$, where $\langle \cdot \rangle$ indicates the expected value (under the typical assumption of uncorrelated noise, \mathbf{C}_n is a diagonal matrix with the i th entry representing the variance of the i th datum). The method of regularization allows linear inverse problems to be formulated which are underdetermined in an information sense by including *a priori* estimates and uncertainties of the unknown parameters (AEL inversion with both source and sensor positions as unknowns always represents an underdetermined problem). Let $\hat{\mathbf{m}}$ be the *a priori* estimate of the model parameters and $\mathbf{C}_{\hat{\mathbf{m}}} = \langle (\hat{\mathbf{m}} - \mathbf{m})(\hat{\mathbf{m}} - \mathbf{m})^T \rangle$ be the covariance matrix of this estimate (for uncorrelated estimate errors, this matrix is also diagonal with the estimate variances on the main diagonal). The regularized solution is formulated by defining an objective function ψ that combines terms representing the l_2 norms of the data misfit (weighted by the inverse of the data covariance matrix) and the deviation from the prior estimate (weighted by the inverse of the estimate covariance matrix)^{12,26}

$$\psi = (\mathbf{J}\hat{\mathbf{m}} - \mathbf{d})^T \mathbf{C}_n^{-1} (\mathbf{J}\hat{\mathbf{m}} - \mathbf{d}) + (\hat{\mathbf{m}} - \hat{\mathbf{m}})^T \mathbf{C}_{\hat{\mathbf{m}}}^{-1} (\hat{\mathbf{m}} - \hat{\mathbf{m}}). \quad (4)$$

The data and *a priori* information are optimally applied by determining the model $\hat{\mathbf{m}}$ that minimizes ψ (i.e., by setting $\partial\psi/\partial\hat{\mathbf{m}} = 0$) to yield

$$\hat{\mathbf{m}} = \hat{\mathbf{m}} + [\mathbf{J}^T \mathbf{C}_n^{-1} \mathbf{J} + \mathbf{C}_{\hat{\mathbf{m}}}^{-1}]^{-1} \mathbf{J}^T \mathbf{C}_n^{-1} (\mathbf{d} - \mathbf{J}\hat{\mathbf{m}}). \quad (5)$$

Optimal AEL survey design is based not on the solution (5) to the AEL inverse problem, but rather on minimizing an estimate of the uncertainty in the solution. For a linear inverse problem and Gaussian noise, the marginal *a posteriori* probability distributions of the parameters are also Gaussian with variances given by the diagonal entries of the solution covariance matrix²⁷

$$\mathbf{C}_{\hat{\mathbf{m}}} = \langle (\hat{\mathbf{m}} - \mathbf{m})(\hat{\mathbf{m}} - \mathbf{m})^T \rangle. \quad (6)$$

Noting that the true model may be represented

$$\mathbf{m} = \mathbf{m} + [\mathbf{J}^T \mathbf{C}_n^{-1} \mathbf{J} + \mathbf{C}_{\hat{\mathbf{m}}}^{-1}]^{-1} \mathbf{J}^T \mathbf{C}_n^{-1} (\mathbf{d} - \mathbf{n} - \mathbf{J}\mathbf{m}) \quad (7)$$

and substituting (5) and (7) into (6) leads to

$$\mathbf{C}_{\hat{\mathbf{m}}} = [\mathbf{J}^T \mathbf{C}_n^{-1} \mathbf{J} + \mathbf{C}_{\hat{\mathbf{m}}}^{-1}]^{-1}. \quad (8)$$

Hence, the standard deviation σ_i of parameter \hat{m}_i is given by

$$\sigma_i = \sqrt{\langle \mathbf{C}_{\hat{\mathbf{m}}} \rangle_{ii}}. \quad (9)$$

Note that the parameter error estimates σ_i defined by (8) and (9) depend on the data uncertainties through \mathbf{C}_n , on the uncertainties in the *a priori* source-position and sound-speed estimates through $\mathbf{C}_{\hat{\mathbf{m}}}$, and on the source-sensor geometry through \mathbf{J} . The above uncertainty analysis is exact only for linear inverse problems; however, for problems that are only weakly nonlinear, it can provide accurate estimates of the parameter uncertainties. The validity of the linear approximation to the AEL inverse problem is demonstrated by numerical simulation in Sec. II.

The following section defines an AEL error measure based on the *a posteriori* uncertainties of the recovered sensor positions from (9). Optimal AEL source configurations can then be determined by minimizing this error measure with respect to the source positions, which is considered in Sec. IC. Note that this procedure is based on an *expected* inverse problem (i.e., a particular configuration for the sensor array is assumed), but there is no data to invert. Obviously, in practical applications the sensor positions are not accurately known (or there would be no need for AEL), and hence the expected sensor positions on which the minimization is based will be in error. The effects of these errors are considered in Sec. II and shown to be small.

B. AEL error measures

A number of different AEL error measures can be defined using the standard deviations of the individual sensor-position parameters given by (9). Let $\sigma_x, \sigma_y, \sigma_z$ represent the standard deviations of the x, y, z Cartesian coordinates of the sensor positions and $\sigma_r = [\sigma_x^2 + \sigma_y^2 + \sigma_z^2]^{1/2}$. The error measure that is considered primarily in this paper is

$$E = \frac{1}{N_S} \sum_r \sigma_r, \quad (10)$$

where N_S is the total number of sensors to be localized. This measure represents the mean 3-D error of the sensor positions. The source configuration that minimizes this error measure will provide the sensor-position estimates that are the most accurate on average; however, the sensor-position errors are not controlled individually. An alternative is to

minimize the maximum sensor-position error defined by

$$E_M = \text{Max}_r \{ \sigma_r \}. \quad (11)$$

The source configuration that minimizes E_M will provide the set of sensor-position estimates that has the smallest maximum error. This tends to reduce discrepancies in the sensor-position errors, but results in a larger mean (and total) error. Other error measures can also be devised and may be appropriate for specific AEL objectives. For example, if accurate sensor depths are deemed more important than accurate horizontal positions, the σ_z term in the definition of σ_r could be weighted by a factor greater than one and either E or E_M minimized; however, such cases are not considered further in this paper.

C. AEL optimization

Optimal AEL survey design consists of determining the set of source-position parameters that minimizes the sensor-position error E . This is a strongly nonlinear minimization problem which typically has a degenerate global minimum (due to symmetries) and a large number of local minima, and hence is not amenable to linearized optimization methods. In addition, since the size of the parameter search space increases geometrically with the number of sources, grid-search based solutions quickly become impractical. Global optimization methods, such as simulated annealing²³ (SA) and genetic algorithms,²⁴ have been applied to minimization problems associated with geophysical experiment design, but can be relatively inefficient. Recently, hybrid optimization methods have been developed¹⁸ and applied to geophysical¹⁹ and ocean-acoustic²⁰ inverse problems. Hybrid methods combine local and global approaches to produce a more efficient optimization. Here, a hybrid optimization algorithm that combines the local downhill simplex (DHS) method with SA is applied to optimal AEL survey design; the algorithm is similar to that described in detail in Ref. 20. For completeness, the following subsections briefly describe SA, DHS, and the hybrid simplex simulated annealing (SSA) algorithms.

1. Simulated annealing (SA)

Simulated annealing (SA) is a global optimization method that can be applied to minimize a function E with respect to a set of model parameters defined on a given search interval.¹⁸ The algorithm consists of a series of iterations involving random perturbations of the parameters. After each iteration a control parameter, the temperature T , is decreased slightly. Perturbations that decrease E are always accepted; perturbations that increase E are accepted conditionally, with a probability P that decreases with T according to the Boltzmann distribution

$$P(\Delta E) = \exp(-\Delta E/T). \quad (12)$$

Accepting some perturbations that increase E allows the algorithm to escape from local minima in search of a better solution. As T decreases, however, accepting increases in E becomes increasingly improbable, and the algorithm eventually converges. The rate of reducing T and the number and type of perturbations define the annealing schedule. Fast SA

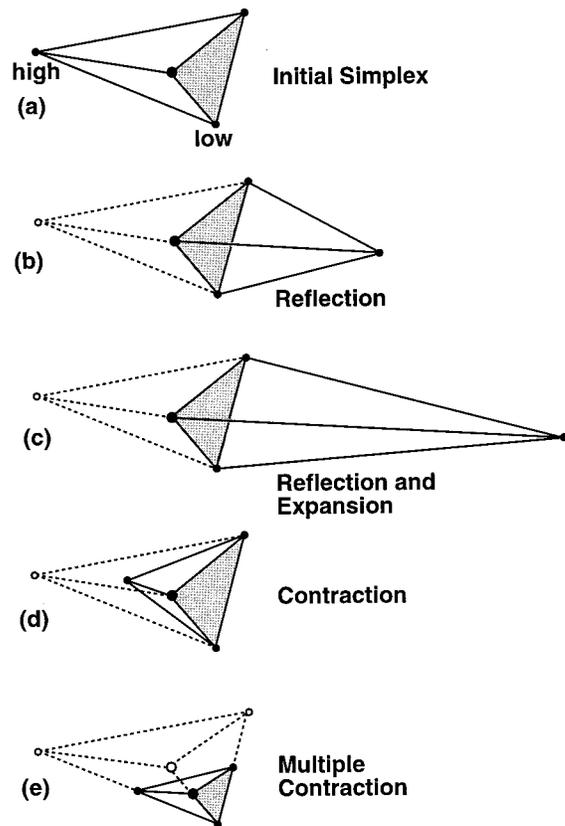


FIG. 2. Types of steps attempted by the DHS algorithm in three dimensions [described in text, after Press *et al.* (Ref. 18)].

(FSA) is based on using a Cauchy distribution to generate the parameter perturbations and reducing the width of the distribution with the temperature.^{28,29} The narrow peak and flat tails of the Cauchy distribution provide concentrated local sampling together with occasional large perturbations, allowing a faster rate of temperature reduction than standard SA.

2. Downhill simplex (DHS)

Global optimization methods widely search the parameter space and avoid becoming trapped in unfavorable local minima. However, since individual steps are computed randomly, these methods can be quite inefficient at moving downhill. In contrast, local (gradient-based) methods move efficiently downhill, but typically become trapped in a local minimum close to the starting model. The DHS method is a local inversion technique based on a geometric scheme for moving downhill in E that does not require the computation of partial derivatives or the solution of systems of equations.¹⁸ The DHS method navigates the search space using a simplex of $M + 1$ models in an M -dimensional parameter space [e.g., Fig. 2(a), for $M = 3$]. The algorithm initially attempts to improve the model with the highest value of E by reflecting it through the opposite face of the simplex [Fig. 2(b)]. If the new model has the lowest E in the simplex, an extension by a factor of 2 in the same direction is attempted [Fig. 2(c)]. If the model obtained by the reflection still has the highest E , the reflection is rejected and a contraction by a factor of 2 along this direction is attempted

[Fig. 2(d)]. If none of these steps decrease E , then a multiple contraction about the lowest- E model is performed [Fig. 2(e)]. This process is repeated until the value of E for each model of the simplex converges to a common value (i.e., the simplex shrinks to a single point at the local minimum).

3. Simplex simulated annealing (SSA)

The goal of hybrid inversion is to combine local and global methods to exploit the advantages of each (i.e., to move efficiently downhill, yet avoid becoming trapped in local minima). Here, a hybrid SSA inversion is developed that incorporates the local DHS method into a global SA search. Unlike standard SA, the SSA inversion operates on a simplex of models rather than on a single model, and instead of employing purely random model perturbations, DHS steps with a random component are applied to perturb the models. To introduce the random component, the DHS steps are not computed directly from the current simplex of models, but rather from a secondary simplex which is formed by applying random perturbations to all the model parameters and E values associated with the current simplex. (The secondary simplex is used only to compute the perturbed models; perturbation acceptance and model updating is based on the current simplex.) The perturbations to the current simplex used to produce the secondary simplex are computed using a Cauchy distribution and reducing the distribution width with temperature as follows.²⁹ Each source-position parameter $u \in \{x', y'\}$ is perturbed according to

$$u \leftarrow u + \xi \Delta, \quad (13)$$

where Δ represents the difference between the upper and lower limits assumed for u and the quantity ξ is a temperature-dependent, Cauchy-distributed random variable computed as

$$\xi = [T_j/T_0]^{1/2} \tan[\pi(\eta - \frac{1}{2})]. \quad (14)$$

In (14), η is a uniform random variable on $[0, 1]$, and T_j is the temperature at the j th step. The perturbation to the value of E associated with each model in the simplex is computed according to

$$E \leftarrow E + \xi \bar{E}, \quad (15)$$

where ξ is computed according to (14) and \bar{E} is the mean value of E for the current simplex. Each new model proposed by a DHS step is evaluated for acceptance based on the probabilistic criterion of SA applied to the actual (not perturbed) energies, before and after perturbation, for that model. This provides a mechanism for accepting uphill steps and escaping from local minima. If any DHS step results in parameter values outside their given search interval, the parameters are set to the interval bound prior to evaluation. After the set of perturbations is complete, the temperature is reduced according to

$$T_j = \beta^j T_0, \quad (16)$$

where β is a constant less than one. An appropriate starting temperature T_0 can be determined by requiring that at least 90% of all perturbations are accepted initially. Appropriate values for β and the number of perturbations per temperature

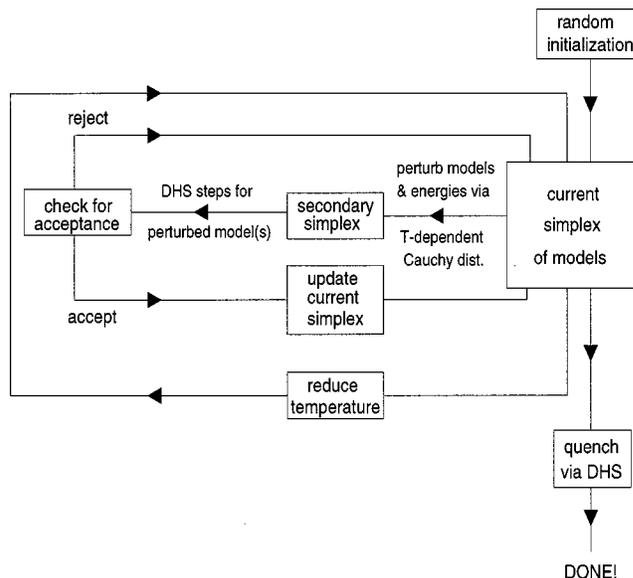


FIG. 3. Block diagram illustrating the SSA algorithm (after Ref. 20).

step are usually straightforward to determine with some experimentation.

At high temperatures where the random component of the perturbations dominates, the SSA method resembles a FSA global search. At low temperatures, where the random component is small, the method resembles the local DHS method. At intermediate temperatures, the method makes a smooth transition between these two endpoints. We have found SSA to be a highly efficient method to determine optimal AEL source configurations, as illustrated in the following section. The efficiency can be improved further by “quenching” the optimization when it approaches convergence (i.e., when E effectively stops decreasing) by switching to a pure DHS algorithm to avoid the slow final convergence typical of SA. A block diagram illustrating the basic SSA algorithm is given in Fig. 3.

II. EXAMPLES

This section presents a number of examples of optimal AEL survey design for both horizontal and vertical sensor arrays. The examples presented here are not intended to exhaustively consider all aspects of AEL design, but rather to illustrate how the approach developed in Sec. II can be applied to determine optimal source configurations and to quantify the effects of the various sources of errors in AEL. In addition, the applicability of linearized uncertainty analysis to AEL and the effects of errors in the expected inverse problem are examined.

A. Optimal AEL for horizontal arrays

1. Nonsynchronized surveys

The first example consists of determining the optimal configuration for water-column acoustic sources used to localize the sensors of a bottom-mounted HLA [e.g., Fig. 1(b)]. The scenario is designed to emulate AEL using lightbulb implosions as acoustic sources, with source positioning provided by differential GPS measurements. In this example, the ocean is 500 m deep with the Arctic sound-speed profile

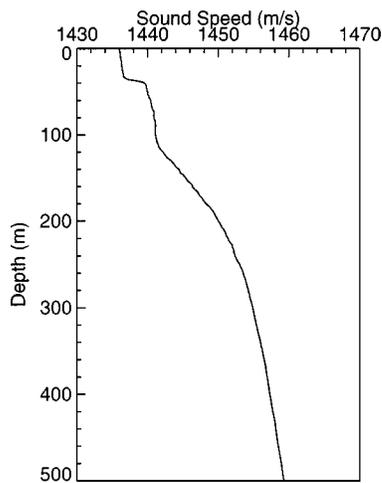


FIG. 4. Arctic sound-speed profile for AEL examples.

shown in Fig. 4. The potential bias of the sound-speed measurements is assumed to introduce a (depth-independent) uncertainty of 2 m/s. The HLA is aligned along the x axis with the eight sensors to be localized equally spaced from $x = -1000$ to 1000 m [see Fig. 5(a)]. The sources and recordings are not synchronized, and hence the source instants are treated as unknowns. The errors on the data are uncorrelated with a standard deviation of 0.5 ms. The acoustic sources are located at a nominal depth of 50 m, representing an approxi-

mate optimal depth for light-bulb sources (source level and bandwidth increase with depth; however, standard bulbs typically cannot withstand depths much greater than 50 m). The source positions are known to within an uncertainty in x' , y' of 5 m and in z' to 2 m; source-position errors are uncorrelated. The uncertainty of the *a priori* estimates of the sensor positions is taken to be infinite, so that only the acoustic data influence these parameters.

The SSA optimization algorithm developed in Sec. I was applied to determine the set of source positions (x' , y') that minimize the mean sensor-position error E given by (10). (Note that source depth z' could also be included in the optimization; however, this is not done here since the depths useful for light-bulb sources fall in a fairly narrow range about the nominal depth of 50 m.) The search interval for the horizontal position parameters was defined to be $|x'| \leq 1500$ m, $|y'| \leq 1000$ m (these constraints ensure that direct acoustic ray paths exist between each source and sensor in the upward-refracting ocean environment of Fig. 4). Figure 5(a)–(c) shows the optimal configurations for four, five, and six sources, and Fig. 5(d)–(f) shows the corresponding *a posteriori* standard deviations in the sensor positions. The form of the optimal configurations obtained appear to be unique: repeating the SSA optimizations with different sequences of random model perturbations produced configurations that were essentially identical to those shown in Fig. 5, up to reflections about the x and/or y axes. The source configurations shown in Fig. 5(a) and (c) exhibit natural symmetries that are an inherent property of the optimal configuration and were in no way built into the optimization algorithm. The configuration in Fig. 5(b) lacks symmetry; however, this may be due to the effect of the constraints, as one source is located at $(x', y') = (1500, 1000)$ m. Figure 5(d)–(f) shows that the sensor-position uncertainties decrease as the number of sources is increased. In particular, a substantial decrease in the mean sensor-position error from $E = 19.8$ m to $E = 7.8$ m is obtained by increasing from four sources (the minimum number for a nonsynchronized survey) to five sources. This indicates that for this scenario, AEL with four sources is not a well-determined inverse problem and would not be recommended.

The annealing schedule of the SSA optimizations for Fig. 5 (and all others in this paper) was based on the requirement that ten model perturbations be accepted at each temperature step, and the temperature be reduced between steps by a factor of $\beta = 0.975$. Each optimization required approximately 5–10 min of computation time on a 200-MHz Pentium PC running IDL (Interactive Data Language). An example of the optimization procedure is given in Fig. 6, which shows the sensor-position error E and the source coordinates x' , y' as a function of temperature step for the four-source case (all models in the simplex are shown). Initially, the source coordinates fluctuate over their entire allowed range and sensor-position errors as large as $E = 10^4$ m are obtained. The error E decreases steadily, although not monotonically, with temperature until approximately step 250. At this point E has essentially stopped decreasing, which indicates the various models in the simplex are simply fluctuating between

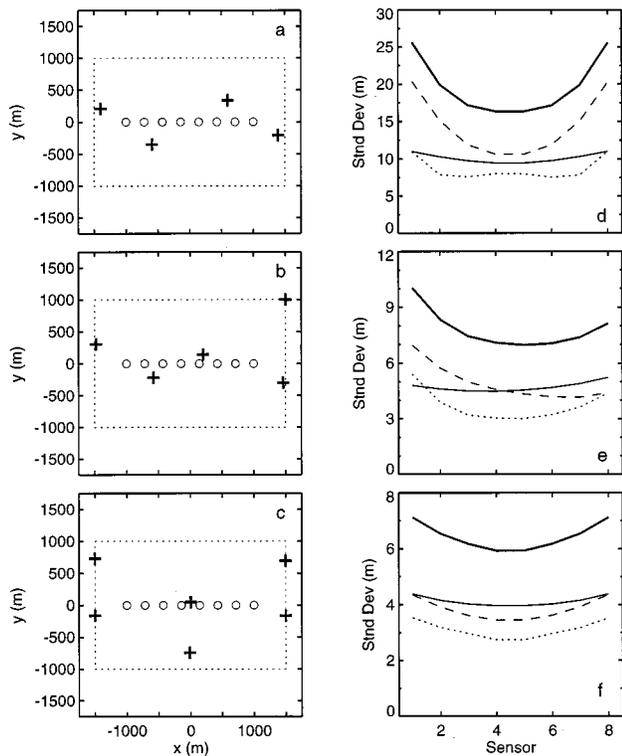


FIG. 5. Optimal AEL source positions (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for four, five, and six sources for a nonsynchronized AEL survey with errors in t of 0.5 ms, uncertainties in x' , y' of 5 m and in z' of 2 m, and a sound-speed uncertainty of 2 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E = 19.7$, 7.8, and 6.4 m, respectively.

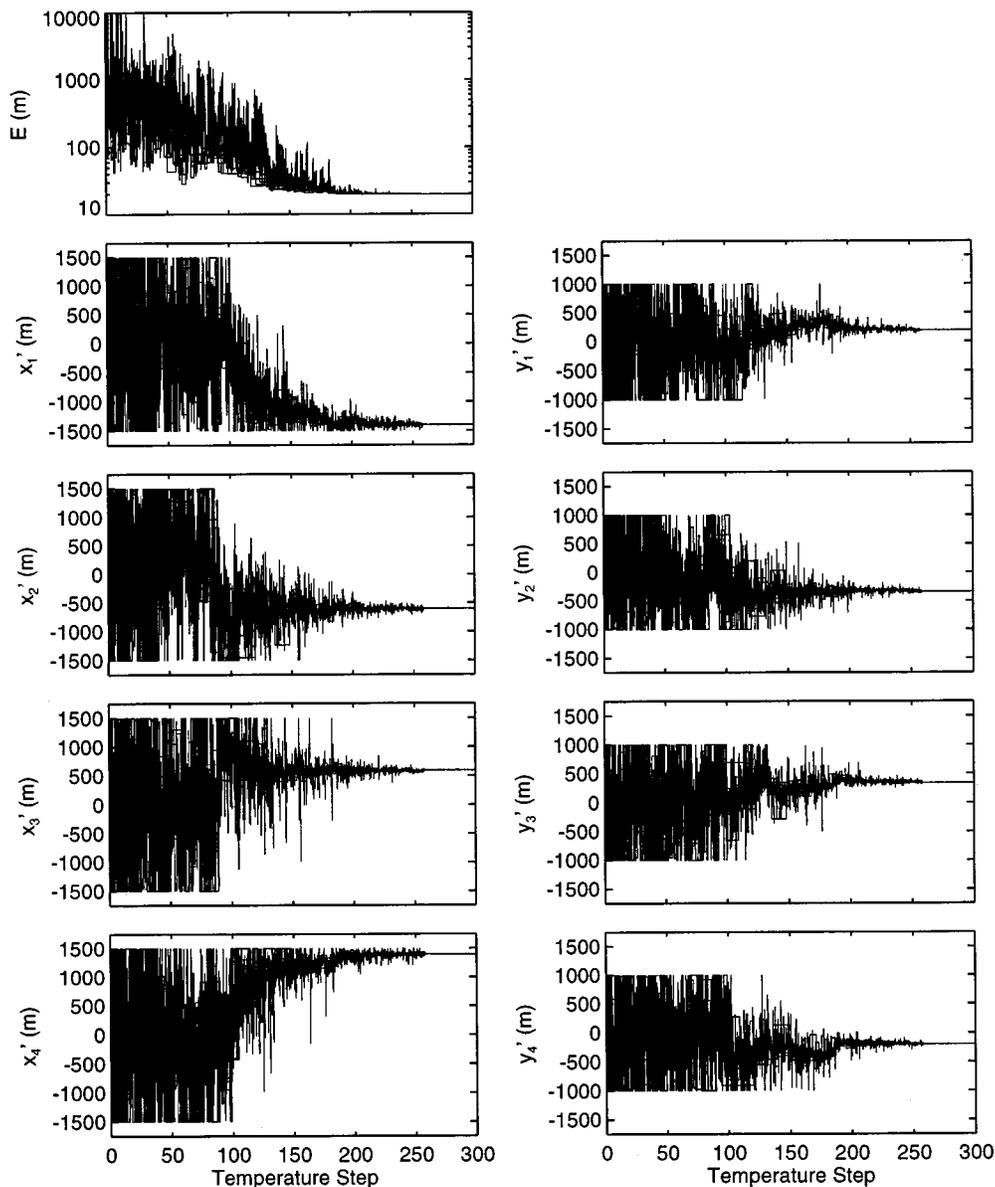


FIG. 6. Convergence of SSA optimization (four-source configuration) for the mean source-position error E and source coordinates $\{x'_i, y'_i, i=1,4\}$. All models in the simplex are shown.

good solutions. The optimization is then quenched to collapse the simplex to the single best model.

The optimal source configurations shown in Fig. 5 were determined by minimizing an error function based on a local linear approximation to the nonlinear AEL inverse problem, as described in Sec. I. To investigate the validity of this approximation, Fig. 7 compares the theoretical Gaussian *a posteriori* error distributions from the linear analysis (smooth curve) to error histograms computed from a nonlinear Monte Carlo simulation for the four-source case. The Monte Carlo simulation consisted of performing 10^5 independent inversions. In each inversion, random errors on the arrival-time data, on the *a priori* source-position estimates, and on the sound-speed bias were drawn from zero-mean Gaussian distributions with standard deviations equal to the uncertainty assigned to that quantity (given above). Each inversion was initiated from a randomly chosen starting model, and the regularized solution (5) was applied iteratively to conver-

gence. Figure 7 shows that the linear and nonlinear error distributions are virtually identical, which indicates that the linear uncertainty analysis is indeed applicable to the nonlinear AEL inverse problem.

As mentioned previously, optimal AEL survey design is based on determining the optimal source configuration corresponding to the expected array configuration. For example, the optimal source configurations shown in Fig. 5 were designed for the eight-element HLA in that figure, which represents the expected sensor configuration. In practice, however, the sensor positions are known only approximately after array deployment, typically to within ten to several hundred meters (hence the need for AEL). Since the expected and true array configurations differ by this amount, it is important to investigate the influence of errors in the expected sensor positions on the accuracy of AEL sensor localization. This issue is examined in Fig. 8 for the optimal five-source configuration [Fig. 5(b)]. Figure 8(a) illustrates

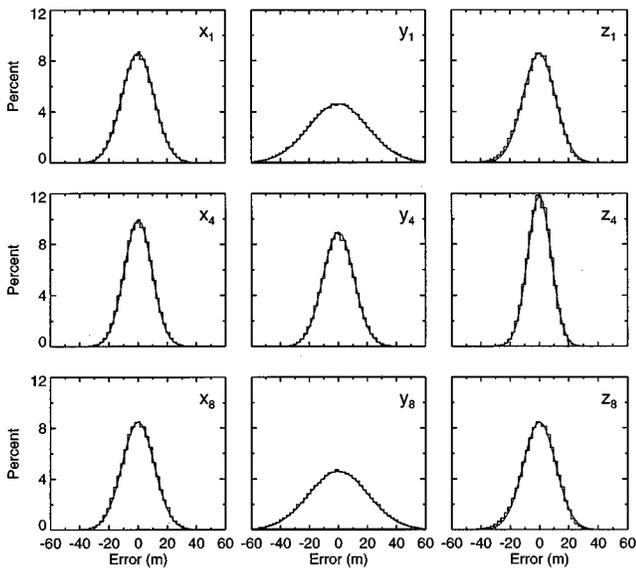


FIG. 7. Marginal *a posteriori* probability distributions for AEL errors in positions of HLA sensors 1, 4, and 8 for the optimum four-source configuration. The smooth curves represent theoretical Gaussian distributions from linear analysis; histograms represent results from nonlinear Monte Carlo simulation.

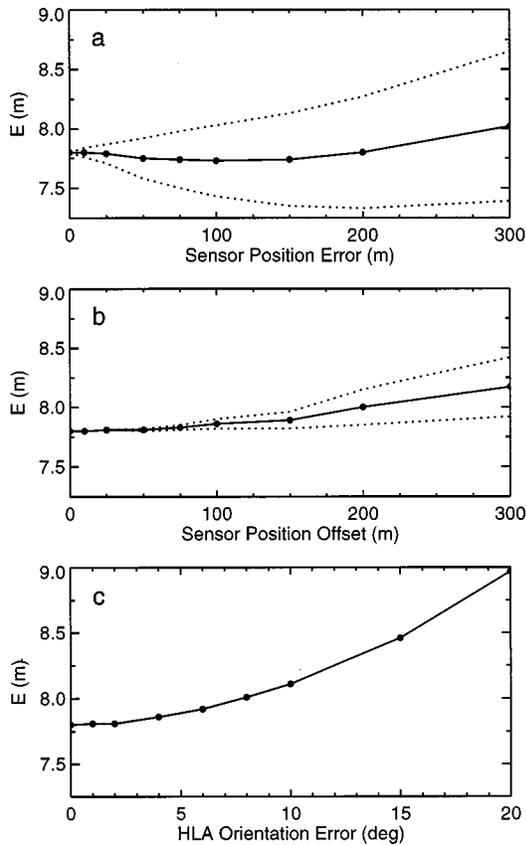


FIG. 8. Effect of errors in expected sensor positions on the AEL error E for the optimal five-source configuration. (a) shows the mean (solid line) and mean \pm one standard deviation (dotted lines) for E computed from 200 realizations of uniform random errors applied to individual sensor positions (horizontal scale indicates the amplitude of the error interval for x and y ; error amplitude for z is ten times smaller). (b) is similar to (a), except that the random errors consist of uniform offsets of the entire HLA in x , y , and z . (c) shows the results of errors in the orientation of the HLA axis.

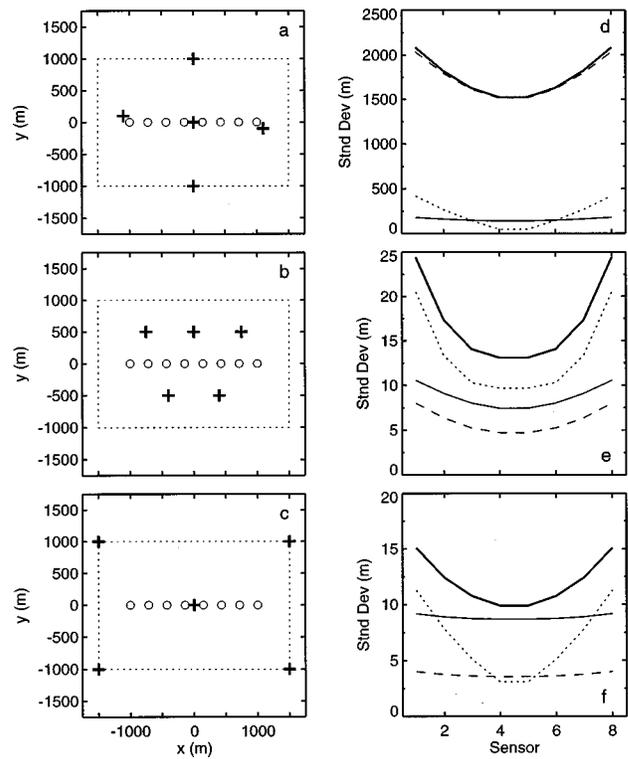


FIG. 9. Examples of nonoptimal configurations of five AEL sources (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for a nonsynchronized AEL survey with errors in t of 0.5 ms, uncertainties in x' , y' of 5 m and in z' of 2 m, and a sound-speed uncertainty of 2 m/s. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E = 1765$, 17.5, and 12.0 m, respectively.

the effect of random errors in the expected positions of individual sensors on the sensor-position error E . To determine each point on this figure, E was computed for 200 realizations of uniform random errors applied to each of the sensor-position coordinates. The resulting mean value for E and the mean \pm one standard deviation are shown in Fig. 8(a). Figure 8(b) is similar to Fig. 8(a), except that in this case the random errors consist of uniform offsets of the entire HLA in x , y and z . Finally, Fig. 8(c) shows the effect of errors in the orientation of the HLA axis from 0 to 20 degrees (note that for an orientation error of 20 degrees the positions of the sensors at either end of the HLA are in error by approximately 350 m). Since Fig. 8(c) does not represent the average of a collection of random errors, there is no standard deviation associated with the curve as in Fig. 8(a) and (b). Figure 8 shows that typical-to-large errors in the expected source configuration result in relatively small errors ($\leq 10\%$) in the sensor localization when the optimal source configuration is employed. It was found that combining the various types of errors considered in Fig. 8 typically lead to smaller values of E than those shown. The effect of errors in the expected source positions was also found to decrease as the number of sources was increased (i.e., the effect was somewhat greater for the optimal four-source configuration, but was smaller for the six-source configuration).

While the effect of errors in the expected sensor positions are relatively small, the use of nonoptimal source con-

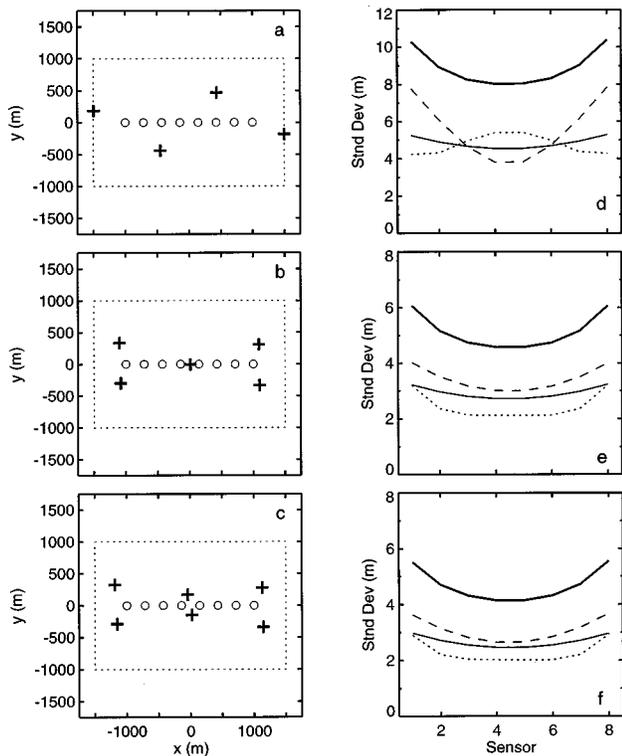


FIG. 10. Optimal AEL source positions (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for four, five, and six sources for a nonsynchronized AEL survey with errors in t of 0.05 ms, uncertainties in x' , y' of 5 m, and in z' of 2 m, and a sound-speed uncertainty of 2 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E=8.9$, 5.1, and 4.7 m, respectively.

figurations can lead to large AEL errors. For example, Fig. 9 shows three five-source configurations that are substantially worse than the optimal configuration [Fig. 5(b) and (e), $E=7.8$ m]. The source configuration in Fig. 9(a) provides exceedingly poor sensor localization, as indicated by the large *a posteriori* uncertainties in Fig. 9(d) and the high value of $E=1765$ m. Performing AEL with this source configuration is clearly worthless. This case illustrates that without qualitative analysis, seemingly reasonable source configurations can produce poor AEL results (it is interesting to note, however, that this configurations provides good sensor localization for synchronized AEL surveys). Figure 9(b) and (c) shows two other source configurations which, although substantial improvements on the configuration in Fig. 9(a), provide sensor localization that is significantly worse than the optimal configuration [$E=17.5$ and 12.0 m for Fig. 9(b) and (c) respectively].

Since the sensor-position error E depends on the errors associated with the measured data, the *a priori* source positions, and the sound-speed bias, the optimal source configuration can change with changes in these factors. The optimal configurations shown in Fig. 5 were computed for errors in t of 0.5 ms, uncertainties in x' , y' of 5 m and in z' of 2 m, and a sound-speed uncertainty of 2 m/s. For comparison, optimal configurations are shown in Fig. 10 for errors in t of 0.05 ms, and in Fig. 11 for uncertainties in x' , y' , z' of 1 m.

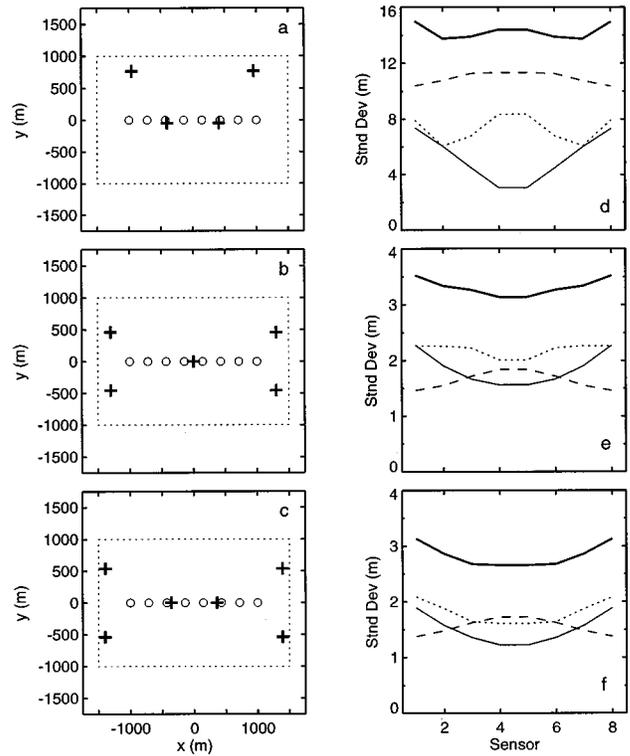


FIG. 11. Optimal AEL source positions (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for four, five, and six sources for a nonsynchronized AEL survey with errors in t of 0.5 ms, uncertainties in x' , y' , z' of 1 m, and a sound-speed uncertainty of 2 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E=14.3$, 3.3, and 2.8 m, respectively.

Changing the sound-speed error was found to have an insignificant effect on the optimal configuration for nonsynchronized AEL, likely because the effect of the sound-speed error is partially absorbed in the computed source instants. Comparison of Figs. 5, 10, and 11 indicates that, depending on the number of sources, the source configurations may or may not change with changes in data or source-position errors. Note that for given data and source-position errors, the actual difference in sensor localization is not great for the various configurations of five sources or of six sources, as documented in Table I. This indicates that, for reasonably well-conditioned inversions, near-optimal source configurations can be determined even if the data or source-position errors are not well known at the design stage.

2. Synchronized surveys

The optimal source configurations shown to this point have been computed for nonsynchronized AEL surveys (i.e., unknown source instants). The additional information provided by measuring source instants allows AEL surveys to be carried out with fewer sources and can lead to more accurate sensor localization. For example, Fig. 12 shows the optimal configurations of three, four, and five sources for a synchronized AEL survey with errors in t of 0.5 ms, uncertainties in x' , y' , z' of 1 m, and a sound-speed uncertainty of 2 m/s. These are the same uncertainties as for the configu-

TABLE I. The AEL errors for various source configurations and errors in the data and source positions. The left-most column gives the data and source-position errors for which the source configuration was optimized. The data errors are 0.5 or 0.05 ms; the source-position errors are 5 m for x' , y' and 2 m for z' , or 1 m for x' , y' , z' . The next three columns give the mean sensor-position error E for these source configurations, and for the data and sensor-position errors given in the column heading. In each case (i.e., each column), E is minimum for the error values for which the configuration was optimized. Note, however, that for five and six sources, the differences are not large.

Configuration optimized for	E (m) 0.5 ms 5/2 m	E (m) 0.05 ms 5/2 m	E (m) 0.5 ms 1 m
Four sources			
0.5 ms, 5/2 m	19.7	9.3	17.7
0.05 ms, 5/2 m	22.0	8.9	20.3
0.5 ms, 1 m	31.1	18.0	14.2
Five sources			
0.5 ms, 5/2 m	7.80	5.84	3.92
0.05 ms, 5/2 m	9.43	5.13	3.46
0.5 ms, 1 m	8.94	5.35	3.31
Six sources			
0.5 ms, 5/2 m	6.44	5.07	3.05
0.05 ms, 5/2 m	7.43	4.68	3.00
0.5 ms, 1 m	6.92	4.91	2.83

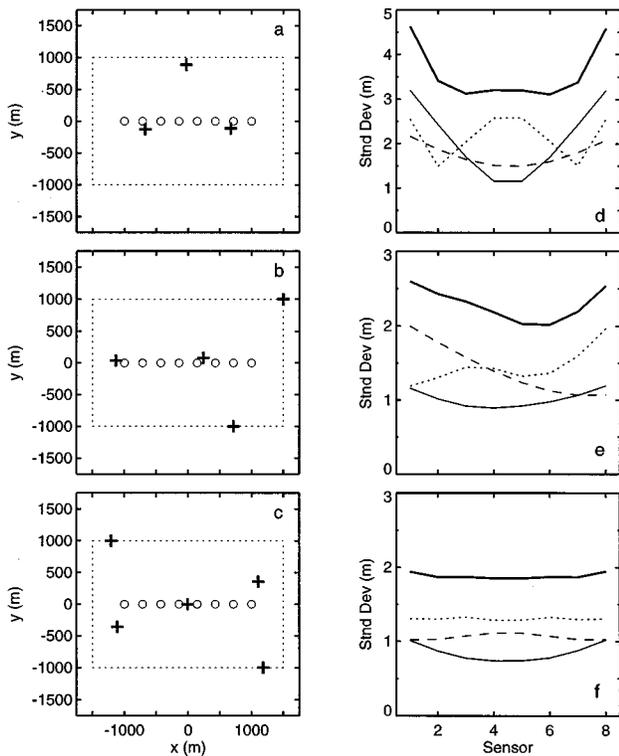


FIG. 12. Optimal AEL source positions (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for three, four, and five sources for a synchronized survey with errors in t of 0.5 ms, uncertainties in x' , y' , z' of 1 m, and a sound-speed uncertainty of 2 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E=3.6$, 2.3, and 1.9 m, respectively.

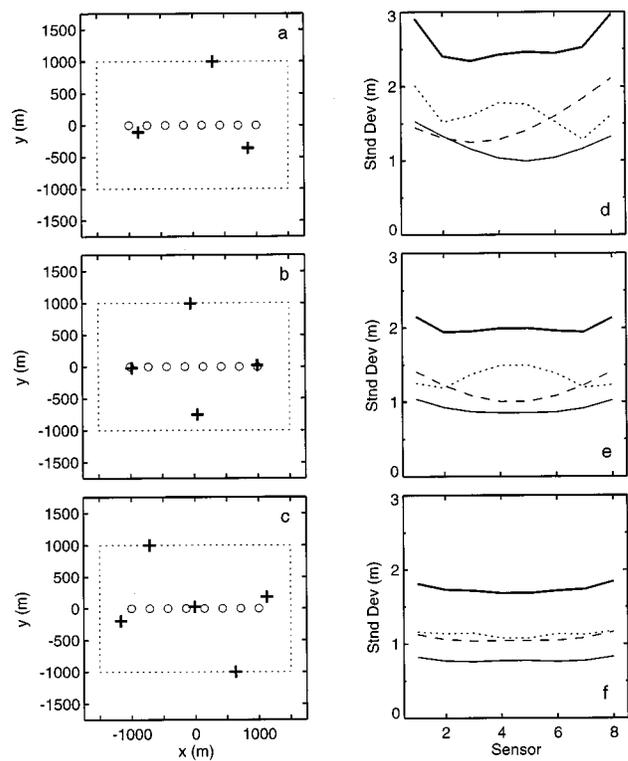


FIG. 13. Optimal AEL source positions (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for three, four, and five sources for a synchronized AEL survey with errors in t of 0.5 ms, uncertainties in x' , y' , z' of 1 m, and a sound-speed uncertainty of 0.25 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E=2.5$, 2.0, and 1.7 m, respectively.

rations in Fig. 11 (nonsynchronized survey), yet the source-position standard deviations shown in Fig. 12(d)–(f) are substantially smaller. The combination of known source instants and small source-position uncertainties was found to be the only case that was significantly affected by the uncertainty in the water sound speed. Figure 13 shows the optimal source configurations computed for the same data and source-position uncertainties as Fig. 12, but with a sound-speed uncertainty of only 0.25 m/s. The configurations shown in Fig. 13 differ from those in Fig. 12, and the sensor-position standard deviations are significantly smaller.

So far, optimal source configurations have been designed for uniform errors on the arrival-time data (i.e., all data have the same standard deviation). The effects of non-uniform errors can be included in the optimization through the data covariance matrix \mathbf{C}_t . An example of this procedure is given in Fig. 14 for a case in which the standard deviation of the i th arrival-time measurement increases linearly with source-sensor range r_i according to $\{\mathbf{C}_t\}_{ii}^{1/2} = mr_i + b$, with $m = 0.002$ ms/m and $b = 0.05$ ms (this relationship could represent empirical observations of the increase in arrival-time errors due to decreasing signal-to-noise ratio with range). The source configurations shown in Fig. 14(a)–(c) are considerably more compact than those in Fig. 13 (uniform errors) as the optimization procedure attempts to reduce the effects of large data errors at long ranges. The resulting sensor-position standard deviations, shown in Fig. 14(d) and

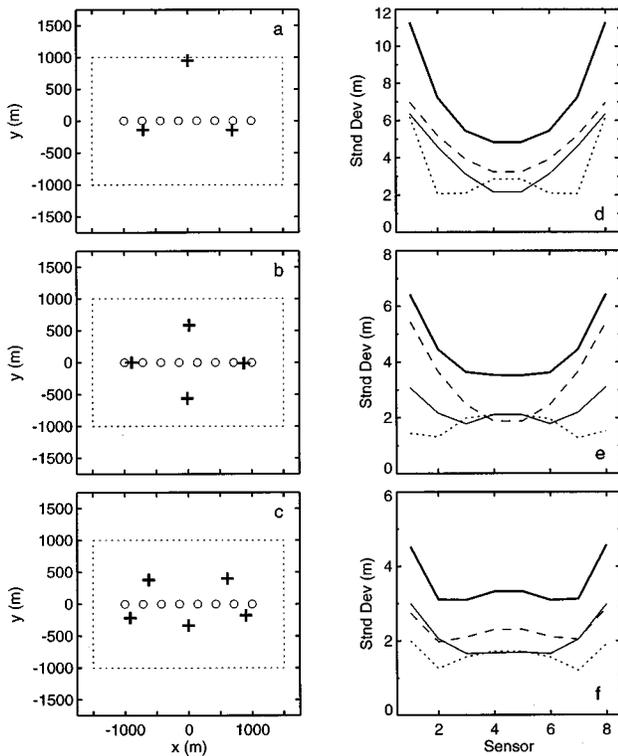


FIG. 14. Optimal AEL source positions (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for three, four, and five sources for a synchronized AEL survey with errors in t that increase linearly with range (see text), uncertainties in x' , y' , z' of 1 m, and a sound-speed uncertainty of 0.25 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E=7.2$, 4.5, and 3.5 m, respectively.

(e), are much larger for the sensors near the ends of the HLA than for those near the array center. In a case like this, it may be preferable to determine the source configuration that minimizes the maximum sensor-position error E_M given by (11), rather than the mean error E given by (10), in an attempt to decrease the discrepancies in the sensor-position uncertainties. The results of minimizing E_M for this case are given in Fig. 15. The optimal source configurations shown in Fig. 15(a)–(c) are less compact than those in Fig. 14, and in each case extend to or beyond the ends of the HLA. This reduces the sensor-position uncertainties near the ends of the HLA at the expense of increasing the uncertainties near the array center, as shown in Fig. 15(d)–(f).

This section has shown a number of examples of both optimal and nonoptimal source configurations for HLA sensor localization. In an attempt to quantify the benefits of employing optimal configurations, Fig. 16(a) shows the ratio E_{ave}/E_{opt} as a function of number of sources for both synchronized and nonsynchronized surveys, where E_{ave} is obtained by averaging E over 1000 random choices of the source configurations and E_{opt} is obtained for the optimal configuration. (While source configuration designed by an experienced investigator should be better than a random choice, the use of E_{ave} provides a consistent baseline for comparison.) Figure 16(a) shows that the benefit of employing the optimal source configuration is greatest for small

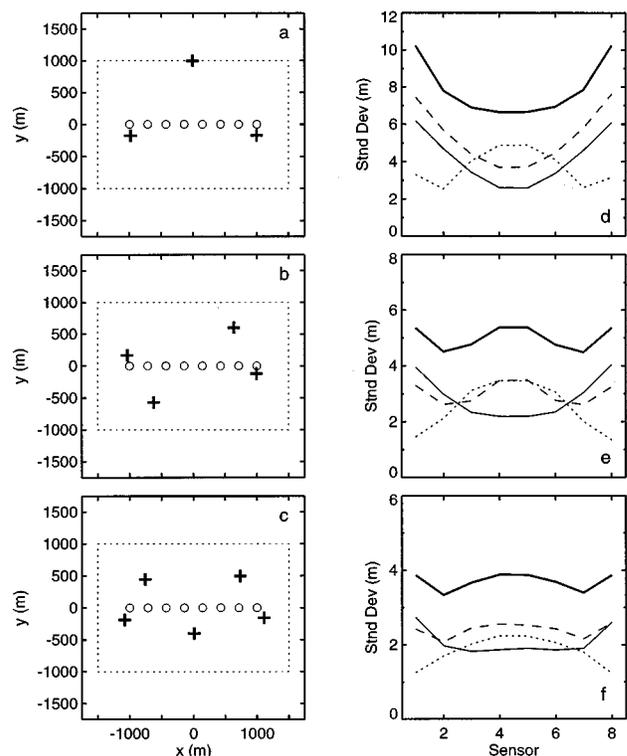


FIG. 15. Source configurations that minimize the maximum source-position error E_M . Source positions (crosses) for localizing HLA sensors (circles) are shown in (a)–(c) for three, four, and five sources for a synchronized AEL survey with errors in t that increase linearly with range, uncertainties in x' , y' , z' of 1 m, and a sound-speed uncertainty of 0.25 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dashed line), and σ_z (dotted line) are shown in (d)–(f) with $E_M=10.2$, 5.4, and 3.9 m (mean source-position errors $E=7.9$, 5.0, and 3.7), respectively.

numbers of sources and for nonsynchronized surveys. The results of synchronized and nonsynchronized surveys are compared further in Fig. 16(b), which shows E_{opt} as a function of the number of sources for these two cases. For small numbers of sources, the results of synchronized surveys are substantially better; however, the difference decreases as the number of sources increases. Plots such as this could be used in AEL survey design. For example, if the maximum acceptable sensor-position error is $E=5$ m, Fig. 16(b) shows that this can be achieved using five sources in a synchronized AEL survey or ten sources in a nonsynchronized survey. A choice between the two surveys could then be made based on the relative effort and cost of synchronization and source deployment.

B. Optimal AEL for vertical arrays

This section considers optimal design for AEL systems that use seafloor acoustic sources (e.g., transponders) to localize the sensors of a vertical array [Fig. 1(a)]. The ocean sound-speed profile for the scenario considered here is shown in Fig. 4, and is assumed to be known to within a potential bias of 1 m/s. Three VLA sensors at depths of 400, 275, and 150 m (referred to as sensors 1, 2 and 3, respectively) are localized to determine the array shape. The mea-

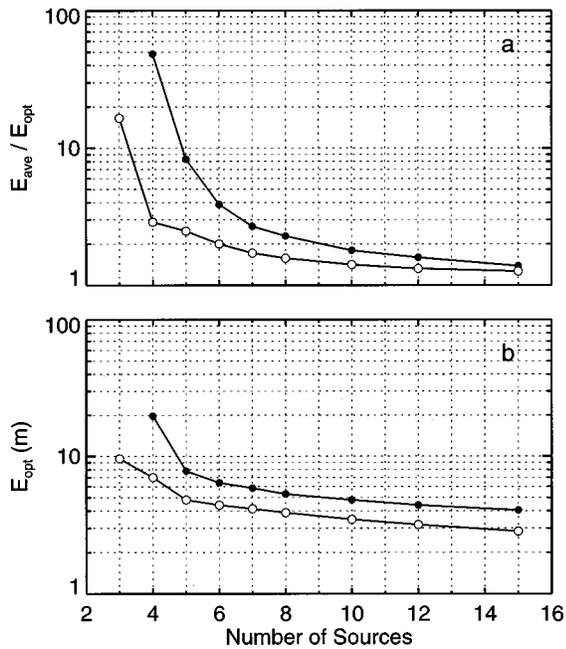


FIG. 16. Comparisons of the results of synchronized (open circles) and nonsynchronized (filled circles) AEL surveys for HLA localization (data errors are 0.5 ms; source-position errors are 5 m for x' , y' and 2 m for z' ; sound-speed errors are 2 m/s). The ratio $E_{\text{ave}}/E_{\text{opt}}$ shown in (a) quantifies the benefits of employing optimal source configurations (see text). (b) shows the mean source-position error E_{opt} for optimal configurations.

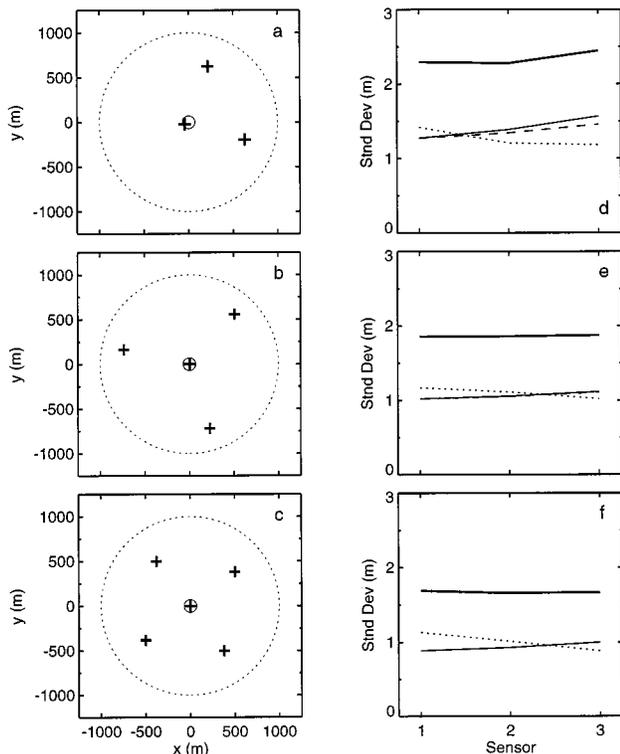


FIG. 17. Optimal AEL source positions (crosses) for localizing three VLA sensors (circle) are shown in (a)–(c) for three, four, and five sources for a synchronized AEL survey with errors in t of 0.5 ms, uncertainties in x' , y' , z' of 1 m, and a sound-speed uncertainty of 1 m/s. Dotted lines represent the constraint on source positions. Corresponding *a posteriori* standard deviations for the sensor positions σ_r (heavy solid line), σ_x (solid line), σ_y (dotted line) are shown in (d)–(f) with mean sensor-position errors of $E=2.3$, 1.9, and 1.7 m, respectively. Note that in (e) and (f) the solid and dashed lines exactly coincide.

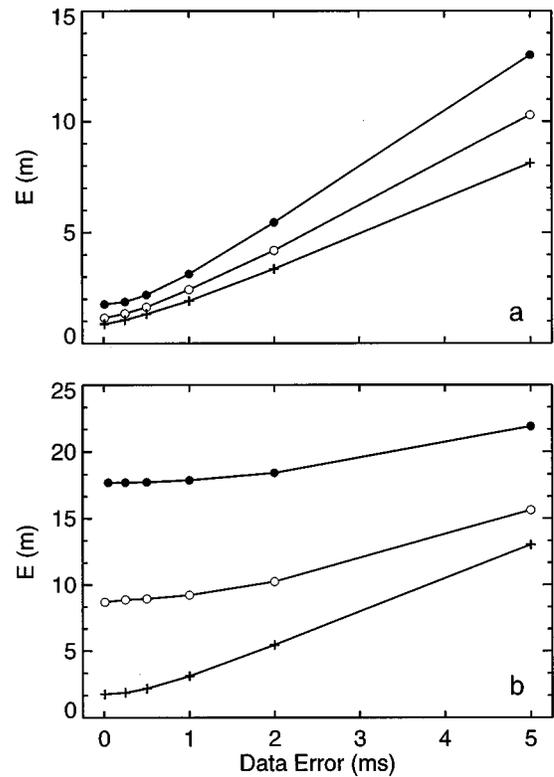


FIG. 18. Mean sensor-position error E for three VLA sensors as a function of data error. (a) shows the results for source-position errors of 1 m and three sources (filled circles), five sources (open circles), and eight sources (crosses). (b) shows the results for three sources and source position errors of 10 m (filled circles), 5 m (open circles), and 1 m (crosses). In each case, the sound-speed error is 1 m/s and the sources are in the optimal configuration.

sured data consist of absolute travel times and the data errors are uncorrelated with a standard deviation of 0.5 ms. The source positions are known to within an uncertainty in x' , y' , z' of 1 m and the source-position errors are uncorrelated. The SSA optimization algorithm was applied to minimize the mean sensor-position error E with the constraint that the sources must be within a radial distance of 1000 m from the VLA. Figure 17 shows the optimal configurations for three, four, and five sources together with the corresponding *a posteriori* standard deviations for each sensor. Repeating the SSA optimization with different random model perturbations produced source configurations that were essentially identical to those in Fig. 17, up to an arbitrary rotation about the origin (i.e., the VLA site).

The radial symmetry of AEL for a vertical array leads to symmetric optimal source configurations for four and five sources [Fig. 17(b) and (c)]. This symmetry allows the precision of the optimization to be examined as follows. The four-source configuration [Fig. 17(b)] consists of one source almost directly below the VLA (range 0.03 m) and three sources separated by 120.0 degrees at ranges of 753, 754, and 755 m; the mean sensor-position error (to machine precision) is $E=1.86268$ m. Assuming that one source directly below the VLA and three others at 120-degree separation and a uniform range represents the optimal configuration, a 1-D search was carried out for the range that minimized E . The resulting range of 754 m and corresponding value of E

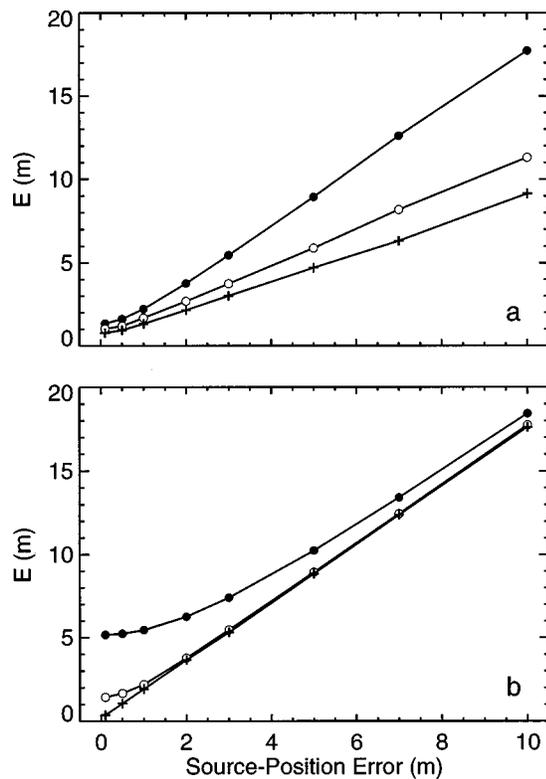


FIG. 19. Mean sensor-position error E for three VLA sensors as a function of source-position error. (a) shows the results for data errors of 0.5 ms and three sources (filled circles), five sources (open circles), and eight sources (crosses). (b) shows the results for three sources and data errors of 2 ms (filled circles), 0.5 ms (open circles), and 0.05 ms (crosses). In each case, the sound-speed error is 1 m/s and the sources are in the optimal configuration.

=1.862 68 m indicates the SSA optimization result is highly precise. A similar analysis of the five-source configuration [Fig. 17(c)] indicated an equally high level of precision.

The ability to determine optimal AEL source configurations as a function of the number of sources and of the errors in the data, source positions, and sound speed allows the effects of each of these factors to be examined in a consistent manner. A modeling study considering these factors can guide the design of an AEL system to meet specific objectives, such as localizing the sensors to within a given uncertainty. An example of this procedure is presented in Figs. 18–21 for the VLA and ocean environment described above. Figure 18 illustrates the variation of the mean sensor-position error E as a function of the standard deviation of the measured data. This relationship is shown in Fig. 18(a) for cases of three, five, and eight sources with 1-m source-position errors, and in Fig. 18(b) for three sources and source-position errors of 1, 5, and 10 m (in each case, E is computed for the optimal source configuration). As would be expected, E increases with data errors, and lower values of E are obtained as the number of sources is increased or the source-position error is decreased.

Figure 19 shows E as a function of the error in the source-position estimates for three, five, and eight sources with data errors of 0.5 ms [Fig. 19(a)], and for three sources with data errors of 2, 0.5, and 0.05 ms [Fig. 19(b)]. (Note that the underlying assumption is that the source-position errors are random; a constant translation common to all

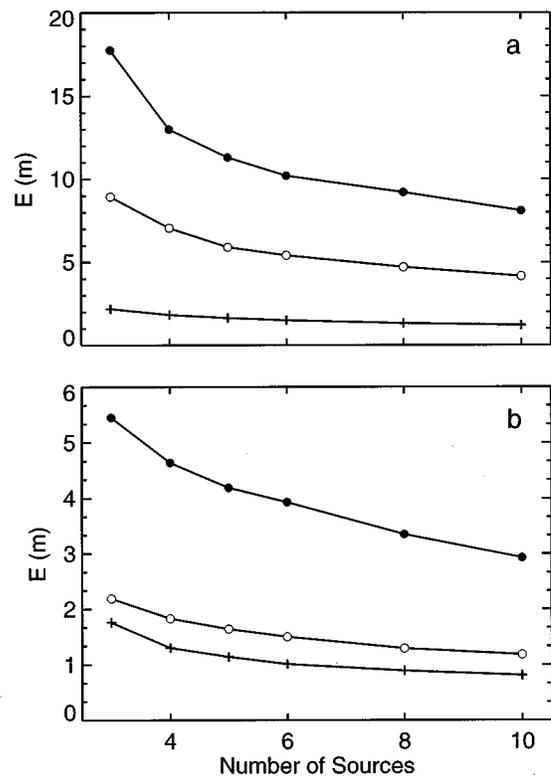


FIG. 20. Mean sensor-position error E for three VLA sensors as a function of number of sources. (a) shows the results for data errors of 0.5 ms and source-position errors of 10 m (filled circles), 5 m (open circles), and 1 m (crosses). (b) shows the results for source-position errors of 1 m and data errors of 2 ms (filled circles), 0.5 ms (open circles), and 0.05 ms (crosses). In each case, the sound-speed error is 1 m/s and the sources are in the optimal configuration.

sources of ≤ 10 m has a negligible effect on the sensor localization.) Figure 19 illustrates the futility of attempting to improve AEL by decreasing the data errors when significant source-position uncertainties exist. For example, Fig. 19(b) shows that for three sources with a source-position uncertainty of 10 m and data errors of 2 ms, the mean sensor-position error is $E=18.5$ m; reducing the data errors by a factor of 40 to 0.05 ms results in only a small improvement in sensor localization to $E=17.6$ m. However, the sensor-position error can be reduced by increasing the number of sources: Fig. 19(a) shows that for source-position errors of 10 m and data errors of 0.5 ms, E is reduced from 17.9 to 9.1 m by increasing the number of sources from three to eight. Figure 20 further investigates the dependence of E on the number of sources for various data and source-position uncertainties. In Figs. 18–20, the standard deviation of the sound-speed bias was taken to be 1 m/s. The results of these figures were found to be relatively insensitive to the sound-speed uncertainty. The effect of sound-speed error is further investigated in Fig. 21, which shows E as a function of this error for various data and source-position uncertainties.

III. SUMMARY

The configuration of source positions is an important aspect of designing AEL surveys that has a substantial effect on the accuracy of sensor localization. In this paper, a method was developed to determine optimal AEL source

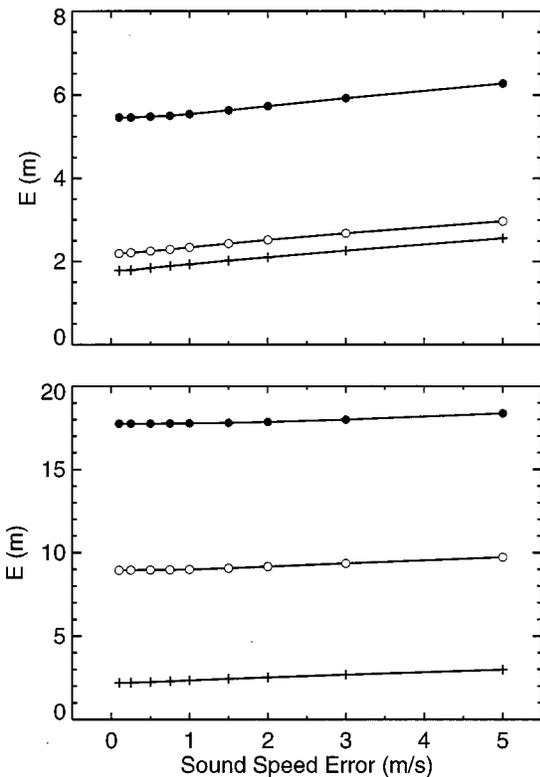


FIG. 21. Mean sensor-position error E for three sources and three VLA sensors as a function of sound-speed errors. (a) shows the results for a source-position error of 1 m and data errors of 2 ms (closed circles), 0.5 ms (open circles), and 0.05 ms (crosses). (b) shows the results for a data error of 0.5 ms and source-position errors of 10 m (closed circles), 5 m (open circles), and 1 m (crosses). In each case, the sources are in the optimal configuration.

configurations. To this end, an error measure was defined based on the *a posteriori* sensor-position uncertainties derived from a general formulation of the AEL inverse problem. Minimizing this sensor-position error with respect to the source positions provides the optimal AEL source configuration. Since this represents a challenging nonlinear minimization problem, an efficient hybrid optimization algorithm was implemented. The method for determining optimal source configurations is general, and can be applied for any sensor configuration, for relative or absolute travel-time data, and for any combination of errors in the data, source positions, and sound speed. It is also straightforward to apply physical constraints to the source positions, or to include the effects of data errors that vary with range. Two error measures, the mean 3-D sensor-position error and the maximum 3-D sensor-position error, were considered and shown to provide sensor localization with somewhat different properties. The results do not appear to be sensitive to the number of sensors being localized.

By definition, the optimal AEL source configuration provides sensor-position errors that are smaller than those for any other source configuration. Heuristic source configurations (i.e., configurations designed by trial-and-error using an investigator's experience and insight) can vary from almost as good to much worse than the optimal configuration. Defining an AEL error measure provides a basis for comparing different source configurations, with the optimal configura-

tion providing an absolute standard for such comparisons. In addition, it was shown that the ability to determine optimal source configurations as a function of the type of data, the number of sources, and the errors in the data, source positions, and sound speed allows the effects of each of these factors to be examined quantitatively in a consistent manner. Such a modeling study can guide in designing an appropriate AEL system to meet specific sensor-localization objectives. Examples were given of these procedures for both horizontal and vertical arrays of sensors.

ACKNOWLEDGMENTS

This work was supported by SPARWARSYSCEN Research Contract No. N66001-93-D-0144, and by the National Science and Engineering Research Council (NSERC) and the Department of National Defense (DND) through a NSERC research grant and the NSERC/DND Ocean Acoustics Chair Program at the University of Victoria.

APPENDIX: FORWARD MAPPING AND PARTIAL DERIVATIVES

To define explicitly the AEL forward mapping, the arrival time t_k corresponding to i th source and j th sensor is given by

$$t_k = T_k(\mathbf{m}) = \tau(x'_i, y'_i, z'_i, x_j, y_j, z_j, c) + t_i^0, \quad (\text{A1})$$

where τ represents the travel time along the direct acoustic ray between source and sensor for an ocean sound-speed profile c , and t_i^0 is the i th source instant. For a sound-speed profile that varies only with depth, the range and travel time along a ray path are given by³⁰

$$r = \int_{z'}^z \frac{pc(u) du}{[1 - p^2 c^2(u)]^{1/2}}, \quad (\text{A2})$$

$$\tau = \int_{z'}^z \frac{du}{c(u)[1 - p^2 c^2(u)]^{1/2}}. \quad (\text{A3})$$

In (A2) and (A3), the ray parameter $p = \cos \theta(u)/c(u)$ is constant along a ray path, and defines the take-off (grazing) angle at the source. The ray parameter for an eigenray connecting source and receiver is determined by searching for the value of p which produces the correct range (to a specified tolerance) using (A2). Our algorithm employs an efficient procedure of determining p for direct-path eigenrays based on Newton's method.¹²

Partial derivatives of the travel time τ (and hence the forward mapping T) with respect to source and receiver coordinates are obtained by differentiating (A3) employing Leibnitz's rule, the chain rule, and the fact that $\partial r / \partial z = \partial r / \partial z' = 0$ to yield

$$\frac{\partial T}{\partial x} = p(x - x')/r, \quad \frac{\partial T}{\partial x'} = p(x' - x)/r, \quad (\text{A4})$$

$$\frac{\partial T}{\partial y} = p(y - y')/r, \quad \frac{\partial T}{\partial y'} = p(y' - y)/r, \quad (\text{A5})$$

$$\frac{\partial T}{\partial z} = [1 - p^2 c^2(z)]^{1/2} / c(z),$$

$$\frac{\partial T}{\partial z'} = -[1 - p^2 c^2(z')]^{1/2} / c(z'),$$
(A6)

To account for bias in the measured sound-speed profile, let $c(u) = c_t(u) + c_b$, where c_t is the true sound speed and c_b is the bias. Differentiating (A3) with respect to c_b leads to

$$\frac{\partial T}{\partial c_b} = - \int_{z'}^z \frac{du}{c_t^2(u) [1 - p^2 c^2(u)]^{1/2}}.$$
(A7)

Finally, if the source instants t^0 in (A1) are unknown (non-synchronized AEL measurements), they are included as parameters in the inversion with partial derivatives

$$\frac{\partial T}{\partial t^0} = 1.$$
(A8)

- ¹W. S. Burdick, *Underwater Acoustic System Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1994).
- ²A. Tolstoy, *Matched Field Processing for Underwater Acoustics* (World Scientific, Singapore, 1993).
- ³A. 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).
- ⁴M. D. Collins, W. A. Kuperman, and H. Schmidt, "Nonlinear inversion for ocean-bottom properties," *J. Acoust. Soc. Am.* **92**, 2770–2783 (1992).
- ⁵S. E. Dosso, M. L. Yeremy, J. M. Ozard, and N. R. Chapman, "Estimation of ocean bottom properties by matched field inversion of acoustic field data," *IEEE J. Ocean Eng.* **18**, 232–239 (1993).
- ⁶B. D. Steinberg, *Principles of Aperture and Array System Design* (Wiley, New York, 1976).
- ⁷W. S. Hodgkiss, D. E. Ensberg, J. J. Murray, G. L. D'Spain, N. O. Booth, and P. W. Schey, "Direct measurement and matched-field inversion approaches to array shape estimation," *IEEE J. Ocean Eng.* **21**, 393–401 (1996).
- ⁸B. J. Sotirin and J. A. Hildebrand, "Large-aperture digital acoustic array," *IEEE J. Ocean Eng.* **13**, 271–281 (1988).
- ⁹B. J. Sotirin and J. A. Hildebrand, "Acoustic navigation of a large-aperture array," *J. Acoust. Soc. Am.* **87**, 154–167 (1990).
- ¹⁰B. J. Sotirin, V. K. McDonald, J. R. Olsen, J. M. Stevenson, and G. O. Pickens, "An integrated array element localization system," in *Proc. Oceans '94* (1994), pp. 727–731.
- ¹¹G. H. Brooke, S. J. Kilistoff, and B. J. Sotirin, "Array element localization algorithms for vertical line arrays," *Proceedings of the Third European Conference on Underwater Acoustics*, edited by J. S. Papadakis (Crete U.P., Heraklion, 1996), pp. 537–542.

- ¹²S. E. Dosso, G. H. Brooke, S. J. Kilistoff, B. J. Sotirin, V. K. McDonald, M. R. Fallat, and N. E. Collison, "High-precision Array Element Localization of Vertical Line Arrays in the Arctic Ocean," *IEEE J. Ocean Eng.* **23**, 365–379 (1998).
- ¹³K. C. Creager and L. M. Dorman, "Location of instruments on the seafloor by joint adjustment of instrument and ship positions," *J. Geophys. Res.* **87**, 8379–8388 (1982).
- ¹⁴W. S. Hodgkiss, "Shape determination of a shallow-water bottomed array," in *Proc. Oceans '89* (1989), pp. 1199–1204.
- ¹⁵S. 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).
- ¹⁶G. J. Heard, M. McDonald, N. R. Chapman, and L. Jaschke, "Underwater light bulb implosions: A useful acoustic source," *Proc. Oceans '97* (1997), Vol. 2, pp. 755–762.
- ¹⁷H. T. Vincent II and S.-L. J. Hu, "Geodetic position estimation of underwater acoustic sensors," *J. Acoust. Soc. Am.* **102**, 3099 (1998).
- ¹⁸W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN* (Cambridge U. P., Cambridge, 1992).
- ¹⁹P. Liu, S. Hartzell, and W. Stephenson, "Non-linear multiparameter inversion using a hybrid global search algorithm: Applications in reflection seismology," *Geophys. J. Int.* **122**, 991–1000 (1995).
- ²⁰M. R. Fallat and S. E. Dosso, "Geoacoustic inversion via local, global and hybrid algorithms," *J. Acoust. Soc. Am.* **105**, 3219–3230 (1999).
- ²¹A. Tolstoy, "Linearization of the matched field processing approach to acoustic tomography," *J. Acoust. Soc. Am.* **91**, 781–787 (1992).
- ²²A. G. Jones and J. H. Foster, "An objective real-time data-adaptive technique for efficient model resolution improvement in magnetotelluric studies," *Geophysics* **51**, 90–97 (1986).
- ²³M. Hardt and F. Scherbaum, "Design of optimum networks for aftershock recording," *Geophys. J. Int.* **117**, 716–726 (1994).
- ²⁴M. Hansruedi and D. E. Boerner, "Optimized and robust experimental design: a non-linear application of EM sounding," *Geophys. J. Int.* **132**, 458–468 (1998).
- ²⁵L. R. Lines and S. Treitel, "Tutorial: A review of least-squares inversion and its application to geophysical inverse problems," *Geophys. Prospect.* **32**, 159–186 (1984).
- ²⁶C. van Schooneveld, "Inverse problems: A tutorial survey," in *Underwater Acoustic Data Processing*, edited by Y. T. Chan (Kluwer, The Netherlands, 1989), pp. 393–411.
- ²⁷W. Munk, P. Worcester, and C. Wunsch, *Ocean Acoustic Tomography* (Cambridge U. P., Cambridge, 1995).
- ²⁸H. Szu and R. Hartley, "Fast simulated annealing," *Phys. Lett. A* **122**, 157–162 (1987).
- ²⁹F. Lane, "Estimation of the kinematic rupture parameters from historical seismograms: An application of simulated annealing to a non-linear optimization problem," Ph.D. thesis, Colorado School of Mines, Golden, CO, 1992.
- ³⁰W. M. Telford, L. P. Geldart, R. E. Sheriff, and D. A. Keys, *Applied Geophysics* (Cambridge U.P., New York, 1976).

Responses of neurons to click-pairs as simulated echoes: Auditory nerve to auditory cortex

Douglas C. Fitzpatrick^{a)} and Shigeyuki Kuwada

University of Connecticut Health Center, Department of Anatomy, Farmington, Connecticut 06030

D. O. Kim and Kourosh Parham

University of Connecticut Health Center, Surgery/Otolaryngology, Farmington, Connecticut 06030

Ranjan Batra

University of Connecticut Health Center, Department of Anatomy, Farmington, Connecticut 06030

(Received 18 September 1998; revised 18 August 1999; accepted 16 September 1999)

When two identical sounds are presented from different locations with a short interval between them, the perception is of a single sound source at the location of the leading sound. This “precedence effect” is an important behavioral phenomenon whose neural basis is being increasingly studied. For this report, neural responses were recorded to paired clicks with varying interstimulus intervals, from several structures of the ascending auditory system in unanesthetized animals. The structures tested were the auditory nerve, anteroventral cochlear nucleus, superior olivary complex, inferior colliculus, and primary auditory cortex. The main finding is a progressive increase in the duration of the suppressive effect of the leading sound (the conditioner) on the response to the lagging sound (the probe). The first major increase occurred between the lower brainstem and inferior colliculus, and the second between the inferior colliculus and auditory cortex. In neurons from the auditory nerve, cochlear nucleus, and superior olivary complex, 50% recovery of the response to the probe occurred, on average, for conditioner and probe intervals of ~ 2 ms. In the inferior colliculus, 50% recovery occurred at an average separation of ~ 7 ms, and in the auditory cortex at ~ 20 ms. Despite these increases in average recovery times, some neurons in every structure showed large responses to the probe within the time window for precedence (~ 1 – 4 ms for clicks). This indicates that during the period of the precedence effect, some information about echoes is retained. At the other extreme, for some cortical neurons the conditioner suppressed the probe response for intervals of up to 300 ms. This is in accord with behavioral results that show dominance of the leading sound for an extended period beyond that of the precedence effect. Other transformations as information ascended included an increased variety in the shapes of the recovery functions in structures subsequent to the nerve, and neurons “tuned” to particular conditioner–probe intervals in the auditory cortex. These latter are reminiscent of neurons tuned to echo delay in bats, and may contribute to the perception of the size of the acoustic space. © 1999 Acoustical Society of America. [S0001-4966(99)07612-2]

PACS numbers: 43.64.Bt, 43.66.Ba, 43.64.Pg [RDF]

INTRODUCTION

When a direct sound and a reflection strike the two ears within a short period of time, they are fused into a single auditory percept, and localization cues from the leading sound dominate those from the lagging sound. We investigated the neural basis of this “precedence effect” by recording responses of neurons in several structures of the ascending auditory system to click-pair stimuli, which are commonly used in behavioral studies. The structures studied were the auditory nerve, anteroventral cochlear nucleus (AVCN), superior olivary complex (SOC), inferior colliculus (IC), and auditory cortex.

Behaviorally, a sequence of perceptual events occurs as the interval between two identical sounds from different locations is increased (see Blauert, 1982; Zurek, 1987; and Litovsky *et al.*, 1999 for reviews). The first sound, simulat-

ing the direct wavefront, is called the “conditioner” and the delayed sound, simulating a reflection or echo, is called the “probe.” The sounds can be presented in a sound field [Fig. 1(A)] or under headphones [Fig. 1(B)]. In a sound field, speakers are generally placed in front of the listener and on either side of the midline, and the sound from the probe speaker is systematically delayed. Under headphones, interaural time delays (ITDs) are commonly used to provide different lateralities of the conditioner and probe. The general effect of varying the “conditioner–probe interval” (CPI) is shown in Fig. 1(C). When the CPI is less than ~ 1 ms, the listener hears a single, fused auditory event and judges its location to be between the conditioner and probe. This is called the window of “summing localization” and the perception follows the physical cues present due to the combination of the two sources (Blauert, 1982). When the CPI is increased beyond 1 ms, the listener continues to hear a fused auditory perception; however, the location is judged to be that of the conditioner. This is the start of the precedence

^{a)}Electronic mail: dcf@neuron.uhc.edu

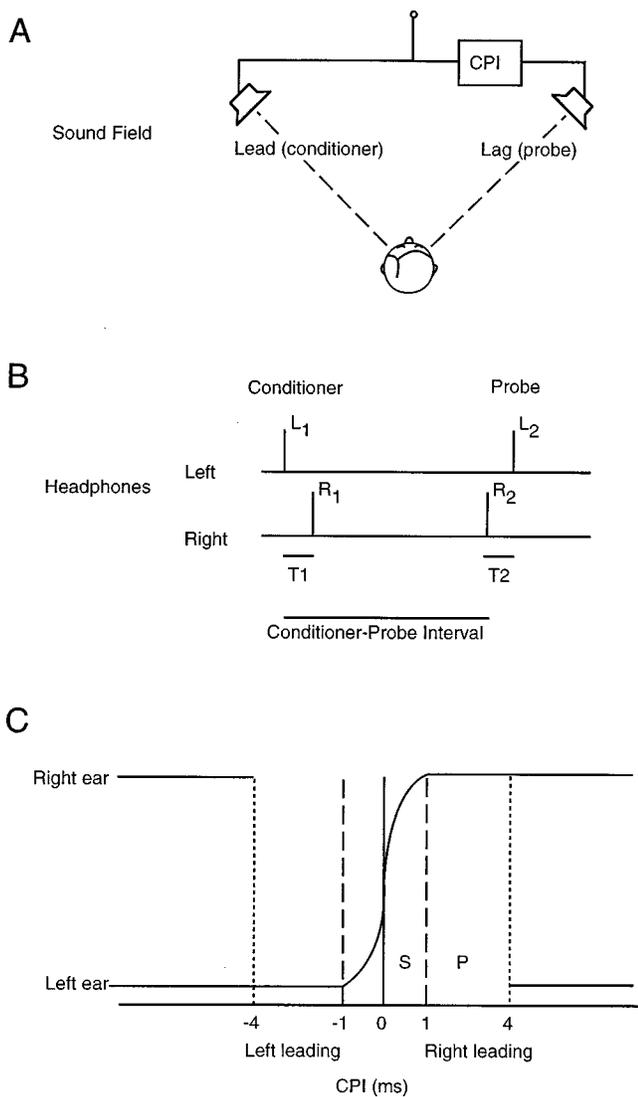


FIG. 1. Basic features of the precedence effect. (A) Typical arrangement of speakers in the sound field. Sound from one speaker lags the other by a variable amount, called the conditioner-probe interval (CPI). (B) Typical stimulus properties using headphones. Each vertical line represents a click to one ear. The ITDs for the conditioner and probe are similar, but opposite in sign. (C) The perceived position of the sound source (solid lines) for different CPIs. From 0 to about 1 ms CPI the two sounds are fused and move from the center towards the position of the conditioner (summing localization window). From about 1 to 4 ms CPI the fused percept is heard at the position of the conditioner (precedence window). At echo threshold and beyond, two separate sound sources are perceived at the locations of the conditioner and the probe.

window. The end of the precedence window occurs at “echo threshold,” or the CPI at which the probe is heard as a separate sound in its own location. For clicks, echo threshold occurs at CPIs of ~2–10 ms, with headphone studies generally yielding shorter values (~2–4 ms) than free-field studies. Echo threshold can be much longer for stimuli of longer duration, such as noise bursts or speech (Haas, 1951; Blauert, 1982). For a considerable period beyond echo threshold, the probe continues to have less salience than the conditioner, e.g., it is heard more faintly than the conditioner (Blauert, 1982). Although the vast majority of behavioral results regarding this sequence of events has been obtained from humans, experiments in rats (Kelly, 1974), cats (Cran-

ford and Oberholzer, 1976; Cranford, 1982; Populin and Yin, 1998), and even crickets (Wytenbach and Hoy, 1993) indicate that similar phenomena occur in other animals.

The neural basis for the precedence effect has recently received considerable attention. For CPIs within the summing localization window, neurons sensitive to sound-source location in owls and cats have been shown to respond to “phantom images” in accord with the physical cues present during this window (Yin, 1994; Keller and Takahashi, 1996a). For larger CPIs, neurons from the nerve to the IC show suppressed responses to a probe in the presence of a conditioner. The period of suppression varies considerably among structures and among neurons within a structure (Yin, 1994; Fitzpatrick *et al.*, 1995; Keller and Takahashi, 1996b; Parham *et al.*, 1996; Wickesberg, 1996; Litovsky and Yin, 1998a, 1998b; Wickesberg and Stevens, 1998; Parham *et al.*, 1998).

Previous studies have considered responses of neurons to the paired-click paradigm in only one or two auditory nuclei at a time, and only up to the level of the IC. Furthermore, some studies employed an anesthetized preparation while others did not use anesthesia. For this report, we synthesize information from subcortical structures and extend the observations to the cortex, all using unanesthetized animals. Because of this approach, the early sections are to an extent a review. However, by comparing the data from each structure directly, we can describe transformations that occur as information ascends the system.

I. METHODS

A. Experimental animals

Two preparations were used. The auditory nerve and AVCN were studied in decerebrate, unanesthetized cats. The SOC, IC, and auditory cortex were examined in unanesthetized rabbits. Details of the surgical and recording methods can be found in previous reports (see Parham *et al.*, 1996, 1998 for the decerebrate cat, and Fitzpatrick *et al.*, 1995 for the unanesthetized rabbit).

B. Acoustic stimuli

Tones, binaural-beat stimuli, and clicks were used. Tones were used to assess characteristic frequencies and best frequency. For the nerve and AVCN, characteristic frequency (cf) was defined as the frequency of minimum threshold using 200-ms-long tones. In the SOC, IC, and auditory cortex, best frequency was determined from an iso-intensity response curve to 75-ms-long tones at a suprathreshold level, usually 40–70 dB SPL.

Binaural-beat stimuli or clicks were used to assess sensitivity to ITDs. The binaural-beat stimuli were created using low-frequency tones (<~2 kHz) or high-frequency tones with sinusoidally amplitude-modulated envelopes (SAM tones). The tones or SAM tones differed at the two ears by 1 Hz in frequency or modulation frequency, respectively. When using clicks to assess sensitivity to ITDs, the ITDs of single clicks was varied between ± 2 ms (positive ITDs indicate the leading stimulus was at the ipsilateral ear). If a neuron was sensitive to ITDs, the best and worst ITDs were

determined (Fitzpatrick *et al.*, 1995). The best ITD was that which elicited the largest response, and the worst ITD was that which elicited the smallest response.

Click-pair stimuli were used to assess the recovery of a probe click to a preceding conditioner. The conditioner and probe clicks had the same amplitude. The clicks were created by a square pulse to the earphones. The pulses were 30- μ s-long for the cat auditory nerve and AVCN, and 100–200- μ s-long for the rabbit SOC, IC, and auditory cortex. The acoustic signals measured in the ear canals lasted 3–4 ms (see Fig. 1 of Parham *et al.*, 1998). When a neuron that was sensitive to ITDs was tested with binaural click pairs, the conditioner and probe initially had the best ITD of the neuron. This configuration is called the “best/best” condition. In some neurons, the effect of a conditioner with the worst ITD was then tested on a probe with the best ITD. This configuration is called the “worst/best” condition.

C. Data analysis

The response to the probe was expressed as the “percent recovery,” defined as the response to the probe relative to that of an identical stimulus presented in isolation. At short CPIs, the response to the conditioner and probe often overlapped, so the response due to the probe had to be corrected for the overlap due to the conditioner. For neurons in the auditory nerve and AVCN, this correction was made according to the method described in Parham *et al.*, 1996. This method uses closely defined time windows based on the duration of the response to a single click to determine the number of spikes to exclude from the response to the probe. This procedure worked well at the level of the nerve and AVCN, but in higher structures there was sometimes variability in the latency and duration of the response to the probe as a function of CPI which precluded the use of such strictly defined windows. In these higher structures, when overlap occurred the response to the probe was determined by summing all of the spikes in the combined conditioner and probe response and subtracting the average response to the conditioner alone. In most neurons, the difference between the two procedures was minimal.

In each structure, the period of the reduction in spontaneous activity following a monaural or binaural click was determined for neurons with spontaneous rates of >5 spikes/s. For ITD-sensitive neurons, the binaural click was set at the best ITD of each neuron. For other neurons studied with binaural clicks the ITD was zero. The spontaneous rate was measured over 5–20 s. A horizontal line at this rate was placed on a smoothed post-stimulus time (PST) histogram (3- or 5-point moving average using 0.5- or 1.0-ms bins) of the response to the single monaural or binaural click, and the period from the end of the excitatory response to the point where the activity returned to the spontaneous rate was determined.

II. RESULTS

This study is based on the responses of 310 neurons to click-pair stimuli. Of these, 42 were in the auditory nerve, 121 in the AVCN, 27 in the SOC, 67 in the IC, and 53 in the

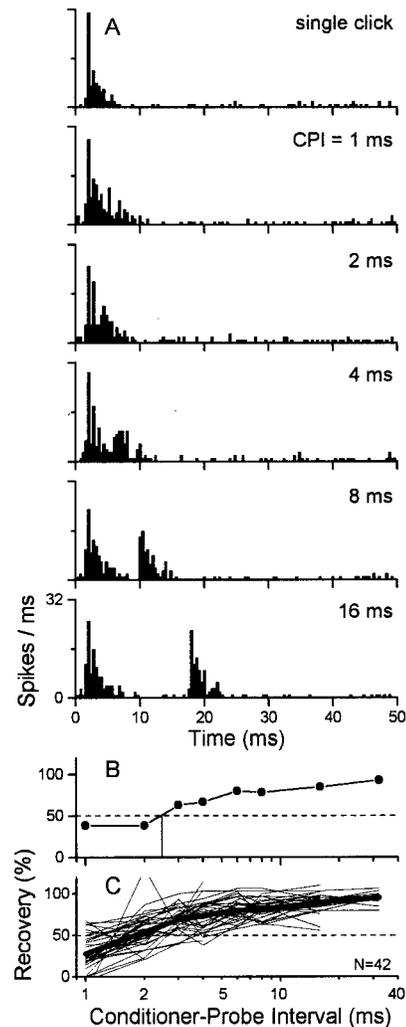


FIG. 2. The effect of varying the CPI in paired clicks on the responses of auditory-nerve fibers. Preparation is the decerebrate cat. (A)–(C) Poststimulus time histograms for one fiber (CF=3.0 kHz, SR=40 spikes/s) at different CPIs (65-dB SPL clicks). (H) Recovery function for the same fiber. (I) Recovery function for all nerve fibers recorded. The heavy line is the average (population) recovery. Data from Parham *et al.*, 1996.

auditory cortex. Most neurons in each structure showed a suppressed response to the probe in the presence of a conditioner. In the following, we will describe the effects observed at each level.

A. Auditory nerve, anteroventral cochlear nucleus, and superior olivary complex

The responses of a typical nerve fiber are shown in Fig. 2. The response to a single click is shown in Fig. 2(A), top panel. The remaining histograms in Fig. 2(A) show responses to click pairs with different CPIs. The response to the probe was smallest at short CPIs, and systematically increased with CPI. By 16 ms CPI the responses to the conditioner and probe were nearly identical. The response to the probe was quantified as a “recovery function” [Fig. 2(B)]. For this neuron, a 50% recovery was reached at a CPI of 2.5 ms, and full recovery by ~ 40 ms.

Recovery functions for the full sample of nerve fibers are shown in Fig. 2(C). At a CPI of 4 ms, i.e., near echo threshold in humans for clicks presented over headphones,

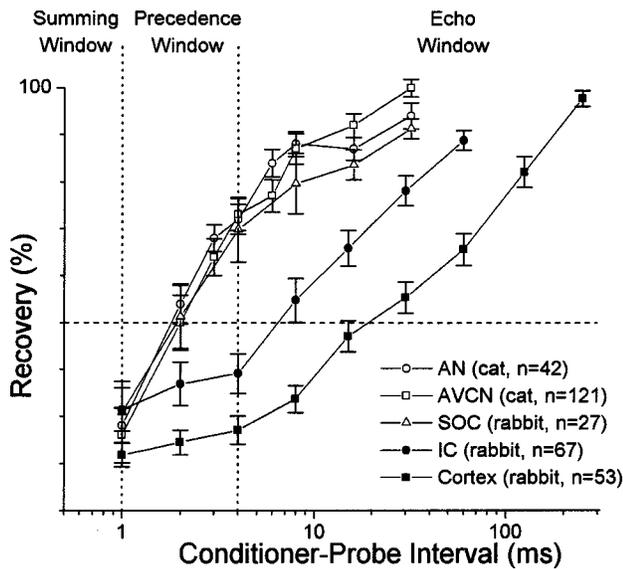


FIG. 3. Population recovery functions for each level compared. Those from the auditory nerve, AVCN, and SOC were all similar and short, with 50% recovery times of ~ 2 ms. By echo threshold (4 ms CPI) the recovery was quite high, near 70%. For the IC, the recovery was much slower, with 50% recovery at ~ 7 ms and $\sim 40\%$ recovery at echo threshold. For the auditory cortex, the recovery was slower yet, with 50% recovery at ~ 20 ms and 25% recovery at echo threshold.

almost all neurons had recovered at least to the 50% level. This is also indicated by the ‘‘population’’ or average, recovery time (heavy line) that showed a 50% level of recovery at 1.8 ms CPI.

The AVCN recordings were from a heterogeneous population of neurons, i.e., those with primary-like responses (putative bushy cells), chopper responses (putative stellate cells), or onset responses (heterogeneous cell types). The SOC neurons studied were also heterogeneous. Most were monaural, although a few were weakly binaural. None was sensitive to ITDs. They were thus presumably located outside of the main nuclei of the SOC. They responded to tones with chopper or off discharge patterns. Despite this heterogeneity of response types, the population recovery functions of AVCN and SOC neurons (Fig. 3, open squares and triangles, respectively) were strikingly similar to the nerve (Fig. 3, open circles). In each structure, 50% recovery occurred at ~ 2 ms CPI, and by echo threshold (4 ms CPI) each population had a high level of recovery ($\sim 70\%$). Full recovery occurred by 30–50 ms CPI.

A difference between the nerve and higher structures was an increase in the diversity of recovery function shapes. In the nerve, functions were monotonic. At higher levels, two additional shapes were noted. Examples of these shapes are shown in Fig. 4 for two neurons from the AVCN. One neuron showed a large response to the probe at all CPIs, starting from 1 ms [Fig. 4(A)–(F)]. Thus, the recovery function was essentially flat [Fig. 4(M), solid circles]. A second neuron [Fig. 4(G)–(L)] showed considerable response to the probe at short CPIs, followed by a decline before the response increased again at longer CPIs. The decline caused a trough-shaped recovery function [Fig. 4(M), open circles].

For both neurons illustrated, the large response to the

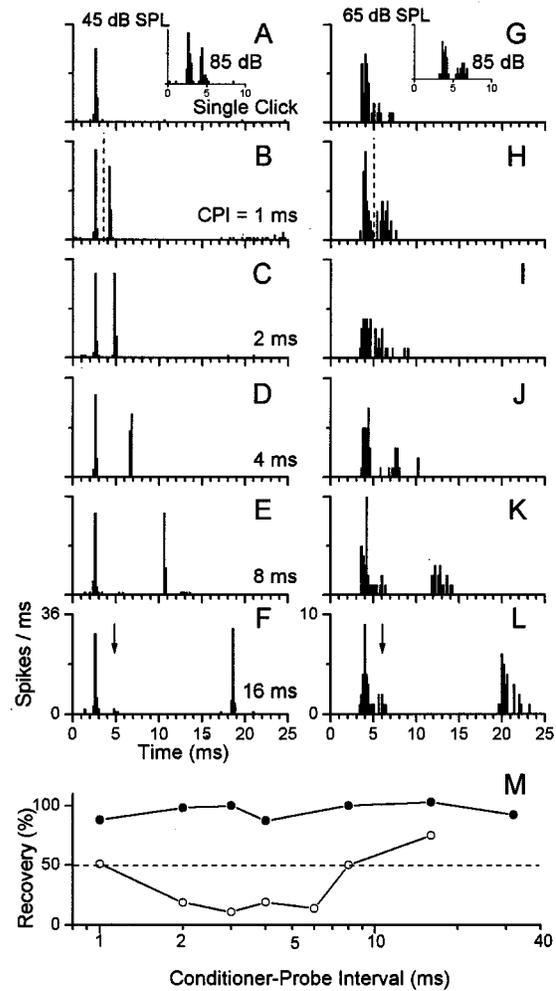


FIG. 4. Examples of AVCN neurons with recovery functions unlike typical nerve fibers. (A)–(F) A neuron (CF=7.81 kHz, SR=42 spikes/s, chopper response to tones) where the response to the probe was strong at all CPIs, even 1 ms (B). Taken at 45 dB SPL. Inset in (A) is the single click response at 85 dB SPL showing a chopping pattern, which is also evident at 45 dB SPL although the second peak is much smaller [arrow in (F)]. (B) and (C) (open circles). (G)–(L) A neuron (CF=2.84 kHz, SR=0 spikes/s, chopper response to tones) where the response to the probe was strong at 1 ms CPI (H) but then declined at longer CPIs. As in the previous example, this neuron showed a chopper response at 85 dB SPL [inset in (G)] that was smaller at the lower SPL [arrow in (L)]. Data from Parham *et al.*, 1998. (M) Recovery functions for these two neurons; that from (A)–(F) is nearly flat (filled circles) while that from (G)–(L) is trough-shaped (unfilled circles).

probe at short CPIs may be related to the temporal structure of the response to a single click. At high SPLs, both neurons showed a chopping response to a single click [insets in Fig. 4(A) and (G)]. At lower sound-pressure levels, the chopping to a single click was less evident, but sometimes a few spikes at the position of the second peak occurred [arrows in Fig. 4(F) and (L)]. For both neurons, the response to the probe at 1 ms CPI [Fig. 4(B) and (H)] was not 1 ms from the peak of the conditioner response (dashed lines). Instead, the increased response in the presence of the probe occurred at the position of the second chopping peak. Thus, the presence of the probe at short CPIs increased the probability of chopping in these neurons, suggesting an interaction with the response to the conditioner.

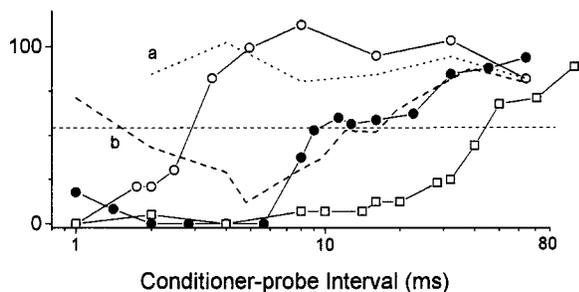


FIG. 5. Examples of recovery functions seen in the IC. Curves (a)–(c) have shapes typical of most neurons and were chosen to span the range of recovery times seen. Neuron *a* had a flat recovery function and neuron *b* had a trough-shaped recovery function.

B. Inferior colliculus

The recovery times of many neurons in the IC were lengthened compared to lower structures. Figure 5 shows examples of recovery functions in the IC. The three examples with symbols span the range of recovery times seen. The neuron with open circles had a relatively short recovery time (2.9 ms to 50% recovery), similar to those typical of most neurons in lower structures. The neuron with closed circles had a longer recovery time (8.8 ms for 50% recovery), which was near the median for the IC. The neuron with open squares had among the longest recovery times seen in the IC (42.3 ms to 50% recovery). In addition to these typical shapes, some neurons in the IC also had flat (neuron *a*) or trough-shaped (neuron *b*) recovery functions, as was also seen in the AVCN and SOC, but not in the auditory nerve. Of the 67 IC neurons, 43(64%) had recovery functions like those in the AN, although some had expanded recovery times. Nineteen neurons (28%) had trough-shaped recovery functions, based on the criteria that they had a recovery >25% at short CPIs (1–2 ms) that declined by at least 1/2 at intermediate CPIs, before increasing again toward full recovery at longer CPIs. Three neurons (4%) had flat recovery functions, based on the criteria that the recovery was always >50%.

The lengthening of recovery times in the IC is evident in the population recovery function (Fig. 3). The population of IC neurons reached 50% recovery at 6.9 ms, which is about three times longer than for lower structures. Complete recovery did not occur until ~100 ms CPI, compared to 30–50 ms in lower structures. The recovery of the IC population during the precedence window was fairly low, with most recovery occurring after echo threshold. The inflection point in the function for the IC at 4 ms CPI was due primarily to neurons with trough-shaped recovery functions, which raised the response to the probe at short CPIs.

As in the AVCN and SOC, a heterogeneous population of neurons was studied in the IC. One large group was sensitive to ITDs in the fine structure of low-frequency sounds (55/67 neurons). These neurons would be expected to mediate the ability to determine the ITD, and by extension the azimuthal location, of conditioners and probes. A smaller group (12/67) was not sensitive to ITDs. Both groups had similar population recovery functions [Fig. 6(A)] that were much longer than for the SOC (dotted line, taken as repre-

sentative of all lower structures). The population recovery functions were also similar among ITD-sensitive neurons that had low best frequencies (<2.5 kHz), with a small preference for higher recoveries at short CPIs in the low best-frequency population [Fig. 6(B)].

In a subset of ITD-sensitive neurons, we compared the effects of conditioners with the worst or best ITD on the responses of probes that had the best ITD. The recovery functions to the different conditioners varied among neurons, with either the best or the worst ITD conditioner more suppressive, or with both having equal suppression (see Fitzpatrick *et al.*, 1995). In the population recovery functions [Fig. 6(C)] the worst ITD was on average slightly more suppressive at short CPIs (1–4 ms), but at longer CPIs the two functions were nearly identical.

The effect of stimulus level was highly variable across IC neurons. In some, higher levels of conditioners and probes evoked stronger suppression; in others there was little effect over a wide range of stimulus levels (up to 40 dB), and in still others lower levels were more suppressive. Among all neurons [Fig. 6(D)] there was a trend for higher stimulus levels to be more suppressive. This result parallels that seen in the auditory nerve and AVCN (Parham *et al.*, 1996, 1998).

Taken together, there were only small differences among different IC populations or the same population studied with different stimuli, and each showed an increased recovery time compared to lower levels. Thus, the increases in the IC appeared to be general rather than specific to responses that might be expected to mediate the precedence effect, such as those sensitive to ITDs.

C. Auditory cortex

A further expansion in recovery times occurred between the IC and the primary auditory cortex (see McMullen and Glaser, 1982, for a description of the auditory cortex in the rabbit). Three examples are shown in Fig. 7(A)–(C) (note different scales on abscissa). As in the IC, some neurons in the cortex had short recovery times, comparable to those of most neurons in the lower brainstem. The neuron in Fig. 7(A) responded weakly to the probe at 1 ms CPI but at 2 ms was recovered to nearly the 50% level. The neuron in Fig. 7(B) had a recovery time near the median for the population. There was little or no response to the probe until the CPI was 16 ms, and at 32 ms the response was over 50% recovered. The longest recovery times in the auditory cortex were much longer than in the IC. The neuron in Fig. 7(C) had no response to the probe until the CPI was greater than 64 ms, and at 192 ms CPI was still not fully recovered.

The recovery functions for the same three neurons are shown in Fig. 8(A) (open circles, closed circles, and open squares, respectively). The recovery functions for two additional neurons are included to illustrate the continuous nature of recovery times among cortical neurons.

As in all lower structures except the nerve, some recovery functions were flat [Fig. 8(B), neurons *a* and *b*] and some were trough-shaped (neurons *c* and *d*). The troughs in cortical neurons sometimes extended to 100 ms CPI (e.g., neuron *d*), which is much greater than the ranges seen in the IC and in the AVCN, where 2–8 ms was typical. In the cortex, 32 of

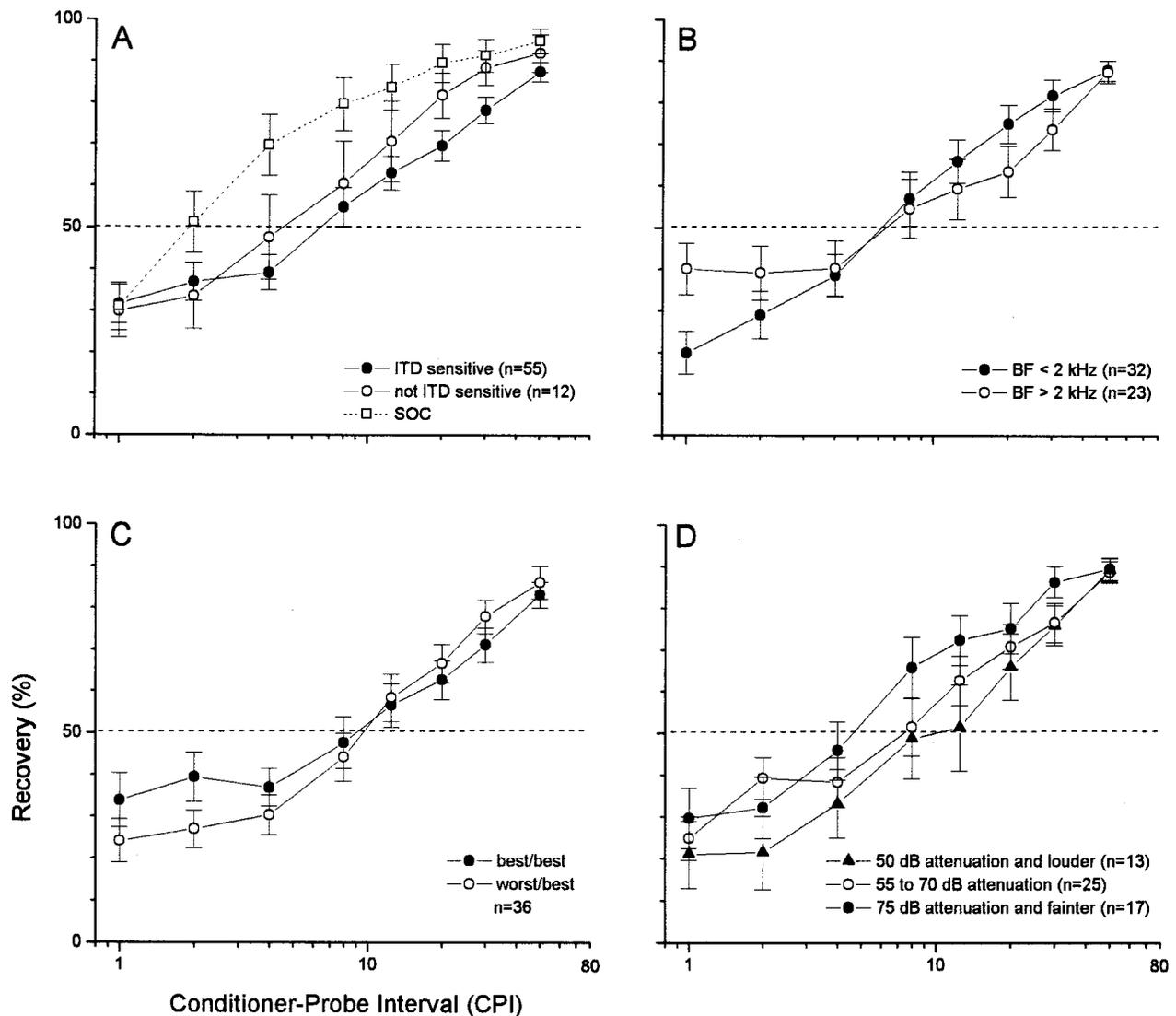


FIG. 6. Comparisons of population-recovery functions for various populations of IC neurons and stimulus conditions. (A) Functions from neurons that were ITD-sensitive (closed circles) compared with ITD-insensitive neurons (open circles). Neurons that were ITD-sensitive were tested with conditioners and probes that had the best ITD of the neuron. The population-recovery function for the SOC (dotted lines), taken as representative of each lower structure (see Fig. 3) is shown for comparison. (B) Functions from neurons with best frequencies below (closed circles) and above 2 kHz (open circles). (C) Functions where the conditioner had either the neuron's best ITD (closed circles) or worst ITD (open circles). The probe always had the neuron's best ITD. (D) Functions taken at different stimulus levels. Stimulus level is expressed in attenuation (*re* 127 dB maximum attenuation).

53 neurons (60%) had monotone recovery functions, although many had expanded recovery times compared to the nerve. Nine neurons (17%) had trough-shaped recovery functions and four neurons (8%) had flat recovery functions, based on the criteria described for the IC neurons.

Unlike neurons seen at lower levels, some cortical neurons were tuned to a particular range of CPIs. Three examples are shown in Fig. 8(C). Each had a different "best" CPI (~2 ms for neuron *a*, 16 ms for neuron *b*, and 40 ms for neuron *c*). Among the eight neurons (15%) that showed such tuning, the best CPIs ranged from 2–70 ms.

A possible basis for the tuned type of recovery function is shown in Fig. 9. In this neuron [neuron *c* of Fig. 8(C)], the "conditioner alone" histogram [Fig. 9(A)] shows the response to 300 repetitions, with the response to the conditioner truncated to emphasize small responses in the period

following the conditioner response. A small volley of excitation occurred about 40 ms after the stimulus (arrow). When the response to the probe overlapped this volley of excitation due to the conditioner [at CPIs of 24 and 32 ms, in Fig. 9(C) and (D), respectively], the response was large. At smaller [16 ms CPI, Fig. 9(B)] and larger CPIs [48 and 64 ms, Fig. 9(E) and (F)], the response to the probe was smaller. The response did not increase again until much longer CPIs.

The distribution of 50% recovery times for cortical neurons is shown in Fig. 8(D). Neurons tuned to CPI were excluded. The distribution is continuous to >128 ms CPI. The median value for 50% recovery was 21.9 ms.

The increase in recovery times in the cortex compared to lower structures is evident in the population recovery function (Fig. 3, closed squares). The CPI for 50% recovery was 19.3 ms, for a threefold increase over the IC. However, at

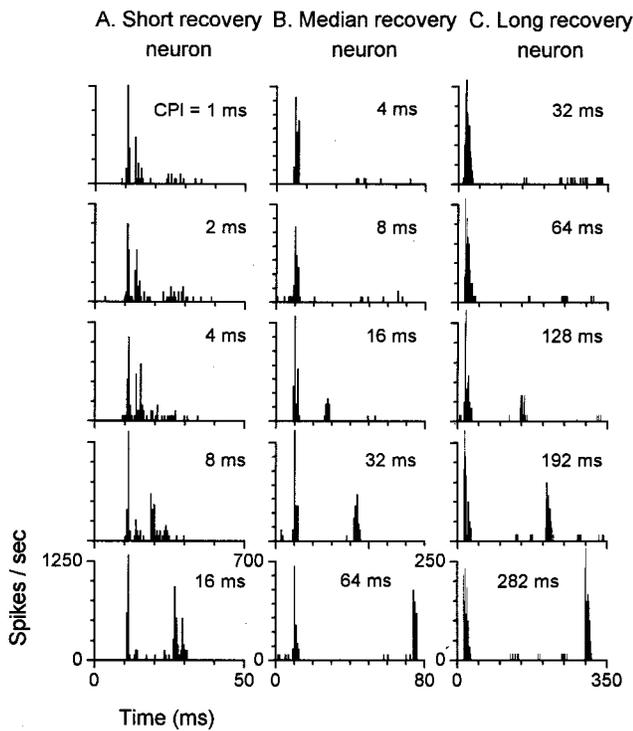


FIG. 7. Examples of responses from the auditory cortex. (A) A neuron (CF=2.2 kHz, SR=1.1 spikes/s) with short recovery times, similar to those found in all lower structures. (B) A neuron (CF=6.2 kHz, SR=7.3 spikes/s) whose recovery times were near the median for the population (about 20 ms for 50% recovery). (C) A neuron (CF=6.2 kHz, SR=0.5 spikes/s) with longer recovery times than lower structures.

echo threshold the response to the probe was still substantial, at ~25%. The time for nearly full recovery extended to ~300 ms.

Different populations of neurons in the auditory cortex are compared in Fig. 10, as was previously done for the IC (see Fig. 6). As in the IC, there was little difference in overall recovery for neurons sensitive to ITDs and those that were not [Fig. 10(A)]. The ITD-sensitive neurons included those with tuning to ITDs in low-frequency sounds ($n = 13$) or in the low-frequency envelopes of high-frequency sounds ($n = 19$). Two neurons were sensitive to ITDs in both types of signals. The population recovery functions were also comparable for neurons with low- or high-frequency tuning [Fig. 10(B)].

The population recovery functions for neurons sensitive to ITDs in low-frequency sounds tested in the best/best compared to the worst/best condition were less similar [Fig. 10(C)] than was the case in the IC [Fig. 6(C)], but the difference was not systematic. At CPIs <8 ms, the conditioner with the worst ITD was more suppressive. At CPIs >8 ms, the conditioner with the best ITD was more suppressive. Of the nine neurons tested in both conditions, four had longer 50% recovery times to the best/best condition, three to the worst/best, and two were similar (within 20%). Thus, as in the IC, there was no clear trend for the best/best or worst/best condition to be more suppressive.

From the nerve to the IC, lower conditioner and probe stimulus levels generally resulted in shorter recovery times. This trend was maintained in the auditory cortex up to CPIs

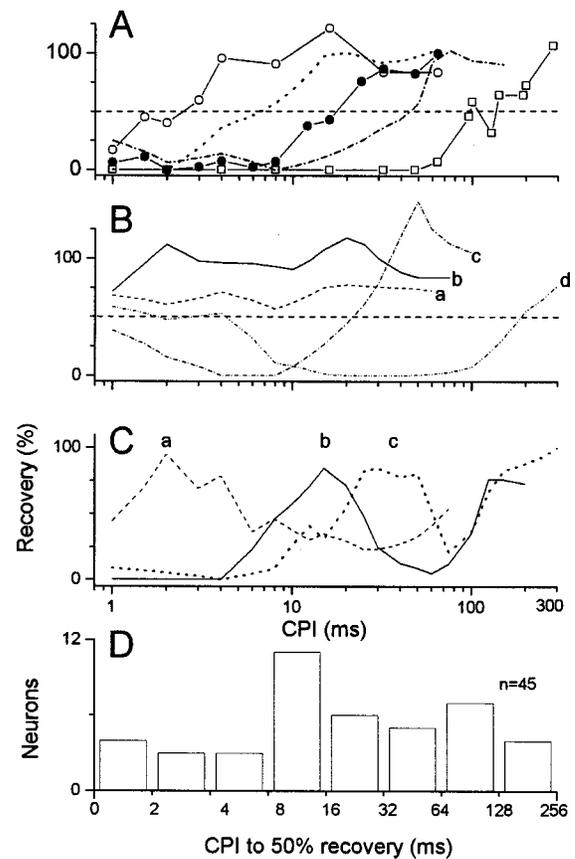


FIG. 8. Examples of recovery functions from the auditory cortex. (A) The three functions with symbols are from the neurons shown in Fig. 7. The remaining two illustrate the continuous distribution of recovery times. (B) Neurons with flat (a,b) and trough-shaped (c,d) recovery functions. (C) Neurons tuned to different CPIs. This type of recovery function was not seen in lower structures. (D) The distribution of 50% recovery times. Neurons tuned to CPI (8/53) were excluded. For some neurons we obtained trials with different stimuli, such as best ITD and worst ITD or different SPLs. All trials were included.

of 8 ms [Fig. 10(D)], but at longer CPIs a trend was not apparent. As in the IC, individual cortical neurons showed a range of effects of stimulus level on recovery, with a higher stimulus level in some neurons being more suppressive, in some equally suppressive, and in some less suppressive than a lower stimulus level.

An interesting feature of the precedence effect is that on the first few trials the conditioner has relatively little ability to mask the location of the probe, but over the course of about ten trials the precedence effect builds up to a maximum (Thurlow and Parks, 1961; Clifton and Freyman, 1989; Freyman *et al.*, 1991). We examined raster displays of the responses to individual trials of conditioners and probes, but saw no evidence for this “build-up” of suppression either in the cortex or in lower structures.

D. Suppression of spontaneous rates

In many neurons with significant spontaneous activity, the response to a single click was followed by a period of suppression of the spontaneous activity. The period of this suppression was determined to assess its correlation with the suppression of responses to probes. As was the case with recovery times, the periods of suppression of spontaneous

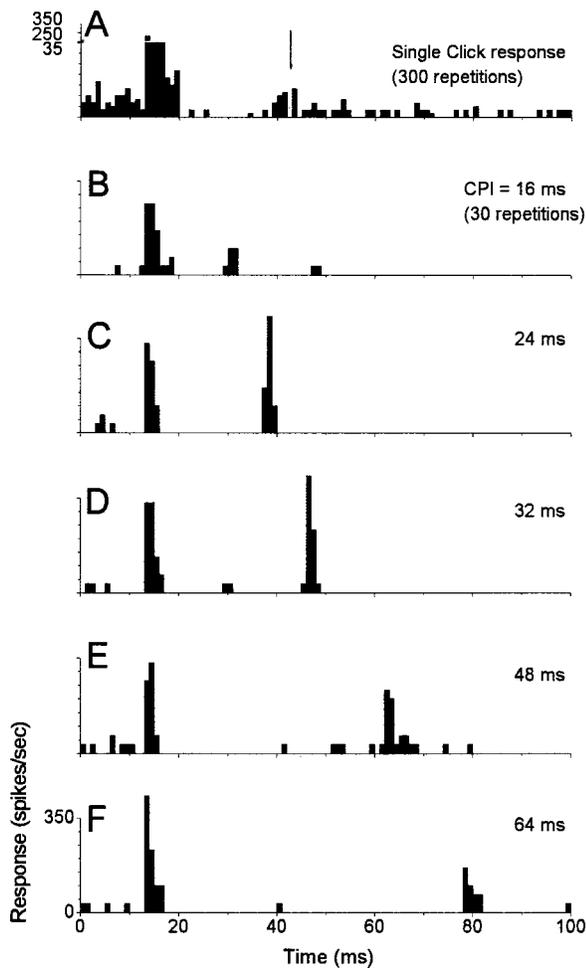


FIG. 9. A mechanism for tuning to CPI. (A) Response to 300 repetitions of a single click. The response to the click has been clipped to highlight a small increase in excitability that occurred at about 40 ms poststimulus (arrow). (B)–(F) Responses to varying CPIs. As the probe responses enter the region of increased excitability the recovery is high (C) and (D), while for shorter and longer CPIs the recovery is lower [(B), (E), and (F)].

activity were short and similar in the auditory nerve, AVCN, and SOC, increased at the level of the IC, and increased again in the cortex [Fig. 11(A)]. Thus, there were similar trends in the two measures of suppressive period following a single click stimulus. Among individual neurons, however, there was little correspondence between the two measures ($r=0.33$ in IC and 0.30 in cortex).

Another factor that varied across brain levels was the spontaneous rate itself [Fig. 11(B)]. In the nerve, AVCN, and SOC the range of spontaneous rates extended to about 100 spikes/s, and the mean values and standard deviations were similar. The rates in the IC and auditory cortex were similar to each other, but lower than the previous levels.

III. DISCUSSION

Our major observation was an increase in recovery times to paired clicks at successively higher stations along the auditory system. One increase occurred between the IC and lower structures, and a second occurred between the IC and the auditory cortex. In the following, we will first compare results in the two species used in this study, the cat and the

rabbit. We will then discuss the increases in recovery times in relation to behaviors surrounding the precedence effect, and consider possible neural mechanisms. Finally, we will discuss other transformations observed as information ascended.

A. Comparison of recovery times to paired clicks in cats and rabbits

For this study, the auditory nerve and AVCN were studied in decerebrate cats, while the SOC, IC, and auditory cortex were studied in unanesthetized rabbits. Both preparations were unanesthetized, an important point of similarity. Neurons in the auditory nerve and AVCN of the cat and in the SOC of the rabbit had recovery times that were quite short, with median 50% recovery at about 2 ms and full recovery of the populations by 40–50 ms. Similarly short recovery times are also reported in the nerve and AVCN of the ketamine-anesthetized chinchilla (Wickesberg and Stevens, 1998; Wickesberg, 1996). Thus, the species and anesthetic state appear to have little effect on recovery times in the auditory nerve and AVCN, and the SOC neurons studied in the unanesthetized rabbit had recovery times that were remarkably similar to these lower structures.

At the level of the IC, species differences and/or anesthetic state appear to be more important. That is, recovery times in the anesthetized cat are on average much larger than in the unanesthetized rabbit (median 50% recovery time of ~20 ms in cats compared to ~6 ms in rabbits—Yin, 1994; Litovsky and Yin, 1998a). While the species difference is a possible cause, barbiturate anesthesia can be expected to increase recovery times, because it known to potentiate inhibition at synapses containing gamma-aminobutyric acid (GABA) (Barker and Ransome, 1978; Richter and Holtman, 1982). The proportion of inhibition due to GABA increases in the IC and cortex compared to lower structures (e.g., Winer *et al.*, 1995). A preliminary report (Reale *et al.*, 1995) indicates that recovery times in the cortex of the anesthetized cat are large compared to the IC of the same preparation (Yin, 1994; Litovsky and Yin, 1998a), and are also larger than in the cortex of the unanesthetized rabbit. Thus, the general trend of an increase in recovery times between the IC and cortex is supported with and without anesthesia, but the quantitative results are quite different.

B. Comparison of the neural responses and behavioral aspects of the precedence effect

Under headphones, echo threshold occurs from ~1–4 ms in different studies, and in the sound field extends to only 5–10 ms (Blauert, 1982; Litovsky *et al.*, 1999). These brief periods can be accounted for by responses in the nerve, AVCN, and SOC, where the response to the probe increased steadily through the precedence window and each population reached a high level of recovery by echo threshold (e.g., about 75% recovery at 4 ms CPI). Thus, the duration of the precedence effect may be based on an ability to detect changes in the most sensitive structures in the brain. Alternatively, despite the overall increases in recovery times in the IC and cortex, many neurons in these structures show a time course of recovery similar to those in lower structures.

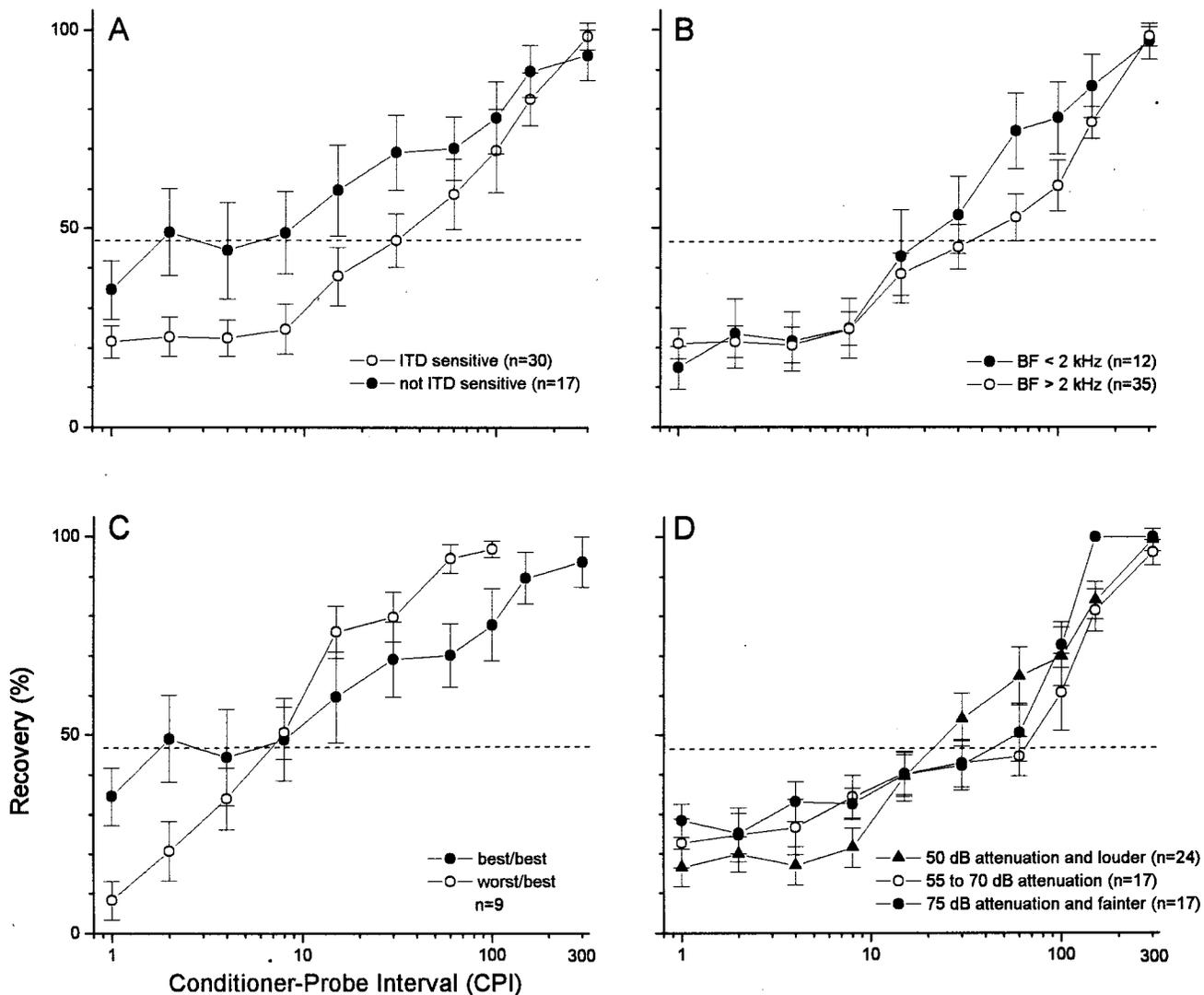


FIG. 10. Comparisons of various population-recovery functions for different populations of cortical neurons of different stimulus conditions. (A) Neurons sensitive to ITDs and those that are not. Responses of ITD-sensitive neurons were to conditioners and probes that had the best ITD, and included neurons tuned to ITDs in low-frequency sounds or the envelopes of high-frequency sounds. (B) Functions from neurons with best frequencies below (closed circles) and above 2 kHz (open circles). (C) Functions where the conditioner had either the neuron's best ITD (closed circles) or worst ITD (open circles). The probe always had the neuron's best ITD. (D) Functions taken at different stimulus levels. Stimulus level is expressed in attenuation (*re* 127 dB maximum attenuation).

It may therefore be that the precedence effect relies on the ability to detect activity in the most sensitive neurons in higher structures of the auditory pathway.

The relatively large degree of responsiveness at short CPIs in each structure is also consistent with the result that information from the probe is not entirely masked during the precedence window. For example, during the precedence window the spatial extent of the fused percept spreads toward the location of the probe (Blauert, 1982; Yost and Soderquist, 1984; Lindemann, 1986; Perrott *et al.*, 1987). This spread is presumably due to the steady increase in firing of neurons tuned to the location of the probe, such that the range of activity across ITDs has increased.

The increases in recovery times in the IC and particularly in the cortex appear to be a phenomenon not directly related to the precedence effect, since the suppression continues well beyond echo threshold. Echo threshold is generally considered to be the end of the precedence effect be-

cause at this point the two sounds are heard separately and near their correct locations. We call the extended period of suppression in neurons the "echo window" (Fig. 3). The long echo windows in the IC and cortex do have behavioral correlates, which include a reduction in the discriminability of the ITD in the probe (Shinn-Cunningham *et al.*, 1993; Tollin and Henning, 1998), and a lower perceived loudness of the probe compared to the conditioner (Blauert, 1983) for periods well beyond echo threshold. Based on the cortical response, we would predict that some dominance of information in the conditioner compared to the probe should extend for up to ~300 ms CPI.

By definition, the precedence effect is restricted to location information. The effect shows an important difference from monaural masking, which is that backward masking of location information in the conditioner (e.g., the ITD) by the probe does not occur (Zurek, 1987). However, we and others have not found systematic differences in recovery times with

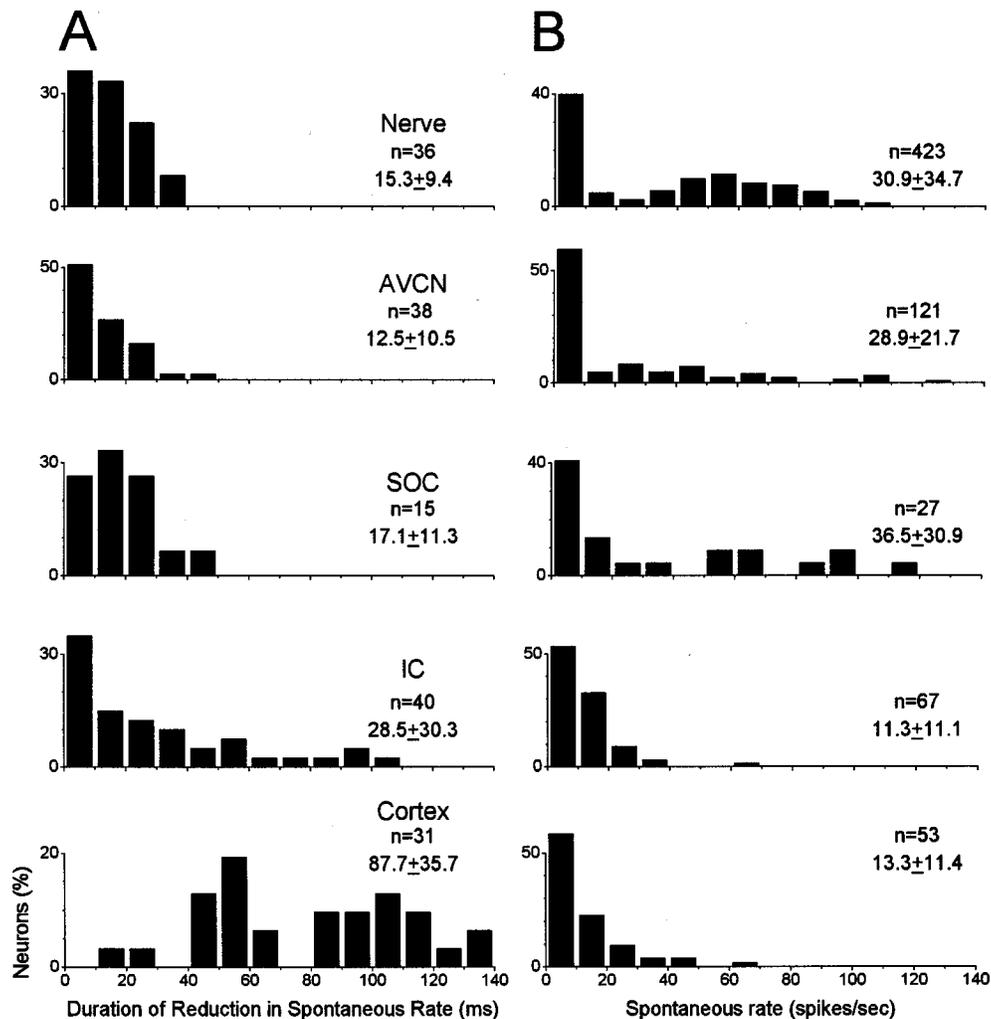


FIG. 11. (A) Distribution of the duration of reduction in spontaneous activity following a single click from each structure (see Methods). The suppressive periods were similar and short in the auditory nerve, AVCN, and SOC, increased in the IC, and increased again in the cortex. These increases parallel the recovery times to paired clicks. (B) Distributions of spontaneous rate. Sources of the original data: nerve—Kim *et al.*, 1991; AVCN—Parham *et al.*, 1998; SOC, IC, and cortex—present paper.

monaural or binaural stimuli (Yin, 1994; Fitzpatrick *et al.*, 1995; Litovsky and Yin, 1998b) or between neurons sensitive to ITDs and those that are not (Fitzpatrick *et al.*, 1995, current study). It therefore has yet to be shown that the neural effect of conditioners on probes is in any way specific to location information.

In humans, some studies support a weak trend for greater suppression of location information about the probe as conditioner locations on ITDs depart from that of the probe (Litovsky and Macmillan, 1994; Shinn-Cunningham *et al.*, 1993). For CPIs within the precedence window in the IC and cortex of the unanesthetized rabbit, there was a slight trend for a conditioner with the worst ITD to be more suppressive than one with the best ITD, in accord with the behavioral results. However, this trend was not followed in all neurons or at larger CPIs. In contrast to the unanesthetized rabbit, in the IC of anesthetized cats there was a strong preference for the best/best condition to be more suppressive than the worst/best condition (up to 89% of neurons studied—Yin, 1994; Litovsky and Yin, 1998b). It is not clear why anesthesia might affect one condition more

than the other, so the reason for this discrepancy is still unknown.

C. Inhibition and recovery times

Inhibition evoked by the conditioner is often invoked as an explanation for the loss of sensitivity to information in the echo (e.g., Zurek, 1980; Lindemann, 1986). In the following, we will assess the influence of inhibition at each level and discuss possible sources.

In the auditory nerve, the gradual recovery to a second click is similar to the relative refractory period of nerve fibers (Gaumond *et al.*, 1982; Parham *et al.*, 1996). This refractory period (~2 ms to 50% recovery, full recovery by ~40 ms) reflects the reduced probability of firing following each spike under a steady-state excitation, e.g., spontaneous activity or a continuous tone of high frequency (> ~5 kHz). Mechanisms for this “discharge history effect” are thought to be intrinsic to a nerve fiber, but hair-cell factors, such as depletion of neurotransmitter, could also potentially affect

the recovery. Thus, inhibition is an unlikely mechanism at the level of the nerve.

The population recovery times in the the AVCN and SOC are highly similar to the nerve. This result is surprising because AVCN and SOC neurons receive inhibitory inputs (Cant, 1992; Wenthold *et al.*, 1987; Helfert *et al.*, 1989). In both nuclei, the inhibition is primarily glycinergic and GABAergic (Cant, 1992; Wenthold *et al.*, 1987; Helfert *et al.*, 1989). In the AVCN of the ketamine-anesthetized chinchilla, a blockade of glycinergic and GABAergic inputs shortened the recovery times to clicks in most neurons, indicating an influence of inhibition in this structure (Backoff *et al.*, 1997). The similarity between the recovery times in the AN and AVCN in the decerebrate cat, however, shows that the duration of inhibitory effects to clicks must be of the same order as the relative refractory period of neurons. A relatively short duration of inhibition (<5 ms) is also suggested in the SOC by recordings in the lateral superior olivary nucleus (LSO) to clicks at either ear (Joris and Yin, 1995; Irvine *et al.*, 1998), and from *in vitro* studies of the medial superior olivary nucleus and LSO (Sanes, 1990; Wu and Kelley, 1992; Grothe and Sanes, 1993). The short recovery times in the periolivary neurons studied here are consistent with the short duration of inhibition seen in other parts of the SOC.

The increases in suppressive periods of the spontaneous activity and recovery times for populations of neurons in the IC are most likely influenced by GABAergic inputs, as opposed to glycinergic. Iontophoretic studies indicate that the effects of glycine are short-lasting (<5 ms) while the effects of GABA are more long-lasting (Yang and Pollak, 1994). Similarly, reversible chemical lesions of the contralateral dorsal nucleus of the lateral lemniscus (DNLL), which provides a major GABAergic input to the IC, eliminated a long-lasting (>5 ms) inhibition evoked by ipsilateral clicks, but left a shorter-lasting inhibition (Kidd and Kelly, 1996). Thus, glycinergic inhibition appears to have a relatively short duration, while GABAergic inhibition derived from the DNLL lasts longer, and could be expected to increase the duration of suppression of spontaneous activity and recovery times in IC neurons. If future studies show that manipulations of the DNLL affect recovery times in IC neurons, it would be a strong indication of a causal relationship between inhibition and neural recovery.

The further increase in suppressive periods at the cortical level is also likely to involve GABAergic pathways, since GABA is by far the predominant inhibitory neurotransmitter above the level of the IC. The auditory thalamus receives GABAergic projections from the IC (Winer *et al.*, 1996; Peruzzi *et al.*, 1997), the reticular nucleus of the thalamus, and, in some species, from intrinsic neurons. In the cortex, GABAergic inhibition is predominantly from intrinsic sources. Our preliminary observations are that at least some neurons in the auditory thalamus have longer recovery times than IC neurons, and comparable to those in the cortex. This suggests that at least some of the increase in recovery times between the IC and cortex occurs in the thalamus.

An important caveat is that not all aspects of the neural responses or behavioral features of the precedence effect can

be easily related to inhibition. Neurally, although a similar increase in the period of suppression of spontaneous rates and the recovery times did occur across structures, we did not see a strong correlation between these measures in individual neurons in the IC or cortex. It may be that the duration of spontaneous reduction does not closely reflect the strength of inhibition or the balance between excitation and inhibition in a particular neuron. For example, strong excitation from a probe may override weak inhibition from a conditioner, even through the inhibition is sufficient to reduce or eliminate spontaneous activity. However, it is also possible that a general postsynaptic hyperpolarization is only one factor in the recovery of responses to probes. In rat cortical slices, stimulation of one set of afferents does not necessarily reduce sensitivity to other sets (Abbott *et al.*, 1997), suggesting that some of the inhibitory or adaptive effects of leading stimuli may occur presynaptically.

D. Other response transformations

Although the nerve, AVCN, and SOC have similar population-recovery functions, individual neurons in the AVCN had flat or trough-shaped recovery functions not seen in the nerve. A larger response to the probe at short CPIs compared to the nerve is a seemingly odd result, since recovery should ultimately be limited by the nerve. Two possible mechanisms for this transformation suggest themselves. First, neurons can achieve a gain in recovery by convergence of inputs. Most nerve fibers show some response to the probe at the shortest CPIs tested (1 ms), and a considerable number are more than 50% recovered, so relatively little gain would be required. A second possibility is suggested by the form of the PST histograms for some neurons with flat or trough-shaped recovery functions (Fig. 4). These neurons had “chopper” responses, and the effect of the probe was to increase the probability of chopping. This suggests an interaction between the probe response and the time course of excitation in the response to the conditioner. These responses occurred only in a subset of chopper neurons, suggesting that the membrane currents that underlie chopping must be near threshold for an effect of the probe to appear.

Another transformation was from a relatively clear trend in the nerve and AVCN for louder conditioners and probes to cause longer periods of recovery, to a smaller overall trend in the IC and cortex with high variability among individual neurons. Behaviorally, the trend with stimulus level is fairly small, although in general the louder the stimulus the weaker the ability of the conditioner to suppress information in the probe (Babkoff and Sutton, 1966; Shinn-Cunningham *et al.*, 1993; Chiang and Freyman, 1998), in contrast to our neural results. The reason for this difference is not yet known. In the anesthetized cat the trend with level in the IC was in accord with the behavioral data (Litovsky and Yin, 1998a), but over smaller ranges of levels (usually 5–20 dB) than were tested in the unanesthetized preparations (up to 65 dB).

A transformation that was not related to recovery times was the change in spontaneous rates across structures. In the nerve, AVCN, and SOC, the range of spontaneous rates extended to ~100 spikes/s and the means were ~30 spikes/s, while in the IC and auditory cortex the range was to <70

spikes/s and the means were ~ 12 spikes/s. Thus, the IC and auditory cortex had similar and overall lower spontaneous rates than lower structures. A similar trend across structures has also been seen in the awake monkey (Ryan *et al.*, 1984). A difference was that in that in the monkey the SOC was reported to have a low, rather than high, average rate as found in the rabbit, but the measure was based on only five neurons.

A final transformation was the emergence of a population of neurons in the auditory cortex with recovery functions that showed tuning to specific ranges of CPIs. We did not encounter these neurons in lower structures, and such neurons have not been reported in previous studies. The tuning is reminiscent of echo-delay tuned neurons in the auditory cortex of echolocating bats that encode target range (e.g., Feng *et al.*, 1978; O'Neill and Suga, 1982). In terrestrial species, such neurons could signal the size of the acoustic space, since larger spaces will yield a larger range of delays than smaller spaces. The range of best CPIs observed was 5 to about 70 ms, corresponding to separations between sound and echo sources of about 2 to 24 m. The presence of such tuning suggests that echo-delay sensitivity in bats may be an elaboration of a system present in other species.

ACKNOWLEDGMENTS

We thank W. D'Angelo for comments on an earlier version of the manuscript and L. Seman for technical assistance. Supported by National Institutes of Health Grant Nos. DC-02178, DC-01366; DC-03948, and DC-00360.

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). "Synaptic depression and cortical gain control [see Comments]." *Science* **275**, 220–224.

Babkoff, H., and Sutton, S. (1966). "End point of lateralization for dichotic clicks." *J. Acoust. Soc. Am.* **41**, 87–102.

Backoff, P. M., Palombi, P. S., and Caspary, D. M. (1997). "Glycinergic and GABAergic inputs affect short-term suppression in the cochlear nucleus." *Hearing Res.* **110**, 155–163.

Barker, J. L., and Ransome, B. R. (1978). "Pentobarbitone pharmacology of mammalian central neurones grown in tissue culture." *J. Physiol. (London)* **280**, 355–372.

Blauert, J. (1982). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA).

Cant, N. B. (1992). "The cochlear nucleus: Neuronal types and their synaptic organization," in *The Mammalian Auditory Pathway: Neuroanatomy*, edited by D. B. Webster, A. N. Popper, and R. R. Fay (Springer, New York), pp. 66–116.

Chiang, Y. C., and Freyman, R. L. (1998). "The influence of broadband noise on the precedence effect." *J. Acoust. Soc. Am.* **104**, 3039–3047.

Clifton, R. K., and Freyman, R. L. (1989). "Effect of click rate and delay on breakdown of the precedence effect." *Percept. Psychophys.* **46**, 139–145.

Cranford, J. L. (1982). "Localization of paired sound sources in cats: Effects of variable arrival times." *J. Acoust. Soc. Am.* **72**, 1309–1311.

Cranford, J. L., and Oberholzer, M. (1976). "Role of neocortex in binaural hearing in the cat. II. The precedence effect in sound localization." *Brain Res.* **111**, 225–239.

Feng, A. S., Simmons, J. A., and Kick, S. A. (1978). "Echo detection and target-ranging neurons in the auditory system of the bat *Eptesicus fuscus*." *Science* **202**, 645–648.

Fitzpatrick, D. C., Kuwada, S., Batra, R., and Trahiotis, C. (1995). "Neural responses to simple simulated echoes in the auditory brainstem of the unanesthetized rabbit." *J. Neurophysiol.* **74**, 2469–2486.

Freyman, R. L., Clifton, R. K., and Litovsky, R. Y. (1991). "Dynamic processes in the precedence effect." *J. Acoust. Soc. Am.* **90**, 874–884.

Gaumond, R. P., Kim, D. O., and Molnar, C. E. (1982). "Response of cochlear nerve fibers to brief acoustic stimuli: Role of discharge-history effects." *J. Acoust. Soc. Am.* **74**, 1392–1398.

Grothe, B., and Sanes, D. H. (1993). "Bilateral inhibition by glycinergic afferents in the medial superior olive." *J. Neurophysiol.* **69**, 1192–1196.

Haas, H. (1951). "On the influence of a single echo on the intelligibility of speech." *Acustica* **1**, 48–58.

Helfert, R. H., Bonneau, J. M., Wentholt, R. J., and Altschuler, R. A. (1989). "GABA and glycine immunoreactivity in the guinea pig superior olivary complex." *Brain Res.* **501**, 269–286.

Irvine, D. R. F., Park, V. N., and McCormick, L. M. (1998). "Contributions of changes in the timing and amplitude of synaptic inputs to neural sensitivity to interaural intensity differences." in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London), pp. 359–367.

Joris, P. X., and Yin, T. C. T. (1995). "Envelope coding in the lateral superior olive. I Sensitivity to interaural time differences." *J. Neurophysiol.* **73**, 1043–1062.

Keller, C. H., and Takahashi, T. T. (1996a). "Binaural cross-correlation predicts the responses of neurons in the owl's auditory space map under conditions simulating summing localization." *J. Neurosci.* **16**, 4300–4309.

Keller, C. H., and Takahashi, T. T. (1996b). "Responses to simulated echoes by neurons in the barn owl's auditory space map." *J. Comp. Physiol. A* **178**, 499–512.

Kelly, J. B. (1974). "Localization of paired sound sources in the rat: Small time differences." *J. Acoust. Soc. Am.* **55**, 1277–1284.

Kidd, S. A., and Kelly, J. B. (1996). "Contribution of the dorsal nucleus of the lateral lemniscus to binaural responses in the inferior colliculus of the rat: Interaural time delays." *J. Neurosci.* **16**, 7390–7397.

Kim, D. O., Parham, K., Sirianni, J. G., and Chang, S. O. (1991). "Spatial response profiles of posteroventral cochlear nucleus neurons and auditory nerve fibers in unanesthetized decerebrate cats: Responses to pure tones." *J. Acoust. Soc. Am.* **89**, 2804–2817.

Lindemann, W. (1986). "Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front." *J. Acoust. Soc. Am.* **80**, 1623–1630.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999). "The precedence effect." *J. Acoust. Soc. Am.* **106**, 1633–1654.

Litovsky, R. Y., and Macmillan, N. A. (1994). "Sound localization precision under conditions of the precedence effect: Effects of azimuth and standard stimuli." *J. Acoust. Soc. Am.* **96**, 752–758.

Litovsky, R. Y., and Yin, T. C. T. (1998a). "Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics." *J. Neurophysiol.* **80**, 1285–301.

Litovsky, R. Y., and Yin, T. C. T. (1998b). "Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural mechanisms." *J. Neurophysiol.* **80**, 1302–1316.

McMullen, N. T., and Glaser, E. M. (1982). "Tonotopic organization of rabbit auditory cortex." *Exp. Neurol.* **75**, 208–220.

O'Neill, W. E., and Suga, N. (1982). "Encoding of target-range information and its representation in the auditory cortex of the mustached bat." *J. Neurosci.* **2**, 17–31.

Parham, K., Zhao, H.-B., and Kim, D. O. (1996). "Responses of auditory nerve fibers of the unanesthetized decerebrate cat to click pairs as simulated echoes." *J. Neurophysiol.* **76**, 17–29.

Parham, K., Zhao, H.-B., Ye, Y., and Kim, D. O. (1998). "Responses of AVCN neurons of the unanesthetized decerebrate cat to click pairs as simulated echoes." *Hearing Res.* **125**, 131–146.

Perrott, D. R., Strybel, T. Z., and Manligas, C. L. (1987). "Conditions under which the Haas precedence effect may or may not occur." *J. Aud. Res.* **27**, 59–72.

Peruzzi, D., Bartlett, E., Smith, P. H., and Oliver, D. L. (1997). "A mono-synaptic GABAergic input from the inferior colliculus to the medial geniculate body in rat." *J. Neurosci.* **17**, 3766–3777.

Populin, L. C., and Yin, T. C. T. (1998). "Behavioral studies of sound localization in the cat." *J. Neurosci.* **18**, 2147–2160.

Reale, R. A., Brugge, J. F., and Hind, J. E. (1995). "Encoding of stimulus direction by AI neurons in the cat: Effects of sounds arriving at different times and from different directions." *Soc. Neurosci. Abs.* **21**, 667.

Richter, J. A., and Holtman, J. R. (1982). "Barbiturates: their in vivo effects and potential biochemical mechanisms." *Neurobiology* **18**, 275–319.

Ryan, A. F., Miller, J. M., Pflingst, B. E., and Martin, G. K. (1984). "Effects

- of reaction time performance on single-unit activity in the central auditory pathway of the rhesus macaque," *J. Neurosci.* **4**, 298–306.
- Sanes, D. H. (1990). "An *in vitro* analysis of sound localization mechanisms in the gerbil lateral superior olive," *J. Acoust. Soc. Am.* **88**, 1442–1453.
- Shinn-Cunningham, B. S., Zurek, P. M., and Durlach, N. I. (1993). "Adjustment and discrimination measurements of the precedence effect," *J. Acoust. Soc. Am.* **93**, 2923–2932.
- Thurlow, W. R., and Parks, T. E. (1961). "Precedence-suppression events for two click sources," *Percept. Mot. Skills* **13**, 7–12.
- Tollin, D. J., and Henning, G. B. (1998). "Some aspects of the lateralization of echoed sound in man. I. The classical interaural-delay based precedence effect," *J. Acoust. Soc. Am.* **104**, 3030–3038.
- Wentholt, R. J., Huie, D., Altschuler, R. A., and Reeks, K. A. (1987). "Glycine immunoreactivity localized in the cochlear nucleus and superior olivary complex," *Neuroscience (NY)* **22**, 897–912.
- Wickesberg, R. E. (1996). "Rapid inhibition in the cochlear nuclear complex of the chinchilla," *J. Acoust. Soc. Am.* **100**, 1691–1702.
- Wickesberg, R. E., and Stevens, H. E. (1998). "Responses of auditory nerve fibers to trains of clicks," *J. Acoust. Soc. Am.* **103**, 1990–1999.
- Winer, J. A., Larue, D. T., and Pollak, G. D. (1995). "GABA and glycine in the central auditory system of the mustache bat: structural substrates for inhibitory neuronal organization," *J. Comp. Neurol.* **355**, 317–353.
- Winer, J. A., Saint Marie, R. L., Larue, D. T., and Oliver, D. L. (1996). "GABAergic feedforward projections from the inferior colliculus to the medial geniculate body," *Proc. Natl. Acad. Sci. USA* **93**, 8005–8010.
- Wu, S. H., and Kelley, J. B. (1992). "Binaural interaction in the lateral superior olive: Time difference sensitivity studied in mouse brain slice," *J. Neurosci.* **10**, 3494–3506.
- Wytenbach, R. A., and Hoy, R. R. (1993). "Demonstration of the precedence effect in an insect," *J. Acoust. Soc. Am.* **94**, 777–784.
- Yang, L., and Pollak, G. D. (1994). "The roles of GABAergic and glycinergic inhibition on binaural processing in the dorsal nucleus of the lateral lemniscus of the mustache bat," *J. Neurophysiol.* **71**, 1999–2013.
- Yin, T. C. T. (1994). "Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat," *J. Neurosci.* **14**, 5170–5186.
- Yost, W. A., and Soderquist, D. R. (1984). "The precedence effect: revisited," *J. Acoust. Soc. Am.* **76**, 1377–1383.
- Zurek, P. M. (1980). "The precedence effect and its possible role in the avoidance of interaural ambiguities," *J. Acoust. Soc. Am.* **67**, 952–964.
- Zurek, P. M. (1987). "The precedence effect," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer, New York), pp. 85–105.

Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio f_2/f_1 ^{a)}

Manfred Mauermann^{b)} and Stefan Uppenkamp^{c)}

AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany

Peter W. J. van Hengel

BCN, Rijksuniversiteit Groningen, 9747 AG Groningen, The Netherlands

Birger Kollmeier

AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany

(Received 10 January 1999; accepted for publication 24 August 1999)

Critical experiments were performed in order to validate the two-source hypothesis of distortion product otoacoustic emissions (DPOAE) generation. Measurements of the spectral fine structure of DPOAE in response to stimulation with two sinusoids have been performed with normal-hearing subjects. The dependence of fine-structure patterns on the frequency ratio f_2/f_1 was investigated by changing f_1 or f_2 only (fixed f_2 or fixed f_1 paradigm, respectively), and by changing both primaries at a fixed ratio and looking at different order DPOAE. When f_2/f_1 is varied in the fixed ratio paradigm, the patterns of $2f_1 - f_2$ fine structure vary considerably more if plotted as a function of f_2 than as a function of f_{DP} . Different order distortion products located at the same characteristic place on the basilar membrane (BM) show similar patterns for both, the fixed- f_2 and f_{DP} paradigms. Fluctuations in DPOAE level up to 20 dB can be observed. In contrast, the results from a fixed- f_{DP} paradigm do not show any fine structure but only an overall dependence of DP level on the frequency ratio, with a maximum for $2f_1 - f_2$ at f_2/f_1 close to 1.2. Similar stimulus configurations used in the experiments have also been used for computer simulations of DPOAE in a nonlinear and active model of the cochlea. Experimental results and model simulations give strong evidence for a two-source model of DPOAE generation: The first source is the initial nonlinear interaction of the primaries close to the f_2 place. The second source is caused by coherent reflection from a re-emission site at the characteristic place of the distortion product frequency. The spectral fine structure of DPOAE observed in the ear canal reflects the interaction of both these sources. © 1999 Acoustical Society of America. [S0001-4966(99)02812-X]

PACS numbers: 43.64.Jb, 43.64.Kc [BLM]

INTRODUCTION

Narrow-band distortion product otoacoustic emissions (DPOAE) are low-level sinusoids recordable in the occluded ear canal at certain combination frequencies during continuous stimulation with two tones. They are the result of the nonlinear interaction of the tones in the cochlea. In human subjects, DPOAE typically exhibit a pronounced spectral fine structure when varying the frequencies of both primaries simultaneously (f_1, f_2) at a fixed frequency ratio f_2/f_1 (Gaskill and Brown, 1990; He and Schmiedt, 1993). The variations of DPOAE level with frequency show a periodicity of about 3/32 octaves (He and Schmiedt, 1993; Mauermann *et al.*, 1997b) in a depth up to 20 dB. DPOAE can be recorded in almost any normal-hearing subject and in sub-

jects with hearing loss up to 50 dB HL (Smurzynski *et al.*, 1990). Because of the narrow-band nature of both stimuli and emissions, they provide a frequency-specific method to explore cochlear mechanics. Therefore, DPOAE are of great interest not only in laboratory studies but also as a diagnostic tool for clinical audiology. However, since it is not as yet completely understood which sources along the cochlear partition contribute to the emission measured in the ear canal, the applicability of DPOAE for, e.g., “objective” audiometry is limited at present.

For DPOAE with frequencies below the primary frequencies ($2f_1 - f_2, 3f_1 - 2f_2$, etc.), it is widely accepted that the generation site is the overlap region of the excitation patterns of the two primaries, which has a maximum close to the characteristic site around f_2 . Although the generation of distortion products due to the interaction of the two primaries is in principle spread over the whole cochlea, a region of about 1 mm around the characteristic place of f_2 has been suggested to give the maximum contribution (van Hengel and Duifhuis, 1999). This region of maximum contribution is referred to as the f_2 site.

^{a)}Parts of this study were presented at the 1998 short papers meeting of the British Society of Audiology in London [Uppenkamp and Mauermann, Br. J. Audiol. 33, 87–88 (A) (1999)].

^{b)}Author to whom correspondence should be addressed. Electronic mail: mani@medi.physik.uni-oldenburg.de

^{c)}Current address: CNBH, Department of Physiology, University of Cambridge, Downing Street, Cambridge CB2 3EG, U.K.

It is still a point of discussion whether the generation site is the only source or to what extent other sources might also contribute to the emission. The DPOAE fine structure found in human subjects is closely related to this question of DPOAE sources. The fine structure might reflect local BM properties of either the generation site or of the re-emission site. It could also result from the interference between two or more sources or even from a combination of both local properties and interference effects. Studying the properties of DPOAE fine structure may result in further insight into BM mechanisms and the location of DP sources.

The patterns of DPOAE fine structure get shifted along the frequency axis when the primary levels are increased (He and Schmiedt, 1993; Mauermann *et al.*, 1997b) or the frequency ratio of the primaries is changed (Mauermann *et al.*, 1997a). On one hand, these shifts may cause some problems in the interpretation of DPOAE measurements, especially DPOAE growth functions. Peaks may change to notches or vice versa. This is most probably the reason for the notches found in human DPOAE growth functions (He and Schmiedt, 1993) and is critical for a direct correlation of DPOAE level to hearing threshold. On the other hand, the correct interpretation of these fine-structure shifts can aid in a detailed understanding of BM mechanisms.

He and Schmiedt (1993) showed that the level-dependent shift of DPOAE fine structure is consistent with the results of Ruggero and Rich (1991) on the shift of the maximum basilar-membrane response in the chinchilla. They interpreted the fine structure as an effect of local BM properties in the region of the primaries and the level-dependent shift as a result of the shift of the primary excitation on the BM (He and Schmiedt, 1993; Sun *et al.*, 1994a, b). Varying only one primary level while holding the other fixed causes pattern shifts in different directions, dependent on whether the primary level at f_1 or at f_2 is varied. He and Schmiedt (1997) argued that these effects strongly support the idea that the DPOAE fine structure might reflect mechanical properties of the overlapping area of the primary excitation.

However, Heitmann *et al.* (1998) showed that the fine structure disappears when the DPOAE is measured with a third tone close to the distortion product frequency (f_{DP}) as a suppressor. This result is interpreted as evidence for an additional second source around the characteristic place of $2f_1 - f_2$, which has a major influence on the fine-structure pattern. The contribution of a second source at the place of $2f_1 - f_2$ is also supported by several experiments on DPOAE suppression. Kummer *et al.* (1995) reported that in some cases a suppressor close to $2f_1 - f_2$ results in more suppression than one close to f_2 . Gaskill and Brown (1996) also found that the DP level is still sensitive to a suppressor near $2f_1 - f_2$, although the major suppression effects they observed were for suppressor frequencies close to f_2 . Brown *et al.* (1996) showed "that it may be legitimate to analyze DP as vector sum of two gross components," (Brown *et al.*, 1996, p. 3263).

Throughout the present paper, experimental results from normal-hearing subjects and computer simulations will be presented examining in detail the properties of DPOAE fine structure for equal-level primaries and varied ratios of the

primary frequencies. Three "critical" experiments have been performed, using different experimental paradigms, that aim to clarify where and how DPOAE fine structure is generated. Experiment 1 investigates DPOAE fine-structure patterns for different f_2/f_1 to determine whether the fine structure is dominated by local properties of the f_2 region or if a supposed re-emission site around f_{DP} is of some importance. Experiment 2 is designed to test the influence of the relative phase between the suggested emission sites by investigating the patterns of different order DPOAE. Finally, in experiment 3 the DPOAE patterns from a f_{DP} -fixed and a f_2 -fixed paradigm are compared to find out if DPOAE fine structure is mainly influenced by one of these two sites.

In addition to the recordings from subjects, all experimental paradigms were also assessed with a computer simulation of DPOAE using a nonlinear and active transmission line model of the cochlea. This model includes an impedance function as suggested by Zweig (1991) which produces excitation patterns with a broad and tall peak. Within the model, statistical fluctuations of stiffness along the cochlear partition are sufficient to create quasiperiodic OAE fine-structure patterns, as reported by Zweig and Shera (1995).

I. METHODS

A. Subjects

Seven normal-hearing subjects, ranging in age from 25 to 30 years, participated in this study. Their hearing thresholds were better than 15 dB HL for all audiometric frequencies in the range 250 Hz–8 kHz, and none of the subjects had a history of any hearing problems. Spontaneous otoacoustic emissions (SOAE) were observed in only one of the subjects (subject SE). DPOAE were recorded from one ear of each subject during several sessions lasting 60–90 min. During the sessions, the subjects were seated comfortably in a sound-insulated booth (IAC-1200 CT).

B. Instrumentation and signal processing

An insert ear probe, type ER-10C, was used to record DPOAE. The microphone output was connected to a low-noise amplifier, type SR560, and then converted to digital form using the 16-bit A/D converters on a signal-processing board (Ariel DSP-32C) in a personal computer. All stimuli were generated digitally. After D/A conversion by the 16-bit D/A converters on the Ariel board and low-pass filtering (Kemo VBF44, 8.5 kHz) they were presented to the subjects via a computer-controlled audiometer. The DSP was used for on-line analysis and signal conditioning of the recorded emissions, including artifact rejection, averaging in the time domain to improve the signal-to-noise ratio, and fast Fourier transform (FFT). For each signal configuration, at least 16 but usually 256 frames were averaged, using a frame length of 186 ms (4096 samples). If required, the number of averages could be increased during the recording to get a sufficient signal-to-noise ratio for the frequencies of interest.

Two sinusoids were generated as even harmonics of the frame rate (5.38 Hz) at a sampling rate of 22 050 Hz. The tones were presented continuously to the subject. To compensate for the ear-canal transfer function, an individual ad-

justment of the primaries to the desired sound-pressure level of 60 dB SPL was performed automatically before each run, taking into account the transfer function of the probe microphone. The variations of attenuation within and between subjects were approximately within a range of 5 dB.

C. Experimental paradigms

1. Dependence of the $2f_1-f_2$ DPOAE on frequency ratio

In experiment 1, the effect of the frequency ratio f_2/f_1 on the fine-structure patterns of the DPOAE at $2f_1-f_2$ was investigated. Fine-structure patterns for this distortion product were recorded in all subjects for seven different frequency ratios f_2/f_1 , fixed at 1.07, 1.1, 1.13, 1.16, 1.19, 1.22, and 1.25. Due to the additional requirement to select the primaries as harmonics of the frame rate, minor deviations up to 0.002 from the desired frequency ratio were present. Recordings were taken covering a frequency range of two octaves ($f_2 = 1-4$ kHz), divided into four sessions covering half an octave each for all of the different frequency ratios specified above. The frequency step between adjacent single recordings was 1/48 octave.

It is assumed that the small changes in the frequency ratio cause only small changes in the fine-structure patterns, i.e., the patterns remain comparable. Consequently, if the fine-structure patterns are mostly influenced by the local properties of the generation site near the characteristic place of f_2 , the patterns for different f_2/f_1 should show a high stability when plotted as a function of f_2 . If, however, the local properties of a presumed re-emission site near the characteristic place of the DPOAE frequency play a major role, the stability of the patterns should be greater when plotted as a function of f_{DP} .

2. Different order DPOAE

In experiment 2, fine-structure patterns for different order DPOAE (e.g., $2f_1-f_2$, $3f_1-2f_2$, and $4f_1-3f_2$) were recorded from six of our seven subjects. The frequency ratios were chosen to get identical frequency pairs either for f_2 and $2f_1-f_2$, f_2 and $3f_1-2f_2$, and f_2 and $4f_1-3f_2$, or for f_1 and $2f_1-f_2$, f_1 and $3f_1-2f_2$, and f_1 and $4f_1-3f_2$, respectively. The condition for identical f_2 and f_{DP} frequencies is fulfilled for the frequency ratios $f_2/f_1 = 1.22$, 1.137, and 1.099. Similarly, the required identity of f_1 and f_{DP} frequencies is fulfilled at frequency ratios $f_2/f_1 = 1.22$, 1.11, and 1.073 (see Fig. 1).

If the DPOAE fine structure is mainly caused by two sources, one at the generation site and the other one at the distortion frequency site, measurements for different order DPOAE with identical f_2 and f_{DP} frequencies should result in very similar patterns, since both f_2 and the observed DP frequency are the same, i.e., the characteristic places of the two assumed sources and hence the phase relation between the two is almost identical (see Fig. 1, left column). With identical f_1 and f_{DP} frequencies, a small variation in the pattern is expected, indicating the influence of the change in the relative phase of the f_2 and f_{DP} components (see Fig. 1, right column).

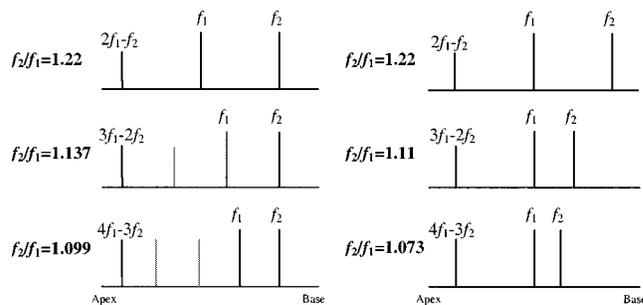


FIG. 1. Sketch of stimulus configuration for the comparison of different order DPOAE fine-structure patterns. For the condition of identical f_2 and f_{DP} frequencies, f_2/f_1 was set to 1.22, 1.137, and 1.099 to get identical DP frequencies $f_{DP} = 0.64f_2$ for $2f_1-f_2$, $3f_1-2f_2$, and $4f_1-3f_2$, respectively. Similarly, to fulfill the condition of identical f_1 and f_{DP} frequencies, f_2/f_1 was set to 1.22, 1.11, and 1.073, resulting in $f_{DP} = 0.78f_1$.

3. Fixed f_2 vs fixed f_{DP}

An additional test for investigating the source of the fine structure was performed during experiment 3. DPOAE were recorded in keeping either f_{DP} or f_2 fixed. This was achieved by varying both f_1 and f_2 while keeping f_{DP} fixed at 2 kHz or varying f_1 and keeping f_2 fixed at 3 kHz resulting in varying f_{DP} . Both paradigms result in a varying frequency ratio f_2/f_1 .

The comparison of the two paradigms should reveal the relative contribution of the two supposed sources. If the fine-structure pattern of the DPOAE is dominated by the contribution from the characteristic place of the distortion product frequency, it is expected that the observable pattern shows much less variation between minima and maxima when f_{DP} is held constant.

D. Analysis

For further analysis, the frequencies in all experiments were transformed to their characteristic places $x(f)$ on the BM using the place-frequency map proposed by Greenwood (1991). Although the frequencies used are almost equally spaced on the Greenwood map, there are some deviations from this. These deviations are mainly due to the fact that the primaries were selected as harmonics of the frame rate. To compensate for that, the data were interpolated using a cubic spline algorithm (MATLAB 5.1) and resampled at 1024 points equally spaced on the Greenwood map. Cross-correlation functions (CCF) were calculated using the data from experiment 1 to quantify the shift between two different fine-structure patterns. The correlation lag giving the maximum of the CCF within the range ± 1 mm was taken as shift between two fine-structure patterns with adjacent frequency ratios. It is assumed that small changes in frequency ratio will cause only small shifts of the overall fine structure. Therefore, the range to look for maxima of the CCF was limited to avoid ambiguities which could be caused by the quasiperiodic shape of the patterns. The computation of CCF was always restricted to the area of actual overlap between the compared patterns.

II. EXPERIMENTAL RESULTS

A. Experiment 1: Dependence on f_2/f_1

Figure 2 shows the effect of variation of f_2/f_1 (seven different f_2/f_1 ratios) on the DP fine-structure patterns for the $2f_1-f_2$ distortion product for four different subjects. The left column shows the results plotted as a function of f_2 place, and the right column shows the same data as a function of $2f_1-f_2$. The labeling of the ordinate holds for the bottom trace only. Each successive trace that corresponds to a different f_2/f_1 ratio is shifted by 20 dB. Note the pronounced shift in the basal direction of successive patterns when plotted as a function of f_2 . This contrasts with the small shift in the apical direction when plotted as a function of $2f_1-f_2$. This is illustrated by the lines in the middle column showing the cumulative sum of correlation lag for the maxima in the CCF between successive patterns. The similarity between adjacent patterns is relatively high for small differences in f_2/f_1 . However, the patterns become more different when the changes in frequency ratio get bigger.

B. Experiment 2: Fine structure of different order DPOAE

Figure 3 shows fine-structure patterns for different order DPOAE characterized by the same distance along the basilar membrane between f_2 and f_{DP} (right column) or f_1 and f_{DP} (left column) for the same four subjects as in Fig. 2. As Fig. 3 illustrates, the patterns for same f_2 and f_{DP} frequencies are very similar, suggesting that the relative phase between the DP components contributing from the characteristic places of f_2 and f_{DP} plays an important role in the DPOAE fine structure. For identical f_1 and f_{DP} the patterns still look similar. However, for most of the subjects the CCF indicates a slight shift in the basal direction for the higher-order DPOAE. This is consistent with the movement of the f_2 place in the apical direction with increasing order in this case.

C. Experiment 3: Fixed f_2 vs fixed f_{DP}

While the results from experiments 1 and 2 indicate that the DPOAE observed in the human ear canal stems from two sources along the cochlea partition, one close to the f_2 site and one at the f_{DP} site, a separation of the contribution of these two sources cannot be achieved using these data. Figure 4 shows DPOAE patterns for four subjects (two of them as in Figs. 2 and 3) obtained with a fixed f_2 (gray line) and with a fixed f_{DP} (black line). The use of a fixed f_2 results in patterns very similar to the ones observed before, using the fixed-ratio paradigm. In contrast, the use of a fixed f_{DP} greatly reduces the fine structure. There remains only an overall dependence of DP level on frequency ratio, with a maximum for $2f_1-f_2$ at f_2-f_1 around 1.2, as reported, e.g., by Harris *et al.* (1992).

III. SIMULATIONS IN A NONLINEAR AND ACTIVE COCHLEA MODEL

A. Description of the model

For computer simulations of the observed effects, a one-dimensional nonlinear and active model of the cochlea was used, based on a model described in previous work (van Hengel *et al.*, 1996; van Hengel and Duifhuis, 1999). In these papers it was shown that the model, which operates in the time domain, is very useful to study nonlinear phenomena such as OAE. The basis for such models has been described in more detail in Duifhuis *et al.* (1985) and van den Raadt and Duifhuis (1990). In previous simulations of DPOAE, it turned out that a possible shortcoming of the model was that it did not produce a high and broad excitation peak for pure-tone stimuli (van Hengel and Duifhuis, 1999). It is claimed by various authors that such a peak is necessary to properly simulate cochlear behavior at low stimulus levels (e.g., Zweig, 1991; de Boer, 1995). Furthermore, it was claimed by Shera and Zweig (1993) and shown by Talmadge *et al.* (1993, 1998) that the impedance function suggested by Zweig (1991), which produces a high and broad excitation peak, also produces a fine structure in various types of simulated emissions when it is combined with a “roughness” in the mechanical parameters of the cochlear partition. This roughness is a random fluctuation of (one of) the parameters describing the mechanics of the sections of the cochlear partition used in the model, and reflects random inhomogeneity in the placement and behavior of cells along the cochlea, especially the outer hair cells. The impedance function described by Zweig (1991) was therefore incorporated in the model, as well as the possibility of introducing roughness. The resulting model consists of 600 sections¹ equally spaced along the length of the cochlea (35 mm). The motion of the cochlear partition in each section is described by the following equation of motion:

$$m\ddot{y}(x) + d(x, v)\dot{y}(x) + s(x)[y(x) + c(v)y(x)|_{t-\tau}] = p(x). \quad (1)$$

This is a normal second-order differential equation of motion for a harmonic oscillator with mass m , damping $d(x, v)$, and stiffness $s(x)$, driven by a pressure force $p(x)$ (x is the position of the oscillator measured from the stapes along the cochlea, y is the displacement, and $v = \dot{y}$ the velocity of the cochlear partition in the vertical direction), except that there is an additional “delayed feedback stiffness” term $s(x)c(v)y(x)|_{t-\tau}$. This term was derived by Zweig (1991) from fits to experimental data on BM excitation patterns. It serves to stabilize the motion of the oscillator, counteracting a negative damping term $d(x, v)$. In order to do so, and to arrive at the desired high and broad peak in the excitation caused by a pure tone, the time delay τ must depend on the resonance frequency $\omega_{\text{res}} = \sqrt{s(x)/m}$ of the oscillator as $\tau = 1.742 \times 2\pi/\omega_{\text{res}}$ (Zweig, 1991). The values of the parameters $d(x, v)$ and $c(v)$ determined by Zweig were $-0.1217\sqrt{ms(x)}$ and 0.1416, respectively. It is important to note that these values are derived from estimated excitation patterns in the cochlea of the squirrel monkey at low levels of stimulation and in the frequency range around 8 kHz.

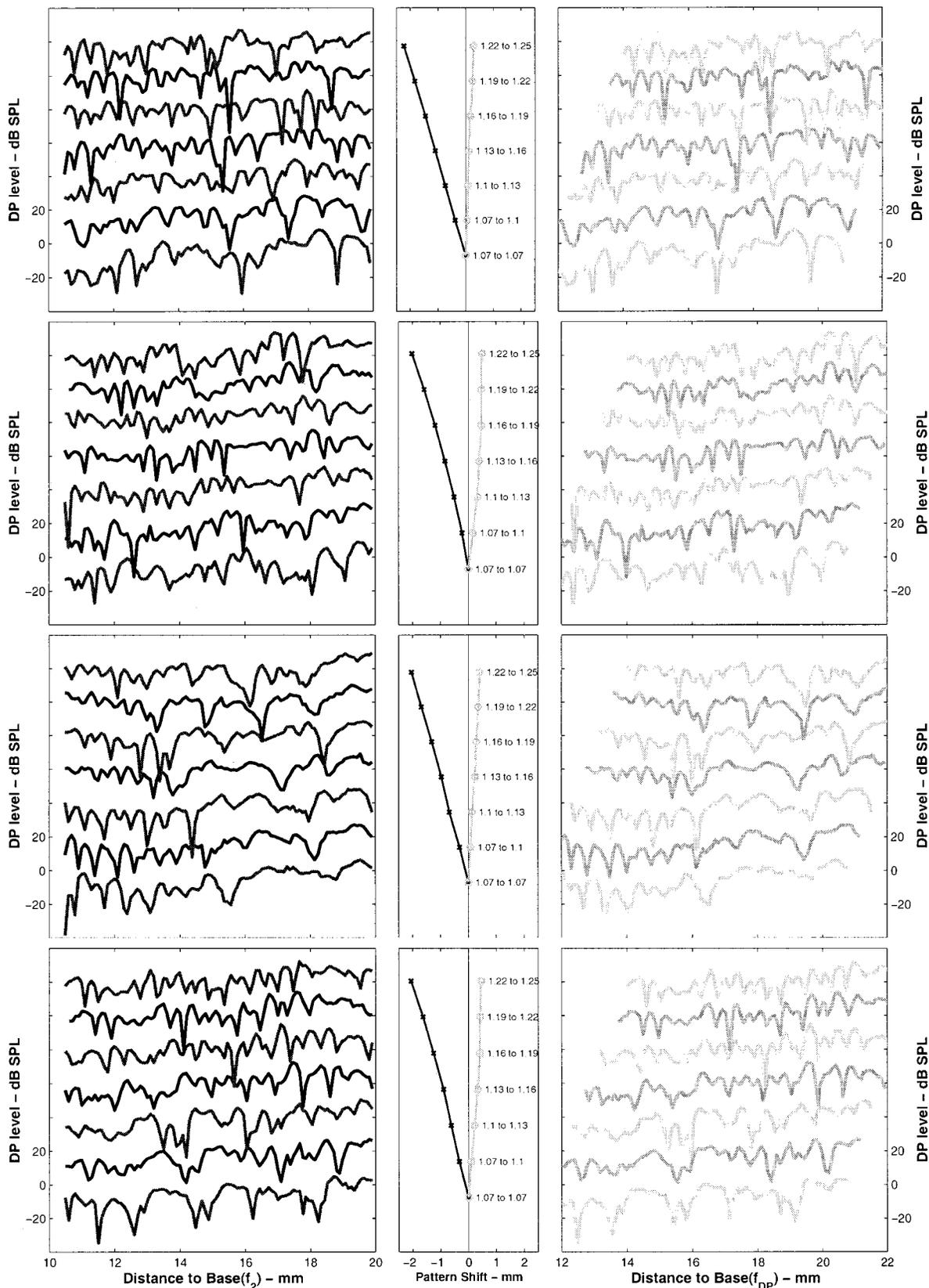


FIG. 2. Experiment 1: Dependence of $2f_1 - f_2$ DPOAE fine structure on the frequency ratio f_2/f_1 for four different subjects (from top to bottom: KI right, MG left, MK right, and MM right). Each row shows DPOAE fine-structure patterns for seven different ratios f_2/f_1 (from bottom to top: 1.07, 1.1, 1.13, 1.16, 1.19, 1.22). Left column: DP level as a function of f_2 place. 10-mm distance to base corresponds to a frequency $f_2 = 4358$ Hz, and 20 mm distance corresponds to $f_2 = 986$ Hz. Right column: same data as a function of $2f_1 - f_2$. 12 mm corresponds to $f_{DP} = 3270$ Hz, and 22 mm corresponds to $f_{DP} = 713$ Hz. The labeling of the ordinate holds for the bottom trace only. Each successive trace is shifted up by 20 dB. Middle column: shift between the different DPOAE patterns when plotted as a function of f_2 (black line) and when plotted as a function of $2f_1 - f_2$ (gray line). This overall shift is quantified by the correlation lag for the maximum of the cross-correlation function between adjacent patterns. Each data point results from cumulative summation of the shifts between adjacent patterns (as indicated by the numbers “1.07 to 1.1,” “1.1 to 1.13,” etc.), starting at the frequency ratio 1.07.

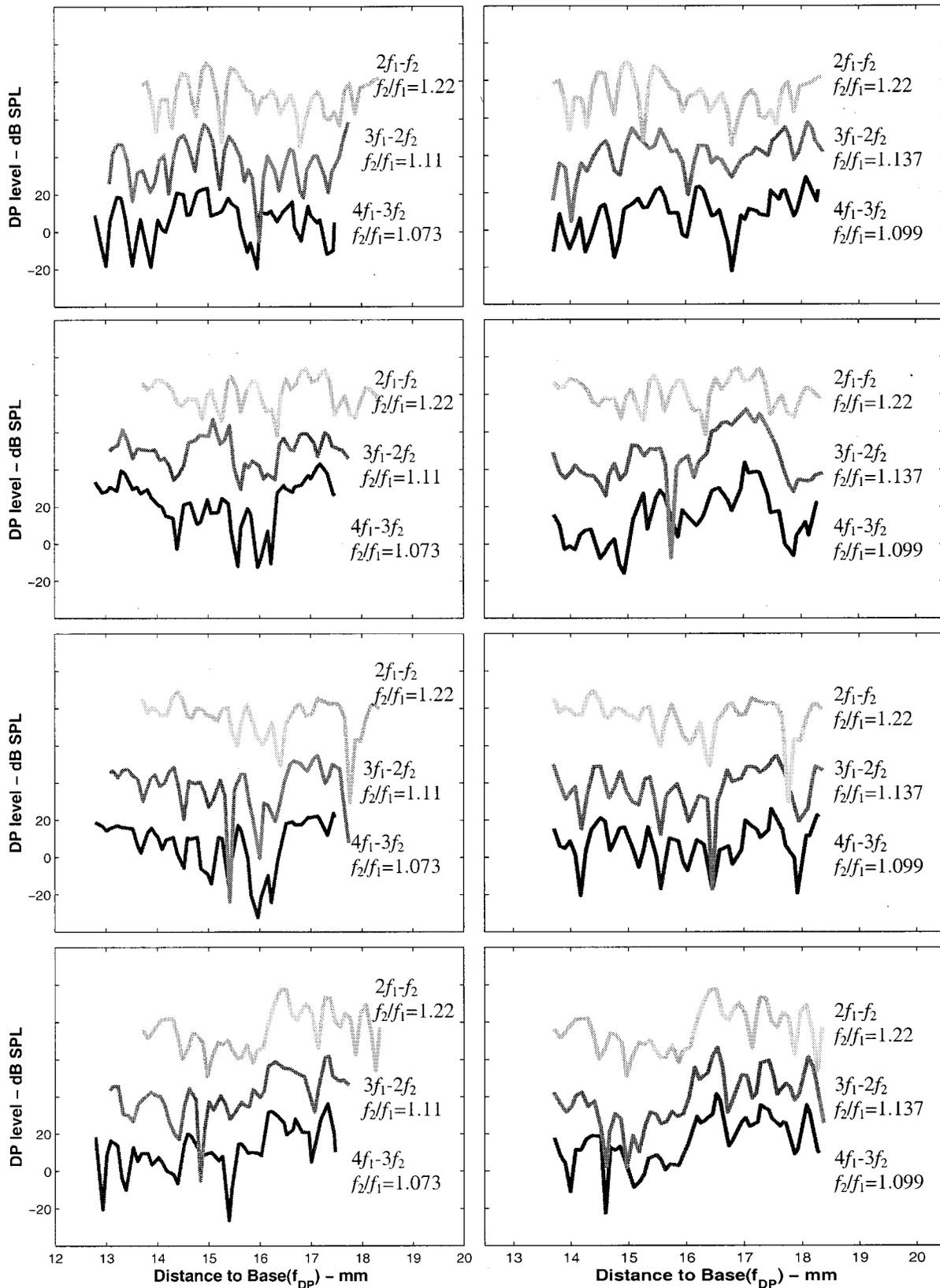


FIG. 3. Experiment 2: Comparison of different order DPOAE fine-structure patterns for four subjects. Left column: Identical f_1 and f_{DP} frequencies, right column: Identical f_2 and f_{DP} frequencies. The frequency ratios f_2/f_1 were chosen according to the scheme in Fig. 1. Note the very similar patterns for all orders of DPOAE. The labeling of the ordinate holds for the bottom trace only. Each successive trace is shifted by 20 dB. 13 mm corresponds to $f_{DP}=2830$ Hz. Same subjects as in Fig. 2.

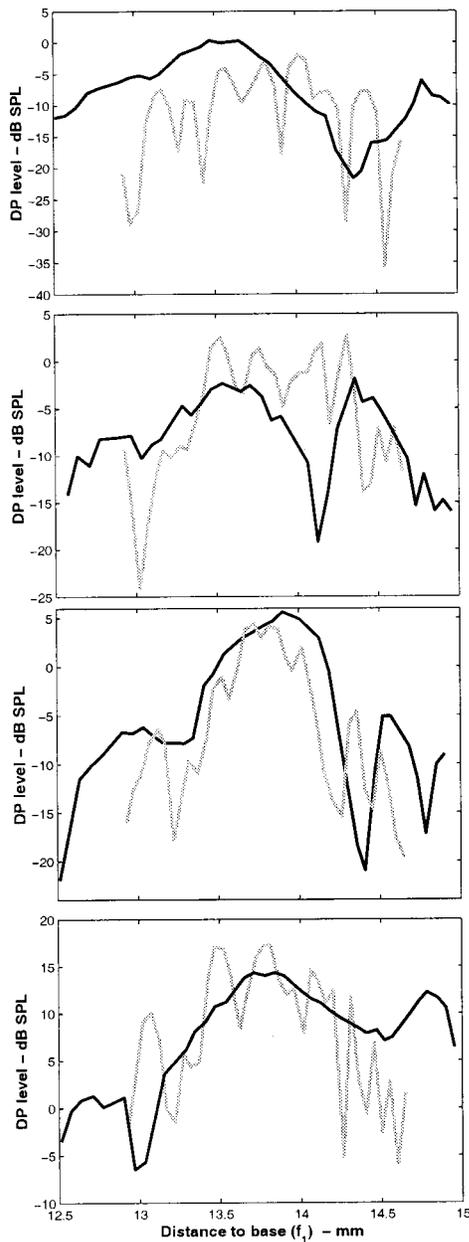


FIG. 4. Experiment 3: DP level as a function of f_1 for four subjects. Black lines: fixed- f_{DP} paradigm, i.e., varying both f_1 and f_2 while keeping f_{DP} fixed at 2 kHz. Gray lines: fixed- f_2 paradigm, i.e., varying f_1 and f_{DP} while f_2 is fixed at 3 kHz. Note the fine structure for the fixed- f_2 paradigm, which is similar to the patterns observable with the fixed-ratio paradigm, while there is no fine structure when using the fixed f_{DP} paradigm. Subjects from top to bottom: MK right, MM right, OW right, and SE right.

These values certainly cannot be used in the vicinity of both stapes and helicotrema, since this would lead to instability (van Hengel, 1993). It is also clear that these values do not hold for higher stimulus levels. Both the negative damping term and the stabilizing delayed feedback stiffness term are thought to result from active, i.e., energy producing, behavior of the outer hair cells. This active behavior must saturate at higher levels. It is therefore logical to capture the nonlinearities present in cochlear mechanics in the terms $d(x, v)$ and $c(v)$.² The nonlinearity was introduced by assuming the following dependence of $d(x, v)$ and $c(v)$ on the velocity v of the section:

$$d(x, v) = \left[d_l + \frac{\beta(d_h - d_l)|v|}{1 + \beta|v|} \right] \sqrt{ms(x)}$$

with $d_l = -0.12$, $d_h = 0.5$, (2a)

$$c(v) = c_l + \frac{-\beta d_l |v|}{1 + \beta|v|} \quad c_l = 0.1416. \quad (2b)$$

In these equations the nonlinear behavior of $d(x, v)$ and $c(v)$ is chosen to be the same, with the damping going to a value of $d_h \sqrt{ms(x)}$ and the delayed feedback stiffness disappearing at high excitation levels. The region in which the nonlinearity plays a role is determined by the parameter β . In all simulations presented here, a value of 0.01 ms/nm was used for this parameter, leading to a compressive growth of the excitation at the characteristic place for a pure-tone stimulus of around 0.3 dB/dB over the range from about 20 to 80 dB SPL stimulus level. Because the mass m was chosen to be 0.375 kg/m², independent of the position x , the stiffness controls the place-frequency map of the cochlear partition. The place-frequency map chosen was

$$f(x) = A \times e^{-ax}, \quad A = 22.508 \text{ kHz}, \quad a = 150 \text{ m}^{-1}. \quad (3)$$

Following the work of Shera and Zweig (1993) and Talmadge and Tubis (1993), the random fluctuations necessary to obtain a fine structure in the emissions were introduced in the stiffness as

$$s(x) = s(x)[1 + r_0 r(x)], \quad (4)$$

where $r(x)$ is a random variable with a Gaussian distribution and r_0 is a scaling parameter that controls the amount of roughness. For the results presented here, a value of 1% was used for r_0 . The equations of motion [Eq. (1)] for all sections were coupled through the fluid, which was assumed linear, incompressible, and inviscid. The coupled system was solved by Gauss elimination and integrated in time using a Runge-Kutta 4 time-integration scheme with a sampling frequency of 150 kHz. Reducing the sampling frequency could lead to instabilities in certain cases, but increasing it did not give significantly different results (differences in emission levels were below 0.5%).

To simulate OAE, the motion of the cochlea sections is coupled to the outside world via a simplified middle ear, consisting of a mass, stiffness, and damping in combination with a transformer. This produces a sound pressure that would result at the eardrum in an open ear canal. Previous studies with this model have shown that emission levels are highly sensitive to conditions at the eardrum. SOAE level may vary up to 30 dB when different loading impedances are added (van den Raadt and Duifhuis, 1993).

B. Simulations

All the experiments described in Sec. II (except experiment 2 for the condition of same f_1 and f_{DP} frequencies) were simulated using this computer model. In contrast to the experiments, all simulations were performed with primary levels of 50 dB SPL. The model output is the sound-pressure level at the eardrum taken from a 30-ms interval beginning 20 ms after stimulus start to avoid onset effects. The data

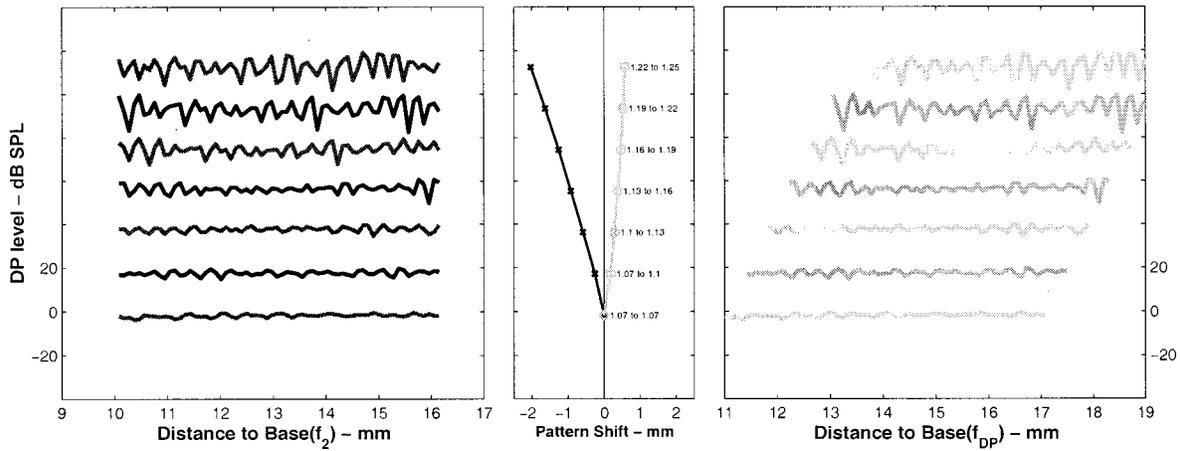


FIG. 5. Computer simulation of experiment 1: Dependence of $2f_1 - f_2$ DPOAE fine structure on the frequency ratio f_2/f_1 . Left column: DP level as a function of f_2 place. 10 mm distance to base corresponds to a frequency $f_2 = 5027$ Hz, and 16-mm distance corresponds to $f_2 = 2045$ Hz, according to the exponential place-frequency map used in the cochlea model. Right column: same data as a function of $2f_1 - f_2$ place. 11 mm corresponds to $f_{DP} = 4327$ Hz, and 19 mm corresponds to $f_{DP} = 1304$ Hz. As in Fig. 2, the shift between adjacent patterns is illustrated by the lines in the middle column of the figure based on cross correlation.

were analyzed using the least-squares-fit method described by Long and Talmadge (1997) to get an estimate of the spectral power of the frequency components of interest. The further analysis of pattern shifts was performed using cross correlation as described for the experimental data.

The simulation results are given in Figs. 5–7. Analogous to the experimental data in Fig. 2, Fig. 5 shows simulated DPOAE fine-structure patterns for different frequency ratios of the primaries plotted as a function of f_2 (left panel) and f_{DP} (right panel). The main experimental result, i.e., the shift of fine-structure patterns when varying the frequency ratio, is replicated very well, both qualitatively and quantitatively. Figure 6 shows fine-structure patterns for three different frequency ratios, which were chosen to get the same frequencies for f_2 and $2f_1 - f_2$, $3f_1 - 2f_2$, or $4f_1 - 3f_2$, respectively. As in the experiments, the patterns for different order DPOAE were almost identical in all these stimulus conditions. Figure 7 shows that, in the model, the fine structure

disappears in a fixed f_{DP} paradigm, similar to the experimental results, while the model still produces a fine structure for fixed f_2 . The only discrepancies between simulations and experimental results in the fine structure are the reduced dynamical range between maxima and minima in the model for small frequency ratios and the slightly smaller period of the frequency-dependent level variations.³

Simulations have also been performed omitting the roughness in the model's stiffness function in either (1) the frequency region above 2 kHz, or (2) below 2 kHz, or (3) with no roughness at all, to get a better understanding of the mechanism creating the fine structure in the model. DPOAE with a high frequency resolution were computed over a frequency range for f_2 from 2483 to 3084 Hz at a frequency ratio $f_2/f_1 = 1.22$. This ensured that the characteristic places of f_{DP} always fell in model sections with characteristic frequencies below 2 kHz, while the characteristic frequencies of

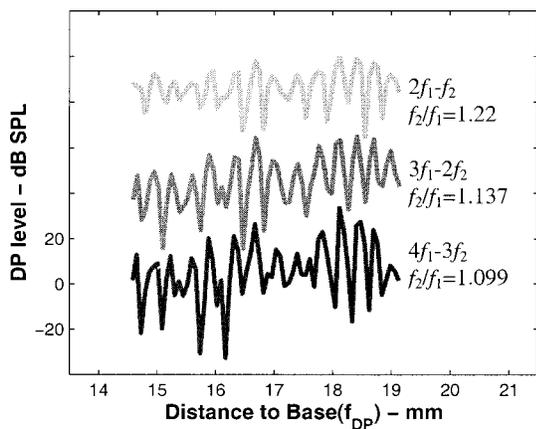


FIG. 6. Computer simulation of experiment 2: Comparison of different order DPOAE fine-structure patterns for identical f_2 and f_{DP} frequencies. The frequency ratios f_2/f_1 were chosen according to the scheme in Fig. 1. Note the very similar patterns for all orders of DPOAE. The labeling of the ordinate holds for the bottom trace only. Each successive trace is shifted by 20 dB. 14 mm corresponds to $f_{DP} = 2760$ Hz, according to the exponential frequency map used in the model.

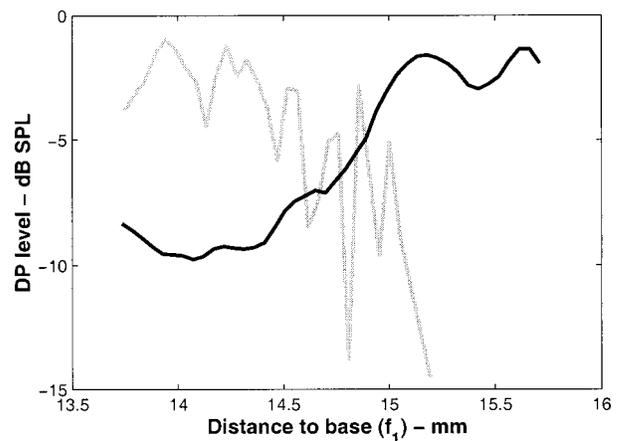


FIG. 7. Computer simulation of experiment 3: DP level as a function of f_1 . Black line: fixed f_{DP} paradigm, i.e., varying both f_1 and f_2 while keeping f_{DP} fixed at 2 kHz. Gray line: fixed f_2 paradigm, i.e., varying f_1 and f_{DP} while f_2 is fixed at 3 kHz. Note that, as for the experimental results, the fine structure for the fixed- f_2 paradigm is similar to the patterns observable with the fixed-ratio paradigm, while there is no fine structure when using the fixed- f_{DP} paradigm.

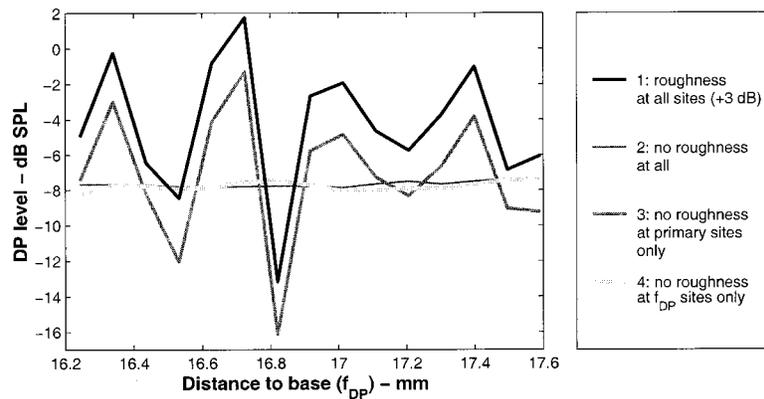


FIG. 8. Computer simulations at a frequency ratio $f_2/f_1=1.22$, but omitting the roughness in different parts of the model cochlea. The primary levels were $L_1=L_2=50$ dB SPL. Line 1 is the reference simulation, a DPOAE fine structure for roughness over the whole cochlea (shifted 3 dB up). Line 2 shows the DPOAE levels for a model cochlea without any roughness in the stiffness function. No fine structure can be observed here. Line 3 shows the DPOAE fine structure using a model cochlea without any fine structure in the region of the primaries. This has no effect on the fine structure, which is almost identical to that for the reference simulation. Line 4 shows DPOAE levels produced by a model cochlea with no roughness at the f_{DP} sites only and the same roughness as in the reference in the primary region. Note that for this condition the DPOAE fine structure disappears completely.

the primaries always fell into sections above 2 kHz. Figure 8 shows the simulation results in these three conditions as well as in the reference condition with roughness over the whole length of the cochlea. The DPOAE fine structure is unaffected by the presence or absence of the roughness in the primary region, while it disappears when the roughness is omitted in the distortion product frequency region. This emphasizes the interpretation of the experimental results that the DPOAE fine structure is mainly influenced by the re-emission components from the characteristic DP places, while emission from the primary component places is almost constant in level and phase over frequency.

IV. DISCUSSION

Similar to recent experimental and theoretical studies on DPOAE fine structure by other authors (e.g., Mauermann *et al.*, 1997a; Heitmann *et al.*, 1998; Talmadge *et al.*, 1998, 1999), our experiments and simulations give further evidence that the fine structure is the result of two sources. Furthermore, the idea is supported that the underlying physical mechanisms of these two sources are different or at least act in a different way (e.g., Shera and Guinan, 1999).

To illustrate this, the results presented in this paper will be interpreted in three steps. The results of experiment 1 show that the DPOAE fine structure is not caused by local mechanical properties of the primary region, but rather that the characteristic site of the distortion product frequency plays a crucial role. Experiment 2 shows that independent of the distortion product order, fine-structure patterns are very similar as long as the characteristic sites of f_2 and of the DP frequencies are the same, i.e., the relative phase between the two emission sites is almost constant in this experiment. Taken together with the findings from experiment 1, this implies that the emission recorded in the ear canal is the vector sum of components from these two sites. The relative phase of these components has at least some influence on the DPOAE fine-structure pattern.

The results of experiment 3 show a quasiperiodic variation when the re-emission site is varied in frequency mono-

tonically while the generation site is held fixed, whereas the quasiperiodic variation in DP level disappears in the case of a fixed re-emission site and a sweep of the f_2 frequency. In terms of a vector summation of two components, this indicates that the component generated at the primary place must be almost constant in level and phase (at least locally) regardless of the frequency. To explain the quasiperiodic variations in the sum of the two components, we have to assume that the re-emission component varies either in phase or in level with increasing frequency. Shera and Zweig (1993) showed that the fine structure of stimulus frequency otoacoustic emissions (SFOAE) can be interpreted as interference of the incoming and outgoing traveling waves with a periodic rotating phase of the cochlear reflectance. It appears reasonable to treat the DPOAE re-emission component in a similar way to SFOAE generation. Therefore, a periodically varying phase of the re-emission component is the most likely explanation for the DPOAE fine structure. Following, furthermore, the arguments of the ‘‘Gedankenexperiment’’ described by Shera and Guinan (1998), the different characteristics of the two components (rotating phase of the f_{DP} component, almost constant phase of the f_2 component) indicate that the underlying mechanisms for these two DPOAE components are different. These authors (Shera and Guinan, 1998, 1999) distinguished two classes of OAE mechanisms, ‘‘linear coherent reflection’’ and ‘‘nonlinear distortion,’’ whereby DPOAE are a combination of both a nonlinear distortion at the generation site and a coherent reflection from the characteristic site of f_{DP} .

The disappearance of fine structure during the fixed f_{DP} experiment shows that there is no coherent reflection from the primary region because there is no rotating phase with frequency. Using the approach suggested by Zweig and Shera (1995) to explain the spectral periodicity of reflection emissions, this had to be expected because the traveling wave from the initial generation site results in constant wavelength only around the f_{DP} site but not in the region around the primaries. Therefore, the contribution from the primary region needs to be generated in a different way,

most probably due to nonlinear distortion. This interpretation of the experiments is confirmed by the computer simulations, which show no effect on the fine structure when removing the roughness (which is necessary for coherent reflections) from the primary region, while the fine structure disappears in simulations when removing the roughness only around the f_{DP} region (see Fig. 8).

The overall good correspondence between simulations and experimental results gives further confirmation for a whole class of two-source interference models like the one used here. This class of models was recently described in detail by Talmadge *et al.* (1998). The simulations with partly removed roughness (see Fig. 8) cannot directly be transformed into an experimental approach because it is impossible to “flatten out” certain areas of the human cochlea. But, a situation close to that might be studied. If a local area of the cochlea is damaged, most probably no broad and tall excitation pattern can be built up there, which in addition to roughness is necessary for coherent reflections (Zweig and Shera, 1995; Talmadge *et al.*, 1999). To get further insight into the mechanisms of DPOAE fine structure, this approach is investigated in the accompanying paper (Mauermann *et al.*, 1999) by looking at the DPOAE fine structure of subjects with frequency-specific hearing losses.

V. CONCLUSIONS

Distortion products recorded in the ear canal cannot be traced back to one single source on the basilar membrane. Instead, DPOAE fine structure reflects the interaction of two components with different underlying physical principles. The first component is due to nonlinear distortion at the primary site close to f_2 and has a nearly constant phase and level. The second component is caused by a coherent reflection from the re-emission site at the characteristic place of f_{DP} and shows a periodically varying amplitude or phase when changing the primary frequencies.

ACKNOWLEDGMENTS

This study was supported by Deutsche Forschungsgemeinschaft, DFG Ko 942/11-2. Many fruitful discussions with G. Long, C. Talmadge, and H. Duifhuis, as well as helpful comments by B. Moore on an earlier version of the manuscript, are gratefully acknowledged.

¹The number of sections had to be increased from the original 400 to at least 600 in order to avoid “wiggles” in the excitation patterns. These “wiggles” were also found by Talmadge and Tubis in their work on a time-domain model involving the “Zweig-impedance” and made them use a spatial discretisation of 4000 sections (Talmadge and Tubis, 1993; and personal communication).

²Of course, other terms could also contain nonlinearity. For example, Furst and Goldstein (1982) argue that the stiffness term should be made nonlinear. For reasons of simplicity only $d(x, v)$ and $c(v)$ were made nonlinear here, since these two terms must certainly change with input level.

³There is another discrepancy between the experimental results and the simulations in the overall shape of the patterns. For the range of frequency ratios observed here, we see a maximum in DPOAE level for human subjects at a frequency ratio around 1.225 (Gaskill and Brown, 1990) while the level is reduced for smaller and larger frequency ratios. This reflects the so-called “second filter” effect (e.g., Brown and Williams, 1993; Allen and Fahey, 1993), which is currently not included in our simulations. With the parameter settings used in this study, the second filter effect produced

by the model does not resemble the shapes found in the experimental data. However, as described in van Hengel and Duifhuis (1999), the second filter behavior can also be simulated using this kind of transmission-line model. Therefore, in the near future attempts will be made to improve the model results by finding parameter values fitting both fine structure and the second filter.

- Allen, J., and Fahey, P. (1993). “A second cochlear-frequency map that correlates distortion product and neural tuning measurements,” *J. Acoust. Soc. Am.* **94**, 809–816.
- 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.
- Brown, A., and Williams, D. (1993). “A second filter in the cochlea” in *Biophysics of Hair Cell Sensory Systems*, edited by J. H. Duifhuis, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 72–77.
- de Boer, E. (1995). “The ‘inverse problem’ solved for a three-dimensional model of the cochlea. I. Analysis,” *J. Acoust. Soc. Am.* **98**, 896–903.
- Duifhuis, H., Hoogstraten, H. W., van Netten, S. M., Diependaal, R. J., and Bialek, W. (1985). “Modelling the cochlear partition with coupled van der Pol-Oscillators,” in *Peripheral Auditory Mechanisms*, edited by J. B. Allen, J. L. Hall, A. E. Hubbard, S. T. Neely, and A. Tubis, Lecture Notes in Biomathematics 64 (Springer, Berlin), pp. 290–297.
- Furst, M., and Goldstein, J. L. (1982). “A cochlear nonlinear transmission-line model compatible with combination tone psychophysics,” *J. Acoust. Soc. Am.* **72**, 717–726.
- Gaskill, S. A., and Brown, A. M. (1990). “The behaviour of the acoustic distortion product, $2f_1-f_2$, from the human ear and its relation to auditory sensitivity,” *J. Acoust. Soc. Am.* **88**, 821–839.
- Gaskill, S. A., and Brown, A. M. (1996). “Suppression of human acoustic distortion product: dual origin of $2f_1-f_2$,” *J. Acoust. Soc. Am.* **100**, 3268–3274.
- Greenwood, D. D. (1991). “Critical bandwidth and consonance in relation to cochlear frequency-position coordinates,” *Hear. Res.* **54**, 164–208.
- Harris, F. P., Probst, R., and Xu, L. (1992). “Suppression of the $2f_1-f_2$ otoacoustic emission in humans,” *Hear. Res.* **64**, 133–141.
- He, N., and Schmiedt, R. A. (1993). “Fine structure of the $2f_1-f_2$ acoustic distortion product: changes with primary levels,” *J. Acoust. Soc. Am.* **94**, 2659–2669.
- He, N., and Schmiedt, R. A. (1997). “Fine structure of the $2f_1-f_2$ acoustic distortion product: effects of primary level and frequency ratios,” *J. Acoust. Soc. Am.* **101**, 3554–3565.
- 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.
- 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.
- Long, G. R., and Talmadge, C. L. (1997). “Spontaneous otoacoustic emission frequency is modulated by heartbeat,” *J. Acoust. Soc. Am.* **102**, 2831–2848.
- Mauermann, M., Uppenkamp, S., and Kollmeier, B. (1997a). “Zusammenhang zwischen unterschiedlichen otoakustischen Emissionen und deren Relation zur Ruhehörschwelle,” in *Fortschritte der Akustik—DAGA 97*, edited by P. Wille (DEGA e.V., Oldenburg), pp. 242–243.
- Mauermann, M., Uppenkamp, S., and Kollmeier, B. (1997b). “Periodizität und Pegelabhängigkeit der spektralen Feinstruktur von Verzerrungsprodukt-Emissionen [Periodicity and dependence on level of the distortion product otoacoustic emission spectral fine-structure],” *Audiol. Akustik* **36**, 92–104.
- Mauermann, M., Uppenkamp, S., van Hengel, P., and Kollmeier, B. (1999). “Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. II. Fine structure for different shapes of cochlear hearing loss,” *J. Acoust. Soc. Am.* **106**, 3484–3491.
- Ruggero, M. A., and Rich, N. C. (1991). “Application of a commercially-manufactured Doppler-shift laser velocimeter to the measurement of basilar-membrane vibration,” *Hear. Res.* **51**, 215–230.
- Shera, C. A., and Zweig, G. (1993). “Order from chaos: resolving the paradox of periodicity in evoked otoacoustic emissions,” 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. 54–63.

- Shera, C. A., and Guinan, J. J. (1998). "Reflection emissions and distortion products arise from fundamentally different mechanisms," in Abstracts of the 21st ARO Midwinter Research Meeting, Abstract No. 344, p. 86.
- 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.
- Smurzynski, J., Leonard, G., Kim, D. O., Lafreniere, D. C., and Jung, M. D. (1990). "Distortion product otoacoustic emissions in normal and impaired adult ears," *Arch. Otolaryngol. Head Neck Surg.* **116**, 1309–1316.
- Sun, X., Schmiedt, R. A., He, N., and Lam, C. F. (1994a). "Modeling the fine structure of the $2f_1-f_2$ acoustic distortion product. I. Model development," *J. Acoust. Soc. Am.* **96**, 2166–2174.
- Sun, X., Schmiedt, R. A., He, N., and Lam, C. F. (1994b). "Modeling the fine structure of the $2f_1-f_2$ acoustic distortion product. II. Model evaluation," *J. Acoust. Soc. Am.* **96**, 2175–2183.
- Talmadge, C. L., and Tubis, A. (1993). "On modeling the connection between spontaneous and evoked otoacoustic emissions," 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. 25–32.
- 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.
- 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.
- Uppenkamp, S., and Mauermann, M. (1999). "Effect of frequency ratio f_2/f_1 on spectral fine structure of distortion product otoacoustic emissions," *Br. J. Audiol.* **33**, 87–88(A).
- van Hengel, P. (1993). Comment to paper presented by Shera and Zweig (1993), 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), p. 62.
- van Hengel, P. W. J., Duifhuis, H., and van den Raadt, M. P. M. G. (1996). "Spatial periodicity in the cochlea: the result of interaction of spontaneous emissions?" *J. Acoust. Soc. Am.* **99**, 3566–3571.
- van Hengel, P. W. J., and Duifhuis, H. (1999). "The generation of distortion products in a nonlinear transmission line model of the cochlea," Proceedings of the Symposium on Recent Developments in Auditory Mechanics, 25–30 July 1999, Sendai, Japan.
- van den Raadt, M. P. M. G., and Duifhuis, H. (1990). "A generalized Van der Pol-oscillator cochlea model," in *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele, Lecture Notes in Biomathematics 87 (Springer, Berlin), pp. 227–234.
- van den Raadt, M. P. M. G., and Duifhuis, H. (1993). "Different boundary conditions in a one-dimensional time domain model," 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), p. 103.
- Zweig, G. (1991). "Finding the impedance of the organ of Corti," *J. Acoust. Soc. Am.* **89**, 1229–1254.
- Zweig, G., and Shera, C. A. (1995). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," *J. Acoust. Soc. Am.* **98**, 2018–2047.

Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. II. Fine structure for different shapes of cochlear hearing loss^{a)}

Manfred Mauermann^{b)} and Stefan Uppenkamp^{c)}

AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany

Peter W. J. van Hengel

BCN, Rijksuniversiteit Groningen, 9747 AG Groningen, The Netherlands

Birger Kollmeier

AG Medizinische Physik, Universität Oldenburg, D-26111 Oldenburg, Germany

(Received 10 January 1999; accepted for publication 24 August 1999)

Distortion product otoacoustic emissions (DPOAE) were recorded from eight human subjects with mild to moderate cochlear hearing loss, using a frequency spacing of 48 primary pairs per octave and at a level $L_1=L_2=60$ dB SPL and with a fixed ratio f_2/f_1 . Subjects with different shapes of hearing thresholds were selected. They included subjects with near-normal hearing within only a limited frequency range, subjects with a notch in the audiogram, and subjects with a mild to moderate high-frequency loss. If the primaries were located in a region of normal or near-normal hearing, but DP frequencies were located in a region of raised thresholds, the distortion product $2f_1-f_2$ was still observable, but the DP fine structure disappeared. If the DP frequencies fell into a region of normal thresholds, fine structure was preserved as long as DPOAE were generated, even in cases of mild hearing loss in the region of the primaries. These experimental results give further strong evidence that, in addition to the initial source in the primary region, there is a second source at the characteristic place of f_{DP} . Simulations in a nonlinear and active computer model for DPOAE generation indicate different generation mechanisms for the two components. The disappearance of DPOAE fine structure might serve as a more sensitive indicator of hearing impairment than the consideration of DP level alone. © 1999 Acoustical Society of America. [S0001-4966(99)02912-4]

PACS numbers: 43.64.Jb, 43.64.Kc [BLM]

INTRODUCTION

The recording of distortion product otoacoustic emissions (DPOAE) is claimed to be useful as an objective audiometric test with a high-frequency selectivity by various clinical studies. In many papers, the reported correlation of audiometric thresholds and DPOAE levels is mainly based on large databases of many subjects (e.g., Nelson and Kimberley, 1992; Gorga *et al.*, 1993; Moulin *et al.*, 1994; Suckfüll *et al.*, 1996). The prediction of individual thresholds based on DPOAE requires a detailed and comprehensive dataset from each individual subject, including growth functions, at multiple stimulus frequencies (Kummer *et al.*, 1998). However, for extensive use as a diagnostic tool, a more detailed understanding of the DPOAE generation mechanisms is still required.

In agreement with theoretical and experimental work reported by other groups (Brown *et al.*, 1996; Gaskill and Brown, 1996; Heitmann *et al.*, 1998; Talmadge *et al.*, 1998,

1999), the experimental results from normal-hearing subjects in the accompanying paper (Mauermann *et al.*, 1999) showed that DPOAE should be interpreted as the vector sum of two sources, one at the initial generation site due to nonlinear distortion close to the f_2 place, the other at the characteristic site of the particular DP frequency of interest. The results from simulations using a nonlinear and active model of the cochlea presented in Mauermann *et al.* (1999) showed that the component from the f_{DP} site is sensitive to the existence of statistical fluctuations in the mechanical properties along the cochlea partition, i.e., roughness, while the initial generation component is not. From the model point of view, this indicates different underlying mechanisms for the generation of the two DPOAE components. However, removing the roughness from certain areas along the cochlear partition—as shown in the computer simulations in Mauermann *et al.* (1999)—cannot directly be transformed into a controlled experiment with human subjects.

Other studies on modeling OAE fine structure (Zweig and Spera, 1995; Talmadge *et al.*, 1999) showed that—in addition to the roughness—the model needs another feature to produce DPOAE fine structure: broad and tall excitation patterns have to be generated to allow coherent reflections. The generation of a broad and tall excitation pattern requires an active feedback mechanism in the model. In the real co-

^{a)}Parts of this study were presented at the 21st Midwinter Research Meeting of the Association for Research in Otolaryngology, 1998 in St. Petersburg Beach, FL [Mauermann *et al.*, Abstract No. 595, p. 149].

^{b)}Author to whom correspondence should be addressed. Electronic mail: mani@medi.physik.uni-oldenburg.de

^{c)}Current address: CNBH, Department of Physiology, University of Cambridge, Downing Street, Cambridge CB2 3EG, U.K.

chlea, this mechanism is most probably related to the motility of outer hair cells (OHC), as found by Brownell *et al.* (1985) and Zenner *et al.* (1985). If the activity of OHC in the cochlea is reduced because of damage to certain areas, most probably no broad and tall excitation pattern can build up there. As a consequence, there would be no coherent reflection from the re-emission site and no DPOAE fine structure would be observable.

This assumption is mainly based on the model results obtained so far (Mauermann *et al.*, 1999), but can be tested in carefully selected hearing-impaired human subjects. In the present paper, results on DPOAE fine structure from subjects with different audiogram shapes will be presented to investigate the effects of damage in different regions of the cochlea in more detail. Our subjects included persons with near-normal hearing only within a limited frequency range (“bandpass listeners”), a notch in the audiogram (“bandstop listeners”), or a hearing loss at high frequencies only (“low-pass listeners”). This allowed measurement of DPOAE while restricting either f_{DP} or the primaries to “normal” or “near-normal” BM regions. It was expected that only the component generated in a region of cochlear damage would be reduced. The experiments were designed to obtain further evidence for the two-source model as discussed in the accompanying paper (Mauermann *et al.*, 1999). To support the arguments, a “hearing-impaired” version of the computer model was also tested, simulating the bandstop listener situation mentioned above.

I. METHODS

A. Subjects

Eight subjects with different types of hearing loss participated in the experiments. They were selected because of the shapes of their audiograms. Subjects HE and HA (63 and 59 years old, respectively) showed hearing loss with a bandpass characteristic, i.e., near-normal threshold within only a small frequency band at 1.5 kHz with raised thresholds for frequencies above and below. The second group of subjects (DI, FM, MM, 29–50 years old) showed a notch in the hearing threshold of about 40 dB centered at 4 kHz. The third group (RH, HL, JK, 59–63 years old) showed a moderate high-frequency hearing loss. All subjects except MM had a stable audiogram for at least 6 months. The notch in the audiogram of MM was caused by a mild sudden hearing loss. The threshold recovered almost completely over a period of 6 months.

B. Instrumentation and experimental procedures

All instrumentation and the experimental methods for recording DPOAE fine structure are described in detail in the accompanying paper (Mauermann *et al.*, 1999). In short, an insert ear probe, type ER-10C, in combination with a signal-processing board Ariel DSP-32C, was used to record DPOAE. All stimuli were generated digitally at a sampling rate of 22.05 kHz and as harmonics of the inverse of the frame length (4096 samples, i.e., harmonics of 5.38 Hz). They were played continuously to the subjects via 16-bit D/A converters, a computer-controlled audiometer, and low-

pass filters at a presentation level of $L_1 = L_2 = 60$ dB SPL. An automatic in-the-ear calibration was performed before each run to adjust the primaries to the desired sound-pressure levels. In most subjects, a frequency ratio of $f_2/f_1 = 1.2$ was used. This was increased to $f_2/f_1 = 1.25$ in some cases to achieve a larger separation of f_{DP} and the primary frequencies.

DPOAE were recorded at a high-frequency resolution of 32 frequencies per octave. The microphone output was amplified, A/D converted and averaging in the time domain of at least 16, and if necessary up to 256 repeated frames were performed for each pair of primaries to increase the signal-to-noise ratio. Again, all results were plotted as a function of f_2 and as a function of f_{DP} . This permits us to relate the data both to the initial generation site near f_2 and to the presumed second source located around the characteristic site of f_{DP} .

In addition to the clinical audiogram, absolute thresholds were measured for five of the eight subjects with an adaptive three-alternative forced-choice (3AFC) two-step method to get a more accurate estimate of the shape of the audiogram. The absolute threshold was measured at a resolution of eight frequencies per octave in the transition regions between near-normal hearing and increased threshold. Sinusoids of 375-ms duration including 45 ms Hanning-shaped ramps at the start and end were used as stimuli. They were played through one of the speakers of the ER10C probe. After threshold detection, the sound pressure in the ear canal was measured with the ER10C probe microphone for a fixed attenuation of the audiometer. The sound-pressure level in dB SPL at threshold was computed from the difference of this attenuator value and the attenuation at hearing threshold.

II. RESULTS

Figures 1–4 show the DPOAE fine-structure patterns and absolute thresholds for eight subjects. In each figure, the left column gives the clinical audiogram. The middle column shows DPOAE results (black line) and absolute threshold (gray line) as a function of f_{DP} to permit a direct comparison around the assumed re-emission site. The right column shows the same data as a function of f_2 to allow a direct comparison of absolute threshold and DPOAE fine structure at the initial generation site close to f_2 . For the middle and right columns, the left ordinate holds for the DPOAE level in dB SPL, while the right ordinate is for the absolute threshold. This threshold was measured in dB SPL for the subjects who did the adaptive procedure (HE, HA, DI, HL, JK), while for the other subjects (MM, FM, RH) the threshold in dB HL from the clinical audiogram (left column) is given. In addition to the threshold data (or audiogram data), the black bars on the top of the middle and right column panels give a sketch of the area with thresholds of 20 dB or better, i.e., frequency ranges with normal or near-normal hearing (as taken from the audiograms).

Figure 1 shows results for two subjects with a near-normal threshold in only a limited frequency band. A reduced DPOAE fine structure with level fluctuations smaller than 5 dB (areas indicated by the gray bar on the top of the

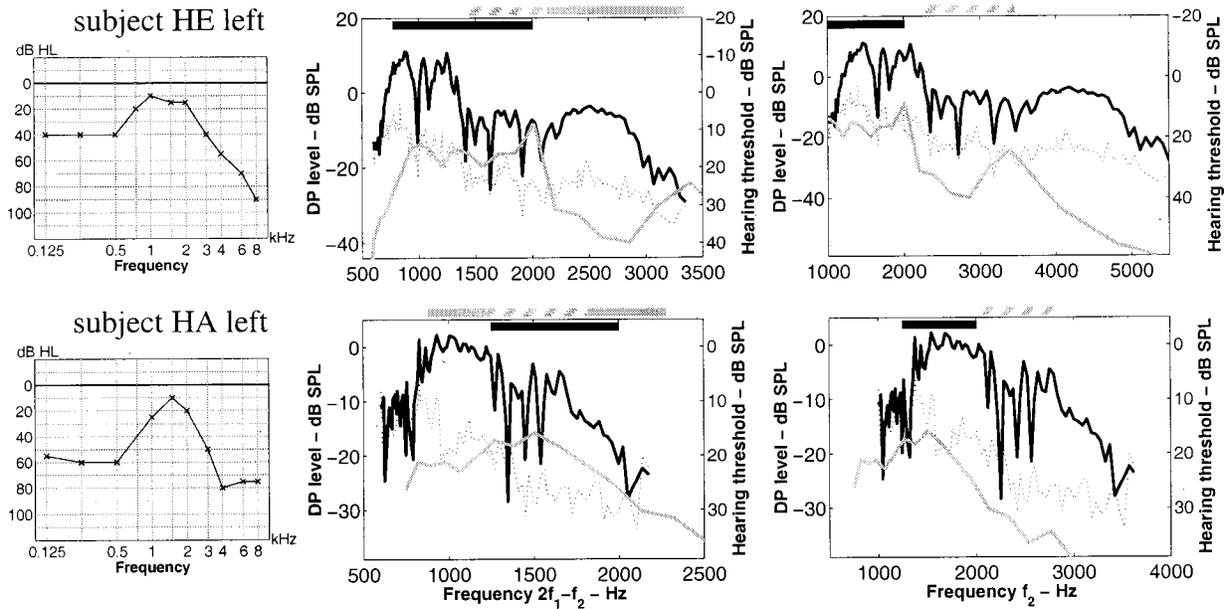


FIG. 1. DPOAE fine structure from two ears with near-normal threshold within a bandlimited frequency region only. Left: clinical audiogram. Middle: absolute threshold (gray line, right ordinate) and DP level as a function of $2f_1-f_2$ (black line, left ordinate). Right: as middle column, but DP level as a function of f_2 . The dotted line represents the noise floor during the DPOAE recording. Frequency regions with near-normal hearing (20 dB HL or better) are indicated by black bars on the top of the plots. Regions with reduced fine structure (less than 5-dB level fluctuations) are indicated by gray bars. Hatched gray bars mark areas with reduced DPOAE level but still pronounced fine structure. Top row: subject HE, left ear. Bottom row: subject HA, left ear. Frequency ratio of the primaries: $f_2/f_1=1.25$; levels of the primaries: $L_1=L_2=60$ dB SPL.

middle column plots) can be observed when f_{DP} falls into a region of raised threshold. On the other hand, when the distortion products are at frequencies with near-normal hearing and the primaries at frequencies of a moderate hearing loss, the DPOAE level is reduced, but a preserved fine structure with level fluctuations of 5 dB or more can still be observed in a certain area (indicated by the hatched gray bars on the top of the plots).

Figure 2 shows the results from four ears of three subjects with notches in their audiograms. As was the case for the data shown in Fig. 1, when the distortion product frequencies fall into the area of hearing loss, the fine structure disappears but nearly no reduction in DPOAE level occurs. If the initial DPOAE generation site, i.e., the area around f_2 , falls into the region of hearing loss while the related f_{DP} frequency covers a region of near-normal hearing, the DPOAE level is reduced but a pronounced fine structure is still observed. If both f_2 and f_{DP} fall into a region of mild to moderate hearing loss, DPOAE level and DPOAE fine structure are reduced. Subject FM (bottom row in Fig. 2) showed a reduced fine structure in a limited frequency band above 4 kHz only. This might be due to a narrow-band notch in the absolute threshold in this frequency region which was not resolved using the clinical audiogram.

Figure 3 shows DPOAE fine-structure patterns for two subjects with a moderate hearing loss only at high frequencies. When plotted as a function of f_2 (right column), a DPOAE level similar to the level in regions with near-normal hearing could still be observed in the region of raised threshold, while the fine structure disappeared as soon as the distortion product frequencies fell into the region of hearing

loss (cf. middle column for plot as a function of $2f_1-f_2$). These cases suggest that the DPOAE fine structure might provide a more sensitive indicator of cochlear damage than the DPOAE level, which is mainly related to the initial generation site close to f_2 .

Figure 4 gives one example of a different effect of hearing loss on DPOAE. This subject also had a high-frequency hearing loss, but the level of DPOAE decreased with the high-frequency hearing loss, while the fine structure was preserved as long as DPOAE are recordable, i.e., below $f_2=4$ kHz. As can be seen in the audiogram, thresholds were normal in the region below 2.5 kHz, i.e., in the region of the related f_{DP} frequencies. Therefore, an effect on DPOAE fine structure is not expected in this particular case of a steep high-frequency hearing loss.

Figure 5 shows again the DPOAE fine structure for subject MM (cf. Fig. 2), this time at three different stages of recovery from the mild sudden hearing loss. After 4 months (dark gray line), the clinical audiogram was almost normal and the DPOAE level had returned to normal (notice the recovery from the notch in the middle trace between 2000 and 3000 Hz for $2f_1-f_2$, corresponding to a f_2 range of 3000–4000 Hz, cf. Fig. 2). A reappearance of fine structure over the whole range could only be observed after 6 months. It is possible that a slight cochlear disorder still affected the DPOAE fine structure after 4 months, while the DPOAE level and the audiogram had almost completely recovered. Again, the fine structure appears to be a more sensitive indicator for local cochlear damage than the consideration of overall DPOAE level alone, revealing even slight disorders not detectable in the clinical audiogram.

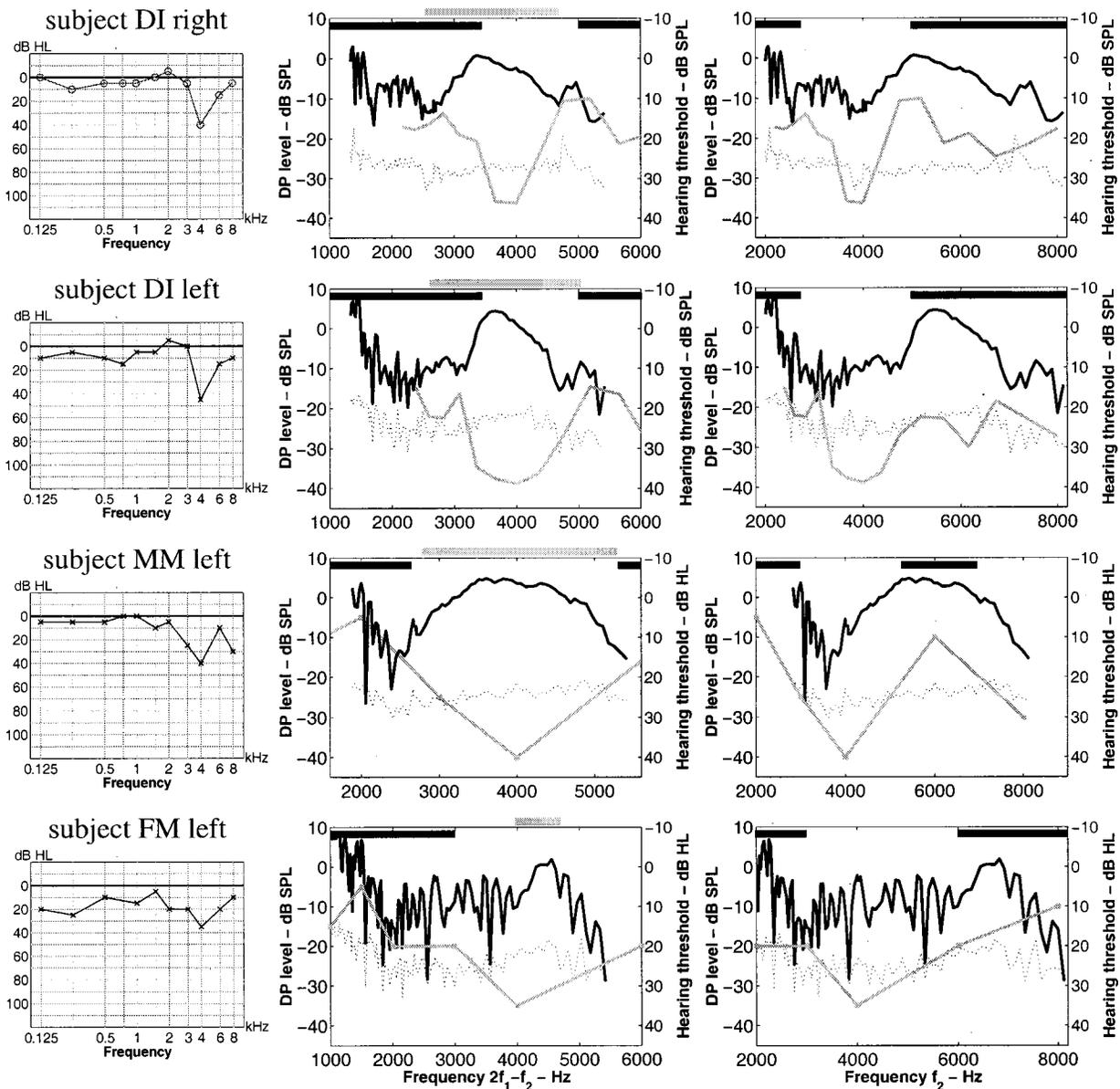


FIG. 2. As Fig. 1, but showing results from four ears with a notch in the audiogram. Top row: subject DI, right ear. Second row: subject DI, left ear. Third row: subject MM, left ear. Fourth row: subject FM, left ear. The notch in the audiogram of subject MM was caused by a mild sudden hearing loss. MM and FM did not perform the adaptive procedure for evaluation of absolute thresholds. Therefore, the threshold curves are taken from the clinical audiograms in these cases. Frequency ratio of the primaries: $f_2/f_1=1.2$; levels of the primaries: $L_1=L_2=60$ dB SPL.

III. SIMULATIONS IN A NONLINEAR AND ACTIVE MODEL OF THE COCHLEA

As shown in the accompanying paper (Mauermann *et al.*, 1999), the behavior of DPOAE fine-structure patterns in different experimental paradigms can be well simulated with a nonlinear and active transmission line model of the cochlea. Here, a “hearing-impaired” version of this computer model is tested to investigate the effects of local changes of the damping function, as an analog of hearing loss in a restricted frequency region. This can be achieved by looking at a partly “passive” cochlea. The biggest hearing loss that can be modeled by making the damping independent of velocity and fixing it to the maximum value that occurs in the “normal-hearing” model is about 40 dB. The value of 40 dB corresponds to the assumed gain of cochlea activity probably due to motility of the OHC (discussed in

Pickles, 1988; Hoth and Lenarz, 1993). Therefore, thresholds distinctively higher than 40 dB probably have to be related to damage of the inner hair cells (IHC), which cannot be simulated in this kind of macromechanical model.

Figure 6 shows one example of a fine-structure pattern, calculated using a model cochlea with raised threshold around 3.7 kHz. The hearing loss in this case was introduced by taking a relatively high positive damping, independent of velocity, for the segments representing the frequencies from 3.6 to 3.8 kHz in the model cochlea (segments 210 to 219 of a total of 600 segments) with smooth transitions (over 3.3 to 3.6 kHz and 3.8 to 4.5 kHz) (see the Appendix). The “delayed feedback stiffness” is smoothly decreased to 0 in this region in proportion to the increased damping (for a more detailed description of the model, see Mauermann *et al.*, 1999). As found in the experiments, the fine structure in the

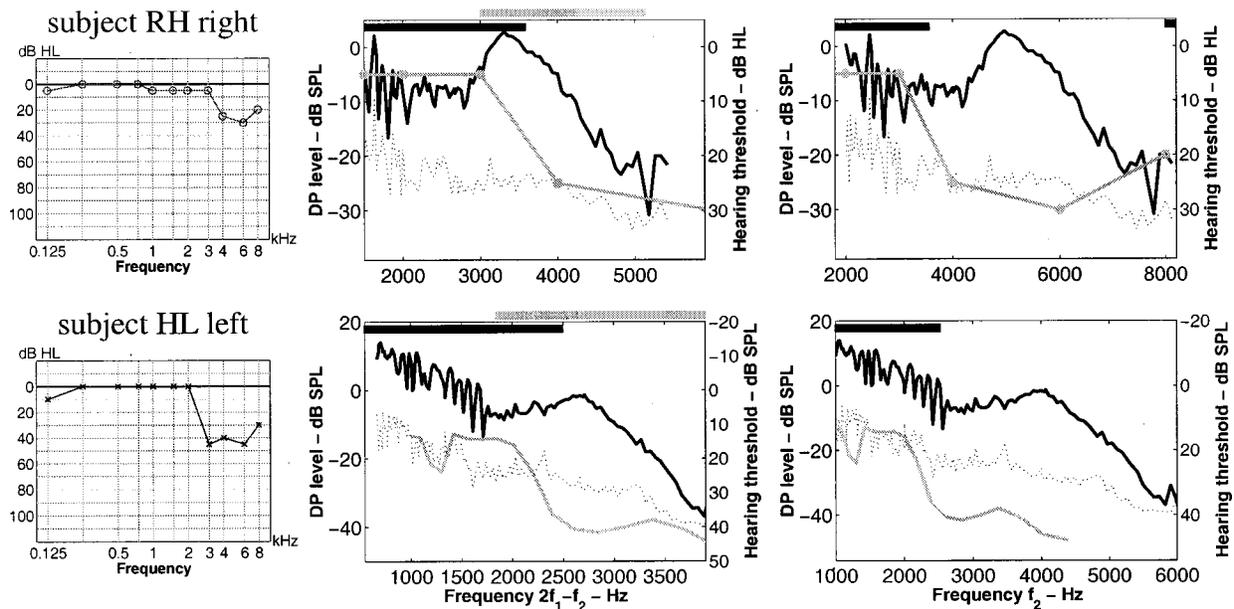


FIG. 3. As Fig. 2, but showing results from two subjects with high-frequency hearing loss. Top row: subject RH, right ear. Bottom row: subject HL, left ear. Absolute threshold was taken from the clinical audiogram for subject RH.

computer simulation disappears as soon as f_{DP} falls into the region of increased damping, i.e., increased threshold, while a reduction of DPOAE level is observable when the primaries cover the region of "hearing loss." The hearing loss was quantified by finding the stimulus level required to give the same excitation as a stimulus level of 0 dB SPL in the normal-hearing version of the model.

IV. DISCUSSION

Most previous studies on DPOAE from hearing-impaired subjects related DPOAE levels to hearing thresholds based only on a comparatively low number of frequencies. In the only study dealing with DPOAE from hearing-impaired subjects with a high-frequency resolution (He and Schmiedt, 1996), it was concluded that a fine structure would always be observable as long as DPOAE can be recorded. Their study included 14 hearing-impaired subjects, all of whom had a more or less steep high-frequency hearing loss. This conclusion of unaffected fine structure holds only for one subject (JK, Fig. 4) from this investigation, while the others showed a substantial decrease of DPOAE fine structure when $2f_1 - f_2$ fell into a region of hearing loss. The subjects presented here were selected because of the particu-

lar shape of their audiograms with either raised thresholds or normal hearing in a limited frequency band only. This sample might not be representative for the clinical population, but it permitted measurements with either the primary frequencies or the distortion product frequencies covering a region of cochlear damage, motivated by the results and computer simulations reported in the accompanying paper (Mauermann *et al.*, 1999). Consequently, our data will be discussed in respect to the two-source model of DPOAE generation. The experiments were not intended as a representative clinical study. Nevertheless, the results might still improve the value of DPOAE as a diagnostic tool.

The initial generation of DPOAE is due to nonlinear distortion at the primary site close to f_2 . The cases reported in Figs. 1–3 exhibit a coincidence of the disappearance of DPOAE fine structure with damage at the characteristic place of f_{DP} . This strongly supports the interpretation of the data in terms of a two-source model of DPOAE generation, as discussed in the accompanying paper for normal-hearing subjects (Mauermann *et al.*, 1999). If only the component generated in the primary region contributes to the emission measured in the ear canal, no fine structure can be observed. When there is also a contribution from the re-emission site at

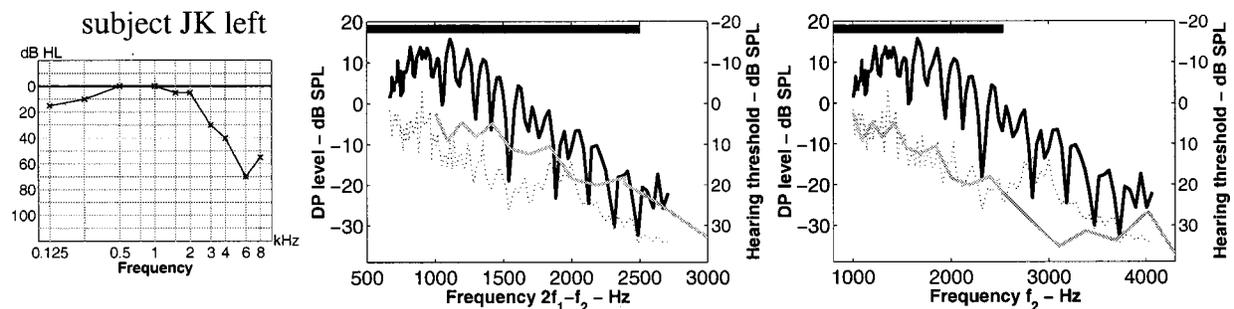


FIG. 4. As Fig. 1, but for subject JK, left ear. This subject also shows a high-frequency hearing loss. In contrast to the data shown in Fig. 3, the DPOAE level decreases with increasing threshold but the fine structure remains unaffected, as long as DPOAE are recordable at all.

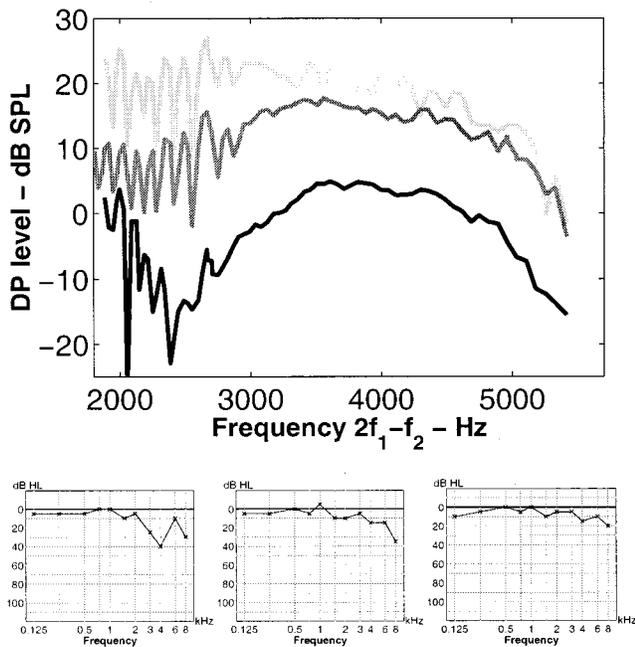


FIG. 5. Recovery of DPOAE fine structure after a mild sudden hearing loss, subject MM, left ear (cf. Fig. 2). Top panel: DPOAE level is plotted as a function of $2f_1 - f_2$ on the day of sudden hearing loss (bottom trace), 4 months later (middle trace), and 6 months later (top trace, complete recovery of absolute threshold). The ordinate holds for the bottom trace only, the other traces are shifted by +10 and +20 dB, respectively. Bottom row, from left to right: the clinical audiograms at the day of sudden hearing loss, 4 months later, and 6 months later.

f_{DP} , a quasiperiodic fine structure is observable, i.e., DPOAE can be treated as the vector sum of two different components, as suggested by Brown *et al.* (1996).

Heitmann *et al.* (1998) showed that the presentation of an additional suppressor tone close to f_{DP} (25 Hz above f_{DP}) causes a disappearance of fine structure due to suppression of the component from the f_{DP} place. As demonstrated here, damage in the DP frequency region has a similar effect to the suppressor. While the addition of a third tone could cause unwanted side effects when investigating the DP generation mechanisms, such as additional distortion products (Harris *et al.*, 1992), the experiments presented here take advantage of the “naturally” reduced cochlear activity.

As already reported in previous studies (e.g., Schlögel *et al.*, 1995), it is not uncommon for subjects with high-frequency hearing loss to have DPOAE not substantially different in level from those from normal-hearing subjects, even when the primaries fall into the region of a hearing loss of 30 dB or more. Two more examples for this are the subjects RH and HL in this study (cf. Fig. 3). However, DPOAE fine structure disappeared in both subjects as soon as f_{DP} fell into a region of even slightly raised thresholds (i.e., above 2.5 kHz for subject HL, above 4 kHz for subject RH). This suggests that moderate cochlear damage can already influence the re-emission component of the DPOAE. A similar interpretation holds for the DPOAE fine structure during recovery from a mild sudden hearing loss shown in Fig. 5. The reduced fine structure after 4 months still might indicate some slight damage, although the threshold—as measured in the clinical audiogram with its accuracy of about ± 5 dB and its

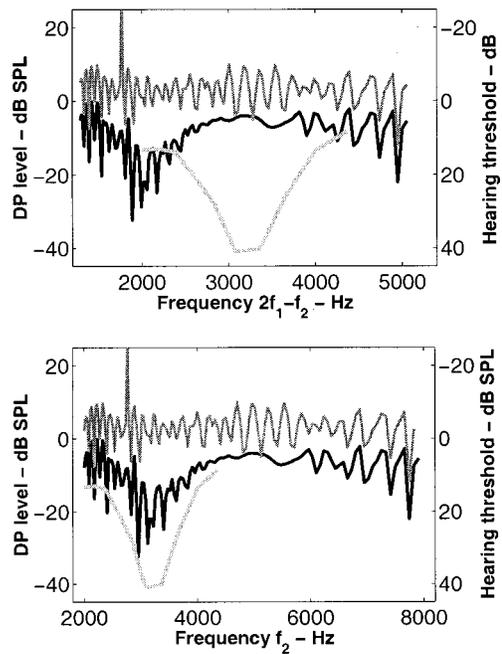


FIG. 6. Simulation of DPOAE fine structure for a frequency ratio $f_2/f_1 = 1.22$ in a version of the transmission-line model with a notch in the audiogram in the region between 3.3 and 4.5 kHz vs the simulated DPOAE fine structure of the normal-hearing model. Black line: DPOAE fine structure of the hearing-impaired model at a frequency resolution of 64 steps per octave. Gray line: DPOAE fine structure of the normal-hearing model (shifted 10 dB up). Light gray line: simulated hearing threshold of the hearing impaired model at eight frequencies per octave.

limited frequency resolution—had already recovered. This high vulnerability to cochlear damage of the DPOAE fine structure has also been noticed in other studies. Engdahl and Kemp (1996) showed that noise exposure causing a temporary threshold shift results in a temporary disappearance of DPOAE fine structure. Furthermore, the fine structure gets reduced during aspirin consumption before an overall DPOAE level reduction can be observed (Rao *et al.*, 1996; Long, 1999). Overall, it appears as if the consideration of fine structure can serve as a more sensitive tool for the detection of slight cochlear damage in certain cases than the DPOAE level alone. However, a prospective clinical study involving more subjects would be required to evaluate this possible application.

In contrast to the subjects discussed so far, the DPOAE from subject JK (Fig. 4) appears to behave differently: DPOAE level decreases with increasing hearing loss while the fine structure is unaffected. This behavior is more in line with the data reported by He and Schmiedt (1996). The seeming contradiction to our other data can be interpreted as follows: When the primary at f_2 falls into a region of a distinct hearing loss, i.e., above 4 kHz for this subject, no measurable initial DPOAE component is generated. Consequently, no reflection component from the re-emission site close to f_{DP} can be recorded. However, when f_2 is below 4 kHz, the corresponding f_{DP} frequency still falls into a region of normal or near-normal hearing for this subject (i.e., threshold of 20 dB HL or better), which is sufficient to create the re-emission component due to coherent reflection. The interaction of the two components generates the fine struc-

ture. A similar explanation would also hold for most of the subjects described by He and Schmiedt (1996), who had steep high-frequency hearing losses.

The effects of frequency-specific hearing loss on DPOAE fine structure can be simulated in a realistic way using the hearing-impaired version of the computer model of the cochlea (cf. Figs. 2 and 6). The very good correspondence between data and simulations even in the case of hearing impairment gives further support for a whole class of similar cochlea models recently described in Talmadge *et al.* (1999). In the model, the generation mechanisms for the two sources are different. The re-emission component can be interpreted as a coherent reflection sensitive (1) to the presence of ‘‘roughness,’’ and (2) to the presence of broad and tall excitation patterns (e.g., Talmadge *et al.*, 1998), which are generated by an active feedback mechanism. The initial generation—which is not sensitive to the presence of roughness (see Mauermann *et al.*, 1999; Talmadge *et al.*, 1998)—is not connected to coherent reflection but should be interpreted as a consequence of nonlinear distortion only.

This view of two different mechanisms is in agreement with the conclusions drawn by Shera and Guinan (1999). They also distinguish two DPOAE mechanisms: coherent reflection from the f_{DP} site [similar to stimulus frequency otoacoustic emissions (SFOAE)] and nonlinear distortion from the generation site near f_2 . The different effects on DPOAE fine structure and overall level caused by cochlear damage, as reported in our experiments, might reflect an experimental confirmation of two different generation mechanisms.

V. CONCLUSIONS

- (i) Distortion product emissions measured in the human ear canal are produced by two sources, one at the characteristic place of the primaries and the second at the characteristic place of f_{DP} .
- (ii) The DPOAE fine structure is mainly influenced by the local state of the cochlea at the characteristic place of f_{DP} and appears to be a more sensitive indicator of cochlear damage than DPOAE level alone.
- (iii) At least from the model point of view, the initial generation at the site close to f_2 is caused by nonlinear interaction of the primaries while the re-emission from the characteristic site of f_{DP} can be treated as a coherent reflection.
- (iv) The evaluation of fine structure could considerably improve the clinical use of DPOAE, e.g., for early identification of hearing loss or to monitor the recovery from a sudden hearing loss more accurately.

ACKNOWLEDGMENTS

This study was supported by Deutsche Forschungsgemeinschaft, DFG Ko 942/11-2. Helpful comments by Brian Moore and Glenis Long on an earlier version of the manuscript are gratefully acknowledged.

APPENDIX

The nonlinearity in the model is introduced by a nonlinear damping $d(x, v)$ and a stabilizing delayed feedback stiffness $c(v)$ (see Mauermann *et al.*, 1999).

$$d(x, v) = \left[d_l + \frac{\beta(d_h - d_l)|v|}{1 + \beta|v|} \right] \sqrt{ms(x)}, \quad (\text{A1a})$$

$$c(v) = c_l + \frac{-\beta d_l |v|}{1 + \beta|v|}, \quad (\text{A1b})$$

where v is the velocity of basilar membrane (BM) section; m is the mass (0.375 kg/m^2); x is the distance to the base of the cochlea; s is the stiffness of BM section; d_l is the damping parameter, determines the damping at low BM velocities (d_l is -0.12 in the normal-hearing model); d_h is the damping parameter, determines the damping at high BM velocities (d_h is 0.5 in the normal-hearing model); c_l is the parameter of delayed feedback stiffness (c_l is 0.1416 in the normal-hearing model); and β is the parameter to determine the shape of the nonlinear damping function.

The hearing loss was introduced by taking $d(x, v) = d_h$ in the region from 3.6 to 3.8 kHz with smooth transitions (over 3.3 to 3.6 kHz and 3.8 to 4.5 kHz). This could be done by letting $\beta \rightarrow \infty$. Because this is impossible in practice, we let $\beta(d_h - d_l)$ go to 0 and d_l go to d_h simultaneously. In correspondence to the increased damping, the delayed feedback stiffness $c(v)$ is smoothly decreased to 0 in this region.

- Brown, A. M., Harris, F. P., and Beveridge, H. (1996). ‘‘Two sources of acoustic distortion products from the human cochlea,’’ *J. Acoust. Soc. Am.* **100**, 3260–3267.
- Brownell, W., Bader, C. R., Bertrand, D., and Ribaupierre, Y. (1985). ‘‘Evoked mechanical responses of isolated cochlear outer hair cells,’’ *Science* **227**, 194–196.
- Engdahl, B., and Kemp, D. T. (1996). ‘‘The effect of noise exposure on the details of distortion product otoacoustic emissions in humans,’’ *J. Acoust. Soc. Am.* **99**, 1573–1587.
- Gaskill, S., and Brown, A. (1996). ‘‘Suppression of human acoustic distortion product: Dual origin of $2f_1 - f_2$,’’ *J. Acoust. Soc. Am.* **100**, 3268–3274.
- Gorga, M. P., Neely, S. T., Bergman, B., Beauchaine, K. L., Kaminski, J. R., Peters, J., and Jestaedt, W. (1993). ‘‘Otoacoustic emissions from normal-hearing and hearing-impaired subjects: Distortion product responses,’’ *J. Acoust. Soc. Am.* **93**, 2050–2060.
- Harris, F., Probst, R., and Xu, L. (1992). ‘‘Suppression of the $2f_1 - f_2$ otoacoustic emission in humans,’’ *Hear. Res.* **64**, 133–141.
- He, N., and Schmiedt, R. A. (1996). ‘‘Effects of aging on the fine structure of the $2f_1 - f_2$ acoustic distortion product,’’ *J. Acoust. Soc. Am.* **99**, 1002–1015.
- 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.
- Hoth, S., and Lenarz, T. (1993). *Otoakustische Emissionen—Grundlagen und Anwendung* (Thieme, Stuttgart), p. 12.
- Kummer, P., Janssen, T., and Arnold, W. (1998). ‘‘The level and growth behavior of the $2f_1 - f_2$ distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss,’’ *J. Acoust. Soc. Am.* **103**, 3431–3444.
- Long, G. R. (1999). Personal communication.
- Mauermann, M., Uppenkamp, S., van Hengel, P. W. J., and Kollmeier, B. (1999). ‘‘Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure. I. Fine structure and higher order DPOAE as a function of the frequency ratio f_2/f_1 ,’’ *J. Acoust. Soc. Am.* **106**, 3473–3483.

- Moulin, A., Bera, J. C., and Collet, L. (1994). "Distortion product otoacoustic emissions and sensorineural hearing loss," *Audiology* **33**, 305–326.
- Nelson, D. A., and Kimberley, B. P. (1992). "Distortion product emissions and auditory sensitivity in human ears with normal hearing and cochlear hearing loss," *J. Speech Hear. Res.* **35**, 1142–1159.
- Pickles, J. (1988). *An Introduction to the Physiology of Hearing*, 2nd ed. (Academic, London), p. 149.
- Rao, A., Long, G. R., Narayan, S. S., and Dhar, S. (1996). "Changes in the temporal characteristics of TEOAE and the fine structure of DPOAEs with aspirin consumption," in Abstracts of the 19th ARO Midwinter Research Meeting, Abstract, p. 27.
- 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.
- Schlögel, H., Stephan, K., Böheim, K., and Welzl-Müller, K. (1995). "Distorsionsprodukt otoakustische Emissionen bei normalem Hörvermögen und bei sensorineuraler Schwerhörigkeit," *HNO* **43**, 19–24.
- Suckfüll, M., Schneeweiss, S., Dreher, A., and Schorn, K. (1996). "Evaluation of TEOAE and DPOAE measurements for the assessment of auditory thresholds in sensorineural hearing loss," *Acta Oto-Laryngol.* **116**, 528–533.
- 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.
- 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.
- Zenner, H., Zimmermann, U., and Schmidt, U. (1985). "Reversible contraction of isolated mammalian cochlear hair cells," *Hear. Res.* **18**, 127–133.
- Zweig, G., and Shera, C. A. (1995). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," *J. Acoust. Soc. Am.* **98**, 2018–2047.

Phonemes, intensity and attention: Differential effects on the mismatch negativity (MMN)

Michael D. Szymanski,^{a)} E. William Yund,^{b)} and David L. Woods

Department of Neurology (127), VA Medical Center, Northern California Health Care System,
150 Muir Road, Martinez, California 94553

(Received 28 September 1998; revised 25 August 1999; accepted 16 September 1999)

Auditory event-related potentials (ERPs) to speech sounds were recorded in a demanding selective attention task to measure how the mismatch negativity (MMN) was affected by attention, deviant feature, and task relevance, i.e., whether the feature was target or nontarget type. With vowel-consonant-vowel (VCV) disyllables randomly presented to the right and left ears, subjects attended to the VCVs in one ear. In different conditions, the subjects responded to either intensity or phoneme deviance in the consonant. The position of the deviance within the VCV also varied, being in the first (VC), second (CV), or both (VC and CV) formant-transition regions. The MMN amplitudes were larger for deviants in the attended ear. Task relevance affected the MMNs to intensity and phoneme deviants differently. Target-type intensity deviants yielded larger MMNs than nontarget types. For phoneme deviants there was no main effect of task relevance, but there was a critical interaction with deviance position. The both position gave the largest MMN amplitudes for target-type phoneme deviants, as it did for target- and nontarget-type intensity deviants. The MMN for nontarget-type phoneme deviants, however, showed an inverse pattern such that the MMN for the both position had the smallest amplitude despite its greater spectro-temporal deviance and its greater detectability when it was the target. These data indicate that the MMN reflects differences in phonetic structure as well as differences in acoustic spectral-energy structure of the deviant stimuli. Furthermore, the task relevance effects demonstrate that top-down controls not only affect the amplitude of the MMN, but can reverse the pattern of MMN amplitudes among different stimuli. [S0001-4966(99)07312-9]

PACS numbers: 43.64.Ri, 43.64.Sj, 43.64.Qh [RDF]

INTRODUCTION

The mismatch negativity component (MMN) of the event-related brain potential (ERP) has been interpreted as a response from a preattentive auditory change-detection system. The MMN was originally thought to reflect an acoustic spectral-energy mismatch between an incoming stimulus and the sensory memory trace of previous stimuli (Näätänen *et al.*, 1978; Näätänen, 1990, 1992; Näätänen and Alho, 1995; Schröger, 1997). It is well established that MMNs can be generated in repetitive tone sequences by a change of one or more simple auditory features, including frequency, intensity, location, duration, and pattern (Näätänen, 1992).

Recent studies suggest that reliable MMNs are also generated by phonetic deviation. Näätänen *et al.* (1997) and Dehaene-Lambertz (1997) demonstrated that the MMN is larger to natal language prototypical phonemes compared to foreign-language phonemes. Furthermore, Tremblay *et al.* (1997) found that adults who were trained to discriminate foreign-language phonemes had larger MMNs to foreign-language deviant phonemes than did adults with no discrimination training. It is only the cooccurrence of unfamiliar prephonetic auditory features (or unfamiliar combinations of prephonetic auditory features) in foreign-language phonemes

that mitigates against the stronger conclusion that the phoneme-specific MMN exists.

Näätänen and colleagues (Näätänen, 1992; Näätänen *et al.*, 1993) have consistently maintained that attention is not required to elicit the MMN. Evidence supporting the automatic nature of the MMN includes the finding that the negativity occurs to deviant auditory stimuli even when a person's attention is directed to another task in the auditory (Näätänen *et al.*, 1978) or visual modality (Sams *et al.*, 1985; Woods *et al.*, 1992, 1993; Alho *et al.*, 1994). Typically, the MMN has been recorded while the subject was reading or watching a video recording. Although no study to date has demonstrated a total absence of MMN in the absence of attention, numerous studies have shown that attended deviants evoke increased MMN amplitudes compared to ignored sounds. The specific features for which this attention effect has been found include intensity (Woldorff *et al.*, 1991), frequency (Näätänen *et al.*, 1993; Oades and Dittmann-Balcar, 1995), duration (Alain and Woods, 1994), pattern (Alain and Woods, 1997; Nordby *et al.*, 1988a, b; Schröger *et al.*, 1996), and tones embedded in a narrative (Trejo *et al.*, 1995).

A confound in some earlier studies examining the effects of attention on the MMN (e.g., Woldorff *et al.*, 1991) has been the possible inclusion of target-related deviant components (N2b) that may have contributed to the measured attentional modulation of MMN. To eliminate confounding N2b effects, some later studies have evaluated attention ef-

^{a)}Present address: Department of Radiology, Biomagnetic Imaging Laboratory, UC San Francisco, 513 Parnassus Avenue S-362, San Francisco, CA 94143-0628. Electronic mail: mszyman@mail.radiology.ucsf.edu

^{b)}Electronic mail: yund@ebire.org

fects on nontarget deviants for stimulus features such as intensity (Näätänen *et al.*, 1993), frequency (Trejo *et al.*, 1995), and tonal pattern (Alain and Woods, 1997), and have nonetheless shown attention-related MMN enhancements. Attentional modulation of the MMN has been alternatively interpreted as the result of (1) increased gain of the MMN generator for attended deviants (Näätänen *et al.*, 1993), or (2) inhibition of MMN in the ignored channel (Trejo *et al.*, 1995; Alain and Woods, 1997).

The effect of attention on the speech-evoked MMN has not been studied. There are reasons why the attentional modulation of the speech-evoked MMN might be expected to be either more or less than that of the nonspeech MMN. Since speech is a complex high-information-content signal, attention might be expected to facilitate the phoneme MMN because more sophisticated information processing might be required. Conversely, since speech perception is highly overlearned and automatic, attentional modulation might be expected to be reduced (Kraus *et al.*, 1996; Näätänen *et al.*, 1997). Alternatively, if attention modulates the gain of the MMN generator in either an excitatory (Näätänen *et al.*, 1993) or inhibitory (Trejo *et al.*, 1995; Alain and Woods, 1997) manner, then the effect of attention should be essentially the same for speech and nonspeech signals.

There has been increased interest in using the MMN as a clinical diagnostic tool because this deviant-evoked negativity might provide an objective neural measure of auditory discriminability. One idea has been to employ the MMN as a test of the auditory system's ability to transmit the acoustic information important for understanding spoken language (Picton, 1995; Näätänen, 1995; Kraus *et al.*, 1995a). To date, the MMN has been used to assess the efficacy of phoneme discrimination training and related neural plasticity (Kraus *et al.*, 1995b; Kraus *et al.*, 1996; Tremblay *et al.*, 1997), and some investigators have proposed using MMN to monitor the effectiveness of hearing aid therapy (Picton, 1995) and cochlear implants (Ponton and Don, 1995).

The present study was designed to test the effectiveness of the MMN as an indicator of the transmission of the acoustic information needed for correct phoneme identification. The ultimate goal of the research was to develop MMN techniques for evaluating a signal-processing algorithm's ability to transmit auditory information through a sensorineural hearing impairment. The current study was designed to test the hypothesis that the MMN can be used to measure prephonetic acoustic information under different attention conditions and with variation in the behavioral task.

Evidence from previous MMN studies on phonetic information processing has confounded phonetic and prephonetic acoustic contributions to the MMN because the deviant stimuli have differed from the standards along corresponding acoustic and phonetic dimensions. For instance, deviant vowels have different formant frequencies than standard vowels (Aaltonen *et al.*, 1997; Dehaene-Lambertz, 1997; Näätänen *et al.*, 1997), and deviant-consonant CVs have different formant transitions or voice-onset times than standard CVs (Kraus *et al.*, 1992; Sams *et al.*, 1990; Sandridge and Boothroyd, 1996; Tremblay *et al.*, 1997). In such studies, the observed MMNs could reflect either the acoustic or the pho-

netic deviation between vowels or consonants. The speech sounds used here were vowel-consonant-vowel (VCV) disyllables that included two different deviant features—*phoneme and intensity*—to reveal possible differences between MMNs to speech sounds elicited by phonetic and acoustic deviance. The middle consonant of the VCVs could differ in either phoneme (/b/ or /g/) or intensity (a more intense /b/). Each feature deviance also could occur in one of three temporal positions: during the first formant transition (VC), during the second formant transition (CV), or during both (VC and CV, combined). Comparing the MMNs to phoneme and intensity deviants for the three temporal positions might demonstrate differences between phonetic and acoustic MMNs that would not otherwise be apparent because of differences in the magnitudes of acoustic and phonetic deviance across the deviance positions. For phoneme and intensity deviants, the acoustic deviance would be greatest for the both position. Thus, if phoneme and intensity MMNs were generated in response to simple acoustic deviation, then phoneme and intensity MMNs should be maximal with both-position deviants and should be less for either first or second position alone. Alternatively, if phoneme MMNs are generated by speech-specific phoneme deviance, then the both-position phoneme MMN would be expected to be smaller than either single-position MMN. This prediction is based on the fact that the both-position phoneme deviant is less *phonetically* deviant than are the single-position phoneme deviants: In addition to the deviant consonant that is present in all phoneme deviants, the single-position phoneme deviants also include the conversion of the single middle consonant into a *consonant pair*—the stimulus is a VCV when the VC and CV transitions are phonetically consistent, and a VCCV when they are not.

I. METHODS

A. Subjects

The 13 subjects (five males and eight females) whose data were included in the analysis ranged in age from 20 to 39 years (mean 25.8). All were right handed, except for one ambidextrous male and one left-handed female, and all had normal hearing (thresholds 20 dB HL or less from 250 to 4000 Hz). Subjects completed three experimental sessions, one behavioral session and two recording sessions. The purpose of the behavioral session was to train the subject to detect the various targets under simplified conditions and to verify that each subject could detect the targets under the demanding conditions¹ to be used in the recording sessions (see Sec. I C). Subjects were paid for their participation and gave informed consent, according to DVA guidelines.

B. Stimuli

Vowel-consonant-vowel (VCV) stimuli were synthesized at a 20-kHz sampling rate with a Klatt speech synthesizer, KLSYN88a (MIT and Sensimetrics Corporation), in cascade mode with a 1 ms update interval. The digital stimulus waveforms were converted to analog signals (Data Translation DT2823), filtered (Tucker-Davis Technology, TDT, antialiasing filter FT6-92, 8 kHz), attenuated (TDT PA4),

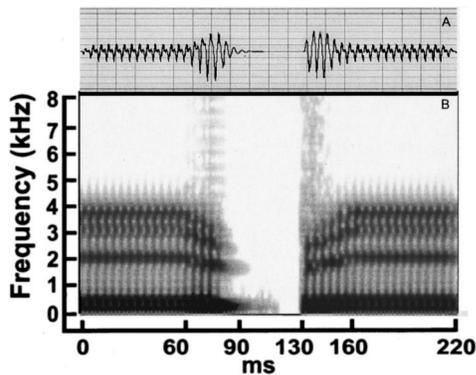


FIG. 1. Waveform (a) and spectrogram (b) of the standard VCV stimulus, /ibi/. The onset of the initial vowel and the offset of the final vowel were shaped with a 10-ms linear ramp. Steady-state vowels were present in the intervals 10–60 ms and 160–210 ms. The VC formant transitions occurred at 60–90 ms, the stop-consonant gap occupies 90–130 ms, and the CV formant transitions occurred at 130–160 ms. Deviant stimuli differed from standards only in the VC and/or CV formant-transition intervals.

and delivered to Etymotic ER-2 earphones through a power amplifier (TDT HB6). The intensity of the steady-state portion of the vowels in the VCV stimuli was 75 dB SPL. The stimulus duration was 220 ms, including 10-ms linear onset and offset shaping.

The female-voice /ibi/, illustrated in Fig. 1, was the standard left-ear stimulus. Like this female-voice /ibi/, all left-ear stimuli had a fundamental frequency (F_0) of 220 Hz and formant frequencies of 330 (F_1), 2000 (F_2), 3000 (F_3), and 3600 Hz (F_4) in the vowel. The right-ear male-voice /ibi/ differed only in F_0 (156 Hz) and F_1 (270 Hz). The difference in voice pitch was included to assist the subjects in attending to stimuli in one ear and ignoring those in the other ear. Neither ear nor voice were intended to be a factor in the analysis, but separate analyses of left-ear-female-voice and right-ear-male-voice stimuli showed the same effects and interactions for the major factors of the analyses (attention, deviant feature, task relevance, and deviance position—see below for factor definitions and descriptions).

There were two types of deviant features: (1) phoneme deviants, where /b/ was changed to /g/, and (2) intensity deviants, where the intensity of /b/ was increased (referred to as /B/). The end points for the formant transitions for /b/ were 1600 (F_2), 2300 (F_3), and 2800 (F_4). Those for /g/ were 2395 (F_2), 3000 (F_3), and 3305 Hz (F_4). In the intensity deviants, the intensity increases were restricted to the formant-transition regions of the VCVs, with the intensity (Klatt AV parameter) maxima occurring from 70–80 ms and at 130 ms. In a brief pilot experiment, a 15-dB increase in AV in these regions (phased in from 60–70 ms and out from 80–90 ms and 130–160 ms) was as detectable as the phoneme deviant. As already mentioned in the Introduction, the phoneme or intensity deviance could occur in three different temporal positions, (first) during the VC formant transitions, (second) during the CV formant transitions, or (both) during VC and CV formant transitions.² Table I lists the different types of deviant stimuli and the temporal onsets of the deviance.

TABLE I. The deviant stimuli used in the present experiment. The standard stimuli were the VCV /ibi/. Listed here are the deviance position, the time of deviance onset relative to the beginning of the initial vowel (time=0 ms), and the notation for each phoneme and intensity deviant.

Deviance position	Onset of deviance (ms)	Phoneme deviant	Intensity deviant
Both	60	/igi/	/iBi/
First	60	/igbi/	/iBbi/
Second	130	/ibgi/	/ibBi/

C. Procedure

The subject's task was to detect one type of deviant feature (intensity or phoneme) in one ear, independent of the position of the deviance. Since preliminary experiments indicated that it was difficult for inexperienced subjects to perform these detection tasks without some training, the experiment began with a training session to ensure that all subjects would be able to accurately detect the target deviants in the EEG recording sessions. In the initial training, stimuli were delivered to only one ear and at a slower rate than in the main experiment (200 ms added to the main-experiment interstimulus interval, ISI). The stimulus sequences used in training included the same proportion of intensity and phoneme deviants as in the main experiment. When the subject scored greater than 90% correct on the one-ear training for intensity and phoneme deviants, training proceeded to sequences with the same random right- and left-ear stimulus presentation and the same stimulus timing as in the main experiment. Subjects had to correctly detect at least 80% of each designated target with at most five false alarms (in a standard block of 800 total stimuli, 60 of which were targets), to continue on to the EEG recording sessions. This training was essential to be sure that subjects could perform the demanding experimental task.

The EEG recording sessions were scheduled on two subsequent days. In each block of 800 stimuli subjects responded to one type of deviant stimulus in one ear (the target), while withholding responses to all other stimuli. Target detection was difficult and subjects spontaneously reported that they would miss targets if they failed to pay close attention to stimuli in the attended ear. Hand of response was counterbalanced across the two sessions for each subject, with subjects pressing the button on the top of a joystick with the thumb of the designated hand. Subjects were informed of their performance at the end of each block. The target was the same within each block and was varied across blocks in a pseudo-random order. Each subject completed ten blocks for each target.

Before each block, the subject was informed of the target, and then the block began with 24 sample targets, 8 for each deviance position. A 3-second interval separated the sample targets from the test block of 800 random stimuli. During the test block, 400 stimuli were presented to each ear. The order was random with the constraint that each deviant type (right- or left-ear phoneme or intensity deviant) was followed by at least two other stimuli before it was repeated. Stimuli in each ear consisted of 70% standards (/ibi/), 15% phoneme deviants (5% each of /igi/, /igbi/, and /ibgi/), and

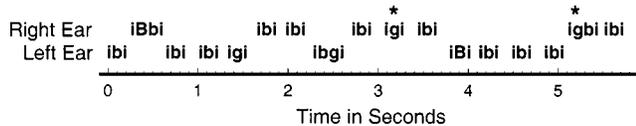


FIG. 2. A 6-s example of the experimental paradigm when the target was phoneme-feature deviance in the right ear. Asterisks indicate target sounds, a both-position and a first-position phoneme deviant—the second-position phoneme deviant /*lgi*/ in the right ear was also one of the targets, but it did occur in this 6-s interval. Note the example begins with a standard left-ear sound, switches to a right-ear nontarget first-position intensity deviant, then three left-ear sounds, the third of which is a target-type both-position phoneme deviant, etc. In other blocks, the right-ear intensity deviants or the left-ear phoneme or intensity deviants were the target sounds.

15% intensity deviants (5% each of /*iBbi*/, /*ibBi*/, and /*iBi*/). The ISI in the test block was varied randomly between 80–125 ms in a rectangular distribution, such that the average rate of stimulus delivery was 3/s (Fig. 2).

D. EEG recording

The EEG (bandpass 0.01 to 100 Hz) was recorded from 30 electrodes: BE (below left eye), LE (lateral to left eye), nose, Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC6, LM (left mastoid), T7, C3, Cz, C4, T8, RM (right mastoid), CP5, CP6, P7, P3, Pz, P4, P8, PO5 (center of O1-P3-P7 triangle), PO6 (center of O2-P4-P8 triangle), O1, and O2. Electrode recordings were referenced to four interconnected, EKG-balanced electrodes at the base of the neck (Woods and Clayworth, 1985). The subject was grounded through an electrode at AFz.

The EEG was sampled at 256 Hz and continuously digitized to disk for later processing. The BE and LE recording channels were used to detect eye movements. During the off-line averaging, epochs with eye movements, signal clipping, or excessive amplitude swings (greater than 75 μ V peak-to-peak) were not included in the averages.

E. Data analysis

Difference waves were obtained by subtracting ERPs for standard VCVs from those of deviants. The MMN amplitudes were measured relative to the 200 ms of prestimulus baseline EEG activity. The analysis of MMN scalp-distribution included 25 electrode sites, all except BE, LE, nose, LM, and RM. The detailed analysis of MMN duration and amplitudes was done at the Fz electrode. In the analysis of MMN duration, mean amplitudes were measured in 20-ms time bins for each subject for each combination of the factors: attention (attended, ignored), deviant feature (phonetic, intensity), deviance position (first, second, both), and task relevance (target-type feature, nontarget-type feature). With subject as the random variable, the voltage in each bin was tested (ANOVA) to determine if it was significantly different from zero. The duration of significant negativity (or positivity) was then defined as the time interval over which at least three consecutive 20-ms bins showed significance (at least two at $p < 0.01$).

Comparison across different factors and/or electrode sites requires an unbiased measure of amplitude. A source of difficulty was that the different deviant features and deviance

TABLE II. Mean RTs, target-detection hit rates, and false alarm responses (FAs) for deviant-phoneme and deviant-intensity targets in each deviance position. The RTs were measured from the onsets of deviance given in Table I. Note that either the phoneme or intensity deviants in one ear were the behavioral targets, independent of the position of the deviance within the consonant portion of the VCV.

Position	RT (ms)		Hit rate (%)	
	Phoneme	Intensity	Phoneme	Intensity
Both	389	375	96	95
First	402	401	93	82
Second	415	401	82	73
All	402	392	90	83
FAs (number per block of 800 stimuli)			2.38	1.35

positions produced MMNs with different latencies and somewhat different shapes. Therefore, to minimize bias in the MMN mean amplitude measurements, the temporal location of the 100-ms measurement interval was determined for each combination of deviant feature and deviance position, based on the intervals of significant negativity, determined as described above. Relative to stimulus onset, the measurement intervals for the intensity feature were 180–280 ms for both and first positions and 240–340 ms for second position. The measurement intervals for the phoneme feature were 200–300 ms (both), 220–320 ms (first), and 280–380 ms (second). These intervals were used for all MMN amplitude measurements for all 25 electrodes in the scalp-distribution analysis and for Fz in the detailed MMN amplitude analysis.

The ANOVAs for MMN amplitudes at the Fz electrode included subjects as the random factor and attention (attended, ignored), deviant feature (phonetic, intensity), deviance position (first, second, both), and task relevance (target type, nontarget type) as within-subjects factors. Separate ANOVAs were performed for the phonetic and intensity features because differences between these deviant features were expected, in particular for deviance position. Scalp distributions were evaluated across the 25 scalp sites, including Greenhouse–Geisser corrections assuming that the 25 sites represented only four independent measures. Although this correction was used to determine significance levels, both the original degrees of freedom and the corrected degrees of freedom will be reported here for each analysis.

II. RESULTS

A. Behavioral reaction times and accuracy

The overall mean reaction time (RT) was 397 ms with 87% of the targets detected and less than two false alarms (FAs) per block of 800 stimuli. Table II shows the RTs and hit rates broken down by deviant feature and deviance position. Significantly more deviant-phoneme targets were detected (90%) compared to deviant-intensity targets (83%) [$F(1,12) = 10.56$, $p < 0.01$]. The reverse trend for slower RTs for phoneme targets (402 versus 392 ms) was not significant. The temporal position of the deviance affected the hit rates and RTs of phoneme and intensity targets in similar ways. The both position produced the highest hit rates [$F(1,12) = 23.21$, $p < 0.0001$] and the fastest RTs [$F(1,12)$

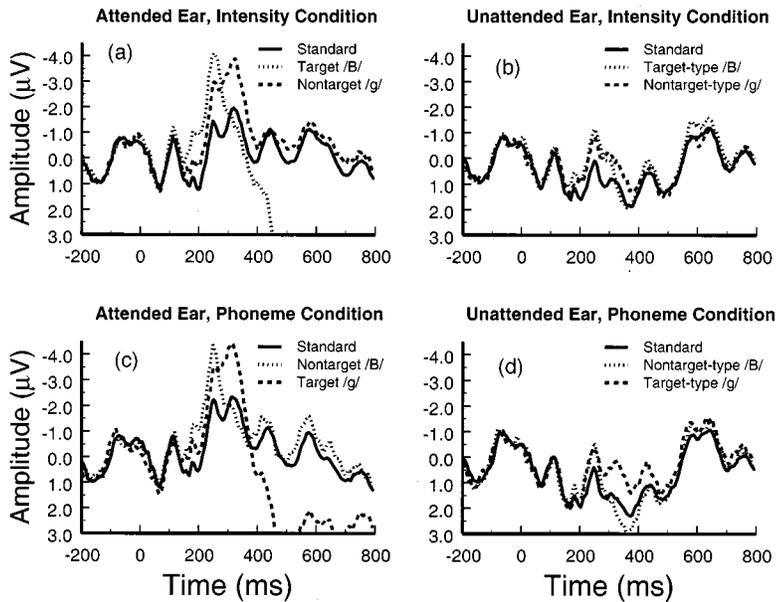


FIG. 3. The ERPs for standard and deviant stimuli at Fz averaged across all subjects. Each panel includes ERPs for a standard stimulus and for the intensity and phoneme deviants that were presented in the same ear under the same condition. Here, the data are collapsed across ear of presentation and deviance position. Responses are shown for stimuli in the attended ear in (a) and (c) and for stimuli in the unattended ear in (b) and (d). The ERPs fail to show the clear pattern of components that would be expected for brief-tone-burst or click stimuli, probably because of their long duration and the amplitude and formant-frequency variation within the stimuli. Expected negative and positive difference components, however, are apparent when the waveforms for deviant and standard stimuli are compared. The ERPs are plotted relative to the average level in the baseline interval (the 200 ms prior to stimulus onset). Time 0 is stimulus onset. The voltage scale was chosen to illustrate the negative differences between deviant and standard ERPs, putting the large positive (P3) response to target stimuli off the bottom of the negative-up voltage ordinate.

= 219.78, $p < 0.0001$]. The performance difference between first- and second-position targets was significant for hit rate [$F(1,12) = 28.09$, $p < 0.0001$], but not for RT.

B. ERPs for standard and deviant stimuli

The ERPs for standard and deviant stimuli are illustrated in Fig. 3. Panels (a) and (c) include ERPs to stimuli in the attended ear, while (b) and (d) include ERPs to stimuli in the unattended ear. Similarly, (a) and (b) show ERPs when the attended intensity deviants were targets and (c) and (d) when the attended phoneme deviants were targets. Responses were averaged across ear of delivery and deviance position. As can be seen in Fig. 3, the ERPs to the standard VCVs were of low amplitude, never exceeding $\pm 2.5 \mu\text{V}$. This low-amplitude and relatively indistinct standard-VCV ERP is not surprising given the rapid presentation rate (3/s), the long VCV duration (220 ms), and the complex intensity and formant-frequency modulation that occurs in the VCV stimulus. This intensity modulation of the VCV was shown in Fig. 1: there is an onset at 0 ms, an offset at 90 ms, a second onset at 130 ms, and then the final offset at 220 ms; formant fre-

quencies change in the time ranges 60–90 and 130–160 ms. In spite of these limited ERPs to the standard VCVs, clear negative and positive differences are present in the average ERPs to the deviant stimuli. The MMN was isolated in the difference waves by subtracting the ERPs for the standard VCV from the ERPs for the deviants. The MMN onset occurred at 150–200 ms. This relatively long latency of onset reflects the fact that the standard and deviant stimuli were identical from 0 to 60 ms.

C. Time course of the negative-difference components

As expected, the time course of the MMN, as well as the positive response to detected targets (P3b), varied with the deviant feature and the deviance position. Table III gives the intervals over which the difference waves were significantly negative. Virtually all of the deviant stimuli generated broad negative differences in the MMN latency range.³ Although no further statistical tests were done to compare durations for different deviant stimuli or for different factors, some general patterns of time differences are apparent. When comparisons are made within the same deviant feature, both- and

TABLE III. The durations of significant difference-wave negativity for phoneme and intensity deviant features under all combinations of the other three factors (attention, deviance position, and task relevance). Mean voltages of the difference waves were measured in 20-ms intervals for each subject and the significance of the difference from zero was determined (ANOVA). Only durations with at least three consecutive significant 20-ms intervals, including at least two with $p < 0.01$, are listed here. The time intervals are defined relative to stimulus onset, rather than to the onset of the deviance, to facilitate the identification of significant portions of the difference responses in Figs. 7 and 9. (To convert to intervals relative to deviance onset, subtract 60 ms from both- and first-position values and 130 ms from second-position values.)

Deviant stimulus	Durations of significant negativity (ms, relative to stimulus onset)			
	Attended target type	Attended nontarget type	Ignored target type	Ignored nontarget type
/iBi/	160–280	140–280	160–300, 560–720	200–260
/iBbi/	80–280	140–280	140–280	180–260
/ibBi/	160–360	220–360	ns	ns
/igi/	120–340	200–300	200–360	200–300
/igbi/	160–240, 260–340	120–400, 440–700	220–420	220–460
/ibgi/	240–440	220–400	280–380	ns

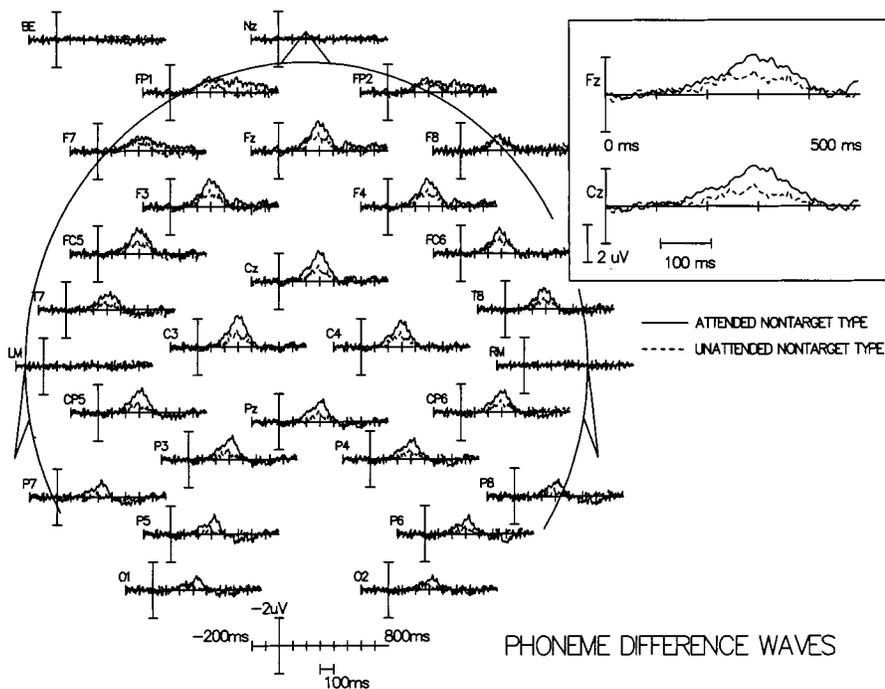


FIG. 4. The MMN in attended and ignored conditions for nontarget-type phoneme deviants averaged across all subjects. The difference waves are plotted relative to the average baseline level (from -200 to 0 ms) and the data are pooled over ear of stimulation and deviance position. Note the fronto-central scalp distribution and the larger MMN amplitude when stimuli were attended. The responses at Fz and Cz are expanded in the call-out box to permit more detailed observation; in these responses combined across deviance position, the peak of the MMN occurs at about 290 ms.

first-position negativities tend to be similar and to start and/or end before second-position negativities. Similarly, when corresponding deviance positions are compared, phoneme-deviant negativities tend to occur later than intensity-deviant negativities.

D. Attention effects on the MMN

A strong attention effect was found for the MMN amplitudes measured at Fz in the 100 -ms time intervals derived from the durations of significant negativity. The MMN had a greater mean amplitude for the attended ear compared to the ignored ear [$F(1,12) = 33.93$, $p < 0.0005$]. Stimuli in the ignored ear generated MMNs of about half the amplitude

($-0.98 \mu\text{V}$) of the MMNs of stimuli in the attended ear ($-1.91 \mu\text{V}$). The effect of attention on MMN amplitude is illustrated in Figs. 4 and 5, for nontarget phoneme and intensity deviants, respectively. These figures show the scalp distribution of the difference waves for attended and ignored nontarget-type deviants. The minimal-amplitude difference waves seen with the present reference-electrode system for the nose (Nz) and mastoid (RM, LM) electrodes indicates that these scalp distributions can be compared directly with those of other studies that used either nose or mastoid reference electrodes. Both deviant features yielded the fronto-central distribution typical of the MMN, with maximum amplitude (peak about $2 \mu\text{V}$) at Fz, Cz, and adjacent sites (F3,

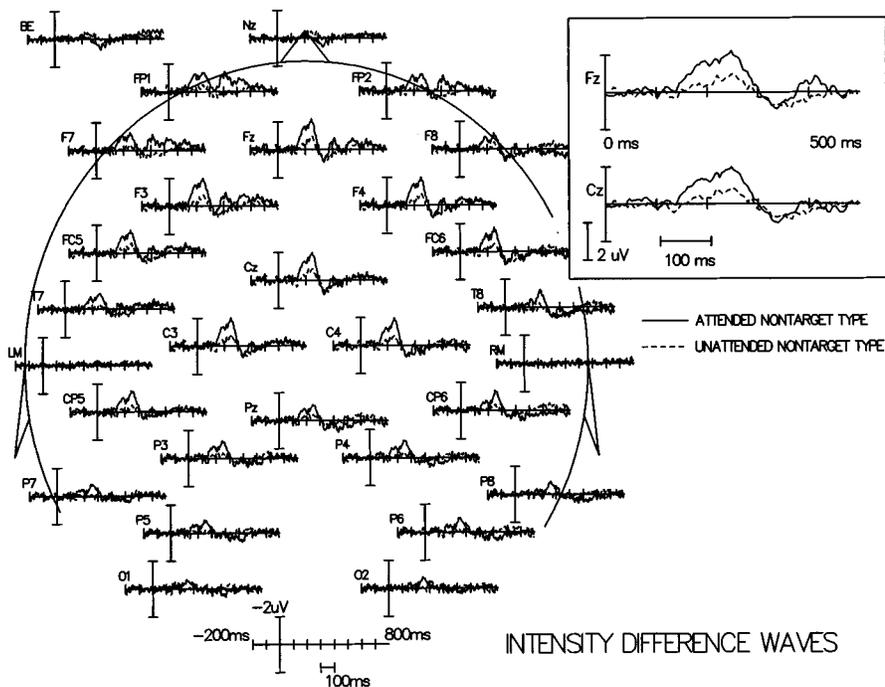


FIG. 5. The MMN in attended and ignored conditions for nontarget-type intensity deviants. Otherwise the same as Fig. 4. Note the fronto-central scalp distribution and the larger MMN amplitude when stimuli were attended, similar to the phoneme data of Fig. 4. The peak of the MMN occurs at about 240 ms, or about 50 ms earlier than for phoneme deviants.

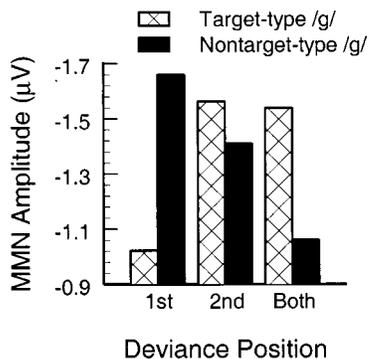


FIG. 6. The interaction of deviance position and task relevance for phoneme deviants. The mean amplitude of the MMN at Fz is plotted for each deviance position for target-type deviants (cross-hatched bars) and nontarget-type deviants (solid bars). See text for further discussion.

F4, C3, C4, FC5, and FC6). Phoneme and intensity MMNs had similar amplitudes and showed similar enhancements with attention, consistent with the lack of a significant main effect of deviant feature and the lack of an interaction between the factors attention and deviant feature. Indeed, attention produced only a main effect, showing no significant interactions with any other factor—deviant feature [$F(1,12) = 1.48, ns$], deviance position [$F(1,12) = 1.14, ns$], or task relevance [$F(1,12) < 1.0$].

In the analysis of the scalp distribution of the MMN, the mean MMN amplitudes were normalized to equate amplitudes at Fz. This normalization procedure eliminates the main effects and interactions among the principal factors of the study (deviant feature, deviance position, attention, and task relevance), but permits an unbiased evaluation of the interactions of scalp location with these factors. As might be assumed from Figs. 4 and 5, the main effect of scalp location was highly significant for nontarget phoneme deviants [$F(24,288) = 10.74, p < 0.001(3,36)$] and for nontarget intensity deviants [$F(24,288) = 8.34, p < 0.005(3,36)$]. The interaction of attention and scalp location, however, was not

significant for either phoneme [$F(24,288) = 1.36, ns(3,36)$] or intensity deviants [$F(24,288) < 1.0$].

E. Deviant feature, deviance position, and task relevance

Since it was anticipated that the effect of deviance position might be different for the deviant features, phoneme and intensity, the remainder of the statistical analysis was carried out separately for these two features. Each of these analyses included the factors deviance position, task relevance, and attention. Attention was included in these analyses to detect interactions with the other critical variables and, since there were none, it will not be discussed further. In the case of the phoneme feature, neither deviance position nor task relevance produced a significant main effect. However, their interaction was significant [$F(2,24) = 4.06, p < 0.05$] and is illustrated in Fig. 6. For the target-type phoneme deviant (cross-hatched bars), the both position yielded a large-amplitude MMN and the first position yielded the smallest amplitude MMN. In contrast, for the nontarget-type phoneme deviant (solid bars), the both position yielded the smallest and the first position yielded the largest MMN. Post-hoc analyses (Tukey, least significant difference test) indicated that MMN amplitude for the first-position nontarget-type deviant was significantly greater than those for the first-position target-type and the both-position nontarget-type deviants (each, $p < 0.05$).

The interaction between deviance position and task relevance for the phoneme deviants is also apparent in the difference waves from the Fz electrode, as shown in Fig. 7. The four panels of the figure show the deviant-phoneme difference waves for (a) targets, i.e., attended target types; (b) attended nontarget types; (c) ignored target types; and (d) ignored nontarget types. The larger MMNs in panels (a) and (b) compared to (c) and (d) represent the attention effect. The attention effect is unimportant in the present context, except to illustrate that the deviance-position-by-task-relevance in-

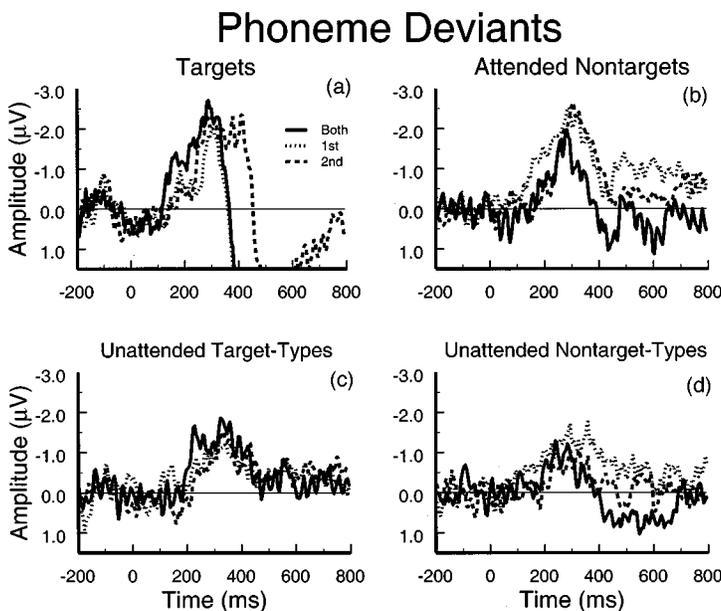


FIG. 7. Average phoneme MMNs at Fz for each deviance position (both in solid, first in dotted, and second in dashed lines) for (a) attended target-type deviants, i.e., targets; (b) attended nontarget types; (c) unattended target types; and (d) unattended nontarget types. Note the same interaction between deviance position and task relevance independent of attention condition. See text for further discussion.

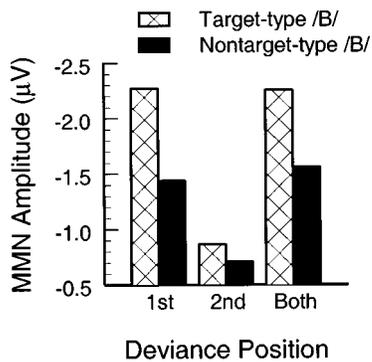


FIG. 8. The effect of deviance position and task relevance for intensity deviants. The mean amplitude of the MMN at Fz is plotted for each deviance position for target-type deviants (cross-hatched bars) and nontarget-type deviants (solid bars). Note the lack of interaction between deviance position and task relevance: Whether the intensity deviant is target or nontarget type, the first and both deviance positions produced equally large MMNs while the second deviance position produced a much smaller response. Compare with the interaction seen for phoneme deviants in Fig. 6. (In comparing Figs. 6 and 8, note the different MMN-amplitude scales; the second-position intensity-deviant responses were smaller, while the first- and both-position target-type intensity-deviant responses were larger, than any of the phoneme-deviant responses.)

teraction occurred independently of attention. The both-position MMN has the largest amplitude in each of the target-type panels [(a) and (c)] and the first-position MMN has the largest amplitude in each of the nontarget-type panels [(b) and (d)]. The occurrence of the interaction in the MMNs to ignored, as well as attended, deviants demonstrates that the source of the interaction is not in recognizing and responding to the target, but rather in the preresponse processing of target- and nontarget-type deviants.

A different pattern of results was obtained for the intensity feature, as shown in Fig. 8. Here, there were significant main effects of task relevance [$F(1,12)=5.67, p<0.05$] and deviance position [$F(2,24)=21.77, p<0.0005$], but their interaction failed to reach significance [$F(2,24)=3.36, ns$]. In the effect of task relevance, target-type intensity deviants yielded larger MMN amplitudes than nontarget-types for

each deviance position. For deviance position, the order of decreasing MMN amplitude was both, first, and second. The similarity of first- and both-position MMN amplitudes within target- and nontarget-type intensity deviants contrasts sharply with the crossover interaction found for the phoneme deviants and illustrated in Fig. 6.

Figure 9 illustrates the main effects of deviance position and task relevance in the Fz difference waves for the intensity deviants. The main effect of deviance position is seen in each panel: both- and first-position MMNs are of similar amplitudes and considerably larger than the second-position response. The main effect of task relevance is seen in comparing the higher-amplitude difference waves in panels (a) and (c) (attended and unattended target types) with the lower-amplitude difference waves in panels (b) and (d) (attended and unattended nontarget types). Again, as was the case for Fig. 7 (phoneme-deviant difference waves), comparison of panels (a) and (b) (attended target and nontarget types) with panels (c) and (d) (ignored target and nontarget types) demonstrates the attention effect that is not of primary interest here.

The scalp-distribution analysis for deviance position (collapsed across attention and task relevance) also showed differences between the MMNs for intensity and phoneme features. For the intensity deviants, there was no significant interaction between electrode location and deviance position. For the phoneme feature, however, there was an interaction of electrode site by deviance position ($S \times P$) [$F(48,576)=2.76, p<0.05(6,72)$].

The pattern of the $S \times P$ interaction for phoneme deviants is interesting because it was right-left asymmetrical. All 11 left-hemisphere electrodes showed the smallest MMN for the both position, while 9 of 11 showed the largest MMN for the first position. The MMN amplitude differences for the both and first positions were less consistent in right-hemisphere electrodes, with 6 of 11 showing larger MMNs for the both position. A secondary analysis was done to verify this right-left asymmetry in the deviance-position ef-

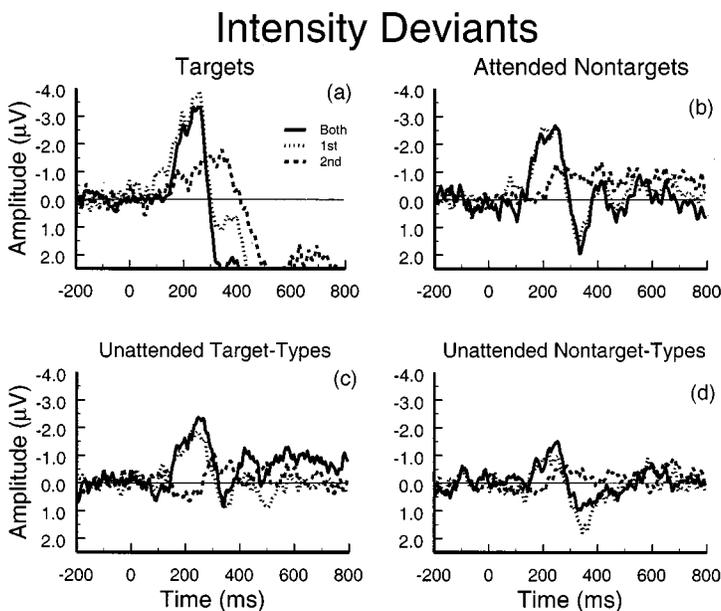


FIG. 9. Average intensity MMNs at Fz for each deviance position (both in solid, first in dotted, and second in dashed lines) for (a) attended target-type deviants, i.e., targets; (b) attended nontarget types; (c) unattended target types; and (d) unattended nontarget types. In contrast to the phoneme MMN of Fig. 7, intensity MMNs show essentially the same deviance-position effect in all four panels. The task-relevance main effect is also apparent, with target-type MMNs greater than nontarget-type MMNs.

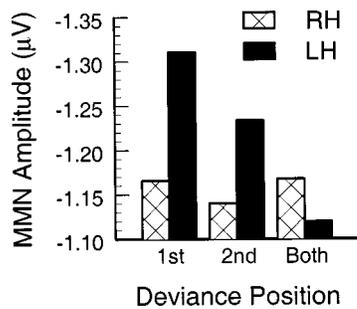


FIG. 10. The interaction of deviance position and electrode-site hemisphere for phoneme deviants. The mean amplitude of the MMN is plotted for each deviance position for right-hemisphere (RH, cross-hatched bars) and left-hemisphere (LH, solid bars) electrode sites. Data were collapsed across attention condition and task relevance, with F4, FC6, C4, CP6, and P4 averaged to obtain the RH response amplitudes and F3, FC5, C3, CP5, and P3 averaged to obtain the LH response amplitudes.

fect for the phoneme deviants, including the average MMN amplitudes for electrode sites F3, FC5, C3, CP5, and P3 representing the left hemisphere (LH), and F4, FC6, C4, CP6, and P4 representing the right hemisphere (RH). These electrode sites were chosen because they showed larger MMN amplitudes than other lateralized sites, as can be seen in Figs. 4 and 5. In the ANOVA, there was no significant main effect of deviance position [$F(2,24) < 1.0$] or laterality [$F(1,12) < 1.0$], but the expected interaction was significant [$F(2,24) = 5.76, p < 0.01$] and is illustrated in Fig. 10. Post-hoc analyses (Tukey, least significant difference test) found significant differences for deviance position over the LH, with the both-position MMN less than the first-position MMN ($p < 0.001$) and the second-position MMN ($p < 0.01$). There were no significant differences among the RH MMNs. In comparisons between LH and RH response amplitudes, LH MMNs were significantly greater than RH MMNs for the first ($p < 0.001$) and second ($p < 0.05$) deviance positions, but the reversal (RH greater than LH) for the both position failed to reach significance. The same secondary analysis was done for intensity deviants, but it will not be considered further because it showed only the main effect of deviance position already described above and illustrated in Figs. 8 and 9, with no laterality effect or interaction.

III. DISCUSSION

There are three major findings in the present study: (1) the large, but relatively simple, non-interactive effect of attention on the MMN; (2) the non-interactive effects of the task relevance and deviance position on the MMN to intensity deviants; and (3) the complex interactions of task relevance and deviance position on the MMN to phoneme deviants. Although attention was not the primary focus of the present study, the results suggest a new explanation for the effects of attention on the MMN. An extensive discussion of attention effects seen in previous studies is required to evaluate this new hypothesis. The effects of task relevance and deviance position for the intensity deviants were quite simple and need to be discussed only briefly. These intensity-deviant results were not surprising, and thus provide a contrasting control condition for the phoneme-deviant results. The effects of task relevance and deviance position for the

phoneme deviants are the most important part of the present results because the purpose of the experiment was to determine the contribution of phoneme processing to MMN generation. In spite of the complexity of these phoneme-deviant results, their implications with respect to the existence of the phoneme-specific MMN seem quite clear. Each major finding and its implications for MMN generation and modulation will be discussed in a separate section below.

A. Attention effects

In the present experiment, the listener had to attend to the stimuli in one ear in order to detect intensity /B/ or phoneme /g/ deviants in that ear. The MMN amplitude for unattended-ear deviants was only 51.3% of that for attended-ear deviants, and this attention effect was independent of the deviant feature (intensity or phoneme), task relevance (target or nontarget type), and deviance position (first, second, or both). In addition, there were no significant interactions that included attention, although the other factors did produce main effects and interactions among each other. Such a simple, non-interactive effect of attention suggests a simple mechanism.

Within the context of the model of MMN generation proposed by Näätänen (1992, 1995), there are three points at which attentional modulation could produce simple effects: (1) the output of the MMN generator, (2) the sensory input signals to the MMN generator, or (3) the memory trace of the standard stimulus. The results of many MMN-attention experiments would not be expected to discriminate among these alternatives. Modulation of the MMN output [alternative (1)] would require a coupling between a process of identifying the stimulus as attended or unattended and the MMN modulator. This alternative (1) is probably the most flexible and least predictive of the three, because it includes components that have unknown properties and are not necessarily involved in any process other than attentional modulation of the MMN.

Modulation of sensory input [alternative (2)] might require early differentiation based on fundamental stimulus features like location or frequency band rather than a more complex stimulus attribute like the direction of frequency glide. Such simple features could be effective because of the possibility of presetting a filter to enhance attended relative to unattended sensory input. Other, more complex features might also facilitate sensory gating, if they can serve as a basis for the perceptual organization of an ongoing sound stream (Bregman, 1990). For example, with subjects listening for information or shadowing one of two voices, Woods *et al.* (1984) demonstrated short-latency selective attention effects in ERPs for unrelated speech probes in the attended voice. Thus, alternative (2) has predictive value in that attention directed to highly complex stimulus attributes should not affect the MMN, whereas attention to more fundamental auditory features should produce attention effects.

Modulation of the memory trace of the standard stimulus [alternative (3)] could provide a basis for the attended deviants to generate higher-amplitude MMNs, if a more accurate or precise memory trace were made of the attended standard stimulus. Even if the comparison of each stimulus

with the memory trace of its standard were completely pre-attentive, a better representation of the attended standard should generate larger difference signals for the attended deviants. In the present experiment, there was no difference between the attention effects on the phoneme (/g/) and intensity (/B/) deviants. This equality would be the expected result if the attention effect were based on a more precise representation of the standard /ibi/, because exactly the same standard would serve for intensity and phoneme deviants. In contrast, consider Alain and Woods (1997), where pure-tone frequency pattern and intensity were the different features in the attended and unattended ears. The pattern-feature MMN showed greater attentional modulation when intensity was the target than did the intensity MMN when frequency pattern was the target.⁴ A sharpened memory trace of frequency pattern would not require more precise intensity information, but a sharpened memory trace of standard-intensity tones would include precise frequency specification, consistent with the asymmetry in the attention effect that was found. Cowan (1995) provides an extensive framework for considering such effects of attention on the memory trace.

The memory-trace modulation model also can account for other differences in attention effects. Woods *et al.* (1992) (and Alho *et al.*, 1992) reported that the MMN was more sensitive to attentional manipulation when the frequency of the deviant (1064 or 1050 Hz) was closer to the frequency of the standard tone (1000 Hz) than when it was more distant (1500 Hz). Clearly, this would be the expected result if the effect of attention were to sharpen the representation of the standard stimulus: A more accurate memory trace of the 1000-Hz standard would be needed to differentiate it from the deviant when the frequency difference was less. Memory-trace modulation also might account for the lack of an attention effect in Alho *et al.* (1994). The auditory task in their attend-auditory conditions was to respond to only the more distant deviant (1300 Hz) and not to respond to the deviant (1050 Hz) that was closer to the standard (1000 Hz). If this task caused the subject to form a less precise memory trace of the standard stimulus than if the close deviant also had been a target (as in Woods *et al.*, 1992), then the memory-trace modulation model would be consistent with the absence of an attention effect.

In some studies of the effect of attention on the MMN, the attended and unattended stimuli were very different. Trejo *et al.* (1995), for example, measured pure-tone MMNs while subjects listened to a mixture of narrative speech and pure tones. MMNs to pure tone deviants were smaller when subjects attended to speech than when they attended to the tones. Alho *et al.* (1992) and Woods *et al.* (1992) also demonstrated that attention to a visual task produced smaller pure-tone MMNs relative to attention to the tones. Furthermore, although attention to vision produced smaller MMNs, the difficulty of the visual task did not affect the MMN amplitude. In the memory-trace modulation model, attention to any different category of stimulus should eliminate attentional sharpening of the memory trace of the standard stimuli. Thus, the model is consistent with the reduction of the pure-tone MMN by attention to speech or a visual task. The insensitivity to visual-task difficulty also seems consis-

tent with the model, as long as the effect of attention is characterized as a sharpening of the memory trace due to attention to those stimuli. In sum, the memory-trace modulation model seems to account for the results of the available MMN-attention studies more adequately than either of the other alternatives.

One final topic should be considered briefly before the discussion of attention effects is concluded. It is difficult to eliminate a possible N2b contribution to the present attention effect, or to the task-relevance effects to be discussed below (Näätänen, 1992; Alho *et al.*, 1994; Alho, 1995). Two response properties can be used to distinguish the N2b from the MMN: (1) the N2b has a longer response latency than the MMN and (2) the N2b does not reverse polarity at the mastoid. Although neither of these properties provides a completely definitive means of separating N2b and MMN components, they can be used together to evaluate the N2b contribution to any particular effect. Briefly, the magnitude of the effect can be measured in an earlier time interval, when the mastoid difference waves (relative to Nz) are positive.⁵ If a similar magnitude effect occurs in this earlier interval, then the effect cannot be attributed to N2b contamination. In the case of the attention effect in the present experiment, the MMN reduction for deviants in the unattended channel was slightly larger in the earlier intervals: 58% versus 54% (in the original interval) for intensity deviants and 48% versus 43% (in the original interval) for phoneme deviants. Thus, the early-interval analysis indicated that N2b contamination was not a major factor in this attention effect.

B. Intensity deviants

The present results demonstrate that the deviance position and the task relevance of the deviant stimulus influenced the MMN for intensity deviants. (In the intensity deviants, the early-interval analysis indicated that neither the deviance-position nor the task-relevance effects could be attributed to N2b contamination.) In the case of deviance position, the intensity-deviant MMN showed the expected result: a larger response for the both position and a smaller response for at least one of the single-position deviants. The single-position deviant that should show the smallest response, of course, is the one that proved more difficult to discriminate from the standard stimulus, i.e., the second-position intensity deviant. That is not to claim that the virtual equality of the MMN amplitudes for the first and both positions was predicted, but merely that the first-position MMN amplitude should have been in the amplitude range determined by the second- and both-position MMNs. Only an order of response amplitude that clearly violated this physical/behavioral order could be taken as evidence that some other factor were controlling the MMN amplitude.

In the case of task relevance, the intensity deviant MMNs had greater amplitudes when the intensity feature was the target type. It is important not to confuse the task relevance of the deviant feature with whether or not the deviant was the target. The target-type deviant was a target or a nontarget, depending on whether it was in the attended or unattended channel. Indeed, the attended nontarget-type deviant was more easily confused with the target. Both subjec-

tive report and false-alarm analysis showed that virtually all of the false-alarm responses were generated by attended nontarget-type deviants, not unattended target-type deviants, indicating that the task-relevance effect was not due to false-alarm or aborted-false-alarm behavioral responses.

Although there was no particular reason to expect that intensity-deviant MMN amplitudes would be larger for target-type compared to nontarget-type deviants, the result is not surprising and is consistent with the attention effect discussed above—target-type deviants and deviants in the attended channel generate higher-amplitude MMNs. It is also interesting that the task relevance effect occurred independent of attention and independent of whether the deviance was in the first, second or both position. From this perspective, the effect of task relevance for the intensity deviants was similar to the effect of attention. It is as if the MMN generator can be set to focus on processing intensity because intensity defines the target. Indeed, it might make sense to adopt this processing-focus explanation, merely as attention focused along the intensity-feature dimension, if the effect of task relevance had not been so different in the case of the phoneme feature.

Few previous MMN studies have manipulated the task relevance of the deviant stimuli, even in the more limited context of targets versus nontargets in a single auditory channel. In a pioneering MMN study, Näätänen *et al.* (1978) found no difference between MMNs for higher- and lower-frequency deviants (relative to an intermediate standard frequency), depending on which deviant frequency was the behavioral target. Although there are many differences between the methods of Näätänen *et al.* (1978) and the present study, the most striking is that their target and nontarget deviants were within the same feature (pure-tone frequency), while the present experiment employed very different features—phoneme and intensity. Another salient difference is the behavioral task: counting the targets in Näätänen *et al.* (1978) and a speeded button press response to each target here. Two other studies have reported at least a trend in the direction of larger MMNs for target-type compared to nontarget-type deviants (as was found here for the intensity deviants). Nordby *et al.* (1988a) used pure tones in which the deviant stimulus either was a different frequency or was delivered at a shorter ISI. Target and nontarget frequency deviants produced the same MMN amplitudes, while the amplitude of the MMN for the ISI nontarget deviant at Fz was 2 μ V less than that for the ISI target. (This difference was not subjected to any direct statistical tests and none of the tests performed revealed it as part of a significant interaction.) Alain and Woods (1997) studied the effect of attention on the MMN to tonal-pattern and intensity deviants, and found significant differences between the MMNs for targets and attended nontargets. Since no differential analysis of responses to unattended target- and nontarget-type deviants was done to show whether or not a similar effect had occurred in the unattended channel, it is not known whether their effect was target specific or was a task-relevance effect as in the present study. Further study of the effect of task relevance on the MMN would be of interest to determine how top-down cog-

nitive control can influence the MMN for both speech and nonspeech stimulus features.

C. Phoneme deviants

In contrast to the MMN amplitudes for intensity deviants, phoneme-deviant MMNs showed no significant amplitude variation with either deviance position or task relevance alone. For the phoneme feature, it was the interaction of task relevance and deviance position that was significant and also very interesting. (In the phoneme deviants, the early-interval analysis indicated that the deviance-position-by-task-relevance interaction could not be attributed to N2b contamination.) When the phoneme deviants were target type, the MMN amplitudes for the different deviance positions were consistent with physical differences and detectability. When the phoneme deviants were nontarget type, however, the MMN amplitudes for the different deviance positions were consistent with phonetic differences and clearly inconsistent with physical differences and detectability. This interaction indicates that the properties of the phonetic processing were adjusted to the demands of the behavioral task. When /g/ was the target type, the focus of phonetic processing was on the presence or absence of the phoneme /g/, and the detectability of the /g/ largely determined the MMN amplitude. Only in the case of the second-position phoneme deviant (/ibgi/) would the occurrence of the /g/ signal both the presence of the /g/ target and the presence of the VCCV instead of the VCV. This would account for the relatively large second-position MMN amplitude for the target-type /ibgi/, in spite of its lower detectability.

The situation was quite different when /B/ was the target. The phonetic processing had no particular focus in this case because there was no deviant-phoneme target—the target was just the standard phoneme at a higher intensity. Without a particular focus, complete automatic phoneme processing occurred and the full phonetic difference determined the MMN amplitude. Thus, the occurrence of the deviant /g/ and the possible occurrence of the deviant VCCV combined to contribute to the MMN, yielding the smallest MMN amplitude for the /igi/ where no VCCV component was present. The largest MMN amplitude was obtained for /igbi/, where the /g/ in the first position contributed to the MMN and then the /b/ in the second position established the VCCV, which further contributed to the MMN. An intermediate MMN amplitude might be expected for the nontarget-type /ibgi/ because there was no deviance through the first position, and then the second position established the presence of the /g/ and the VCCV simultaneously. There have been few other studies of double deviants (Czigler and Winkler, 1996; Levänen *et al.*, 1993; Schröger, 1995; Winkler *et al.*, 1998) and none involving speech sounds or stimuli where one feature change produced double deviance, as in the present study. The data in all of the other studies showed at least a trend for double deviants to yield larger MMNs under some conditions. Only Schröger (1995, measuring MMNs) and Levänen *et al.* (1993, magnetic mismatch) used double deviants with temporally simultaneous deviation of two different features and both found double the mismatch response for the double deviant. In the other two

studies, the onset of deviance for the two features was separated by at least 100 ms (longer than the time between the first and second deviance positions in the present study) and the MMNs reflected less deviance additivity.

Independent of exactly how different components of deviance might combine to produce the MMN, the explanation of the present results in terms of /b-/g/ and VCV-VCCV differences would yield one prediction. The task relevance effect on phoneme deviants should be minimal for the second deviance position, as was seen in Fig. 6. Whether the processing was focused on detecting /g/ (as hypothesized in target-type mode) or was running on automatic (nontarget-type mode), the second-position /g/ signaled its own presence as well as the presence of the VCCV. Thus, MMNs for target- and nontarget-type second-position phoneme deviants should be generated by the simultaneous occurrence of /g/ and VCCV deviation. The situation was different for the first-position /g/. In target-type mode, detecting the first-position /g/ inhibited or masked further processing and only the /g/ deviation contributed to the MMN; while in nontarget-type mode, the sequential occurrence of /g/ and VCCV deviations both contributed to the response.

D. Phoneme-specific MMN

A fundamental question in the present research concerns the nature of the MMN response to speech stimuli. Is the MMN response to deviant speech sounds primarily due to differences detected before or after speech-specific phoneme processing? If the phoneme MMN were due to differences in a general spectral-energy representation of auditory stimuli, prior to phoneme processing, then it could be used to measure the information content of that spectral-energy representation. Such a measure would be very useful in a theoretical context for understanding the nature and information content of the spectral-energy representation. It also would be of clinical value in the case of hearing impairment due to cochlear (or other subcortical) damage. On the other hand, if the phoneme MMN were due to differences in phoneme processing, then it should be of primary use in studying speech perception and speech-perception deficits.

Previous studies have suggested that the MMN may or may not be speech specific. Sharma *et al.* (1993) reported similar MMNs for speech sounds with similar physical differences, whether or not they were exemplars of different phonemes. Sharma and Dorman (1998) also recorded MMNs for different exemplars of the vowel /i/. Maiste *et al.* (1995) found no discontinuity in MMN amplitudes to indicate the location of the category boundary on the /ba-/da/ continuum. In other studies using speech stimuli chosen to emphasize phonetic contributions to the MMN (Aaltonen *et al.*, 1997; Dehaene-Lambertz, 1997; Näätänen *et al.*, 1997; Tremblay *et al.*, 1997), results suggest a phoneme-specific MMN. As noted in the Introduction, however, even these MMNs could have resulted from either pre- or postphonemic differences. The same conclusion would apply to the demonstration of the McGurk effect in the magnetic analog to the MMN (Sams *et al.*, 1991), as long as the McGurk effect could be due to visual modification of prephonetic auditory information.

In contrast, it seems more difficult to explain the present results for the nontarget-type phoneme deviants on the basis of strictly prephonemic differences. In terms of the spectral-energy representation of the phoneme deviants, the both position differed maximally from the standard—as much as the first- and second-position differences combined. Furthermore, the behavioral results and the MMN amplitudes for target-type phoneme deviants were generally consistent with the idea that differences in spectral-energy representations determine both behavioral discrimination and MMN amplitude. The MMN response to the second-position target-type phoneme deviant was larger than expected, but that single point would not be very important without considerable other evidence for the phoneme-specific MMN. This other evidence is provided by the MMN amplitudes for nontarget-type phoneme deviants. Here, the MMN amplitudes were clearly inconsistent with the physical magnitude of stimulus differences and with the behavioral performance. Furthermore, they showed the pattern of the perceived-phoneme differences, where the combination of the presence of the /g/ and the phonetic doubling for the central consonant (VCV to VCCV) can convert a smaller physical/behavioral difference into a larger phonetic difference. The hemispheric asymmetry in the deviance position effect further supported a phonetic source for this MMN, as does the larger than expected MMN for the second-position phoneme deviant even when it is the target type. It is not apparent how an explanation that relies on acoustic differences and excludes the phonetic VCV-VCCV distinction could account for these results.

Connolly and Phillips (1994; Connolly *et al.*, 1995) have reported a phonetic mismatch negativity (PMN) in a paradigm that included meaningful sentences rather than repeated standard syllables with rare phoneme-deviant syllables. The PMN occurred in response to a phonetic mismatch between the initial phoneme of an expected word and that of a substituted, semantically appropriate word with a different initial phoneme. From one perspective, the only difference between the PMN and the MMN of the present study is in how the “expected” phoneme is defined—by linguistic context or by repetition. Winkler *et al.* (1996) demonstrated that only a single exemplar of a pure-tone standard was needed to reinstate its memory-trace across a 9.5 s gap filled with random tones. Perhaps, in the case of meaningful speech, the mere expectation of a phoneme may be sufficient to load the memory trace. Under normal circumstances, such a process would facilitate the understanding of speech by reducing the time spent on phoneme identification.

IV. CONCLUSIONS

Attention can have a major impact on MMN amplitudes, but any explanation must be able to produce a simple, non-interactive increase in the MMN for attended stimuli. A model in which attention improves the memory trace of the standard stimulus would account for these results and also offer an explanation for the many differences in attention effects seen in previous studies.

The task relevance of a deviant feature also affects the MMN amplitude, but this effect and the presence or absence of interactions was different for intensity- and phoneme-

deviant features. These findings indicate that MMN generation is much more complex than a simple preattentive detection of physical deviance.

The magnitudes of perceived phoneme differences may control MMN amplitudes, even when these magnitudes are inconsistent with physical differences and discrimination differences measured in the same listeners during the same experiment. The existence of the phonetic MMN undermines the use of the MMN in measuring the spectral-energy representation of speech sounds in the central auditory pathways and supports its use in studies of speech perception.

ACKNOWLEDGMENTS

The work described above was supported by the Department of Veterans Affairs and by NS32893. The authors thank Lydia Peters for her technical assistance in data collection, and Claude Alain, Kimmo Alho, Keith Ogawa, and Steven Thomas for their helpful comments.

¹Of the initial subject pool of 28, 5 failed to reach criterion performance in the training session and another six failed to complete both recording sessions. In the process of ERP averaging, data of 4 of the remaining 17 subjects showed excessive artifact rejections, leaving the final 13 subjects whose data were included in the analysis. Although less than 50% of the initial subjects produced useable data, there should be no concern about this producing a biased sample population because the loss at the two critical points (5 of 28 and 4 of 17) is not excessive for those situations. All of the subjects were recruited from outside of the laboratory, none knew the purpose of the experiment, and none had previous experience in auditory, or more than one visual, ERP experiment.

²There is some ambiguity concerning the appropriate terminology for these stimuli because phonetic and physical representations seem to conflict. In phonetic terms, /ibi/ and /igi/ are VCVs and should not be called /ibbi/ and /iggi/, while /ibgi/ and /igbi/ are VCCVs and should be so called. In physical terms, /ibi/ and /igi/ differ during both VC and CV formant transitions, suggesting that /ibbi/ and /iggi/ might be used to emphasize this physical difference and the fact that both /ibgi/ and /igbi/ are physically intermediate between /ibi/ and /igi/. The phonetic terminology will be used here because the purpose of the experiment was to determine the role of phonetic differences in MMN generation and because the phonetic terminology makes it easier to explain the phonetic hypothesis that /ibgi/ and /igbi/ deviants might generate larger MMN responses than the /igi/ deviant. For consistency, similar terminology will be used for the intensity deviants (/ibBi/, /iBbi/, and /iBi/).

³Indeed, the fact that the onset of deviance occurred at 60 ms (both and first positions) or 130 ms (second position) indicates that a few of the durations of significant negativity must include initial points of spurious significance (type I errors) that happen to be adjacent to the MMN. Note, for example, that the duration of significant negativity for the attended target-type /iBbi/ begins at 80 ms—only 20 ms after deviance onset. Comparing the attended target-type waveforms for /iBbi/ and /iBi/, which are plotted in Fig. 9(a) for another purpose, supports the suggestion that the initial negativity in the difference wave for /iBbi/ is not a part of the MMN response for that stimulus. The difference waves for /iBbi/ and /iBi/ are very similar in the range 150–300 ms, while the difference wave for /iBbi/ has a seemingly independent negativity in the range 80–150 ms.

⁴Näätänen *et al.* (1993) found a similar attentional modulation of the pure-tone frequency MMN when the target was defined by a difficult-to-detect intensity deviant. Since virtually no intensity MMN was found unless the intensity deviant was the target, there was no measurement of the attentional modulation of the intensity MMN when the frequency deviant was the target. The lack of MMN for difficult-to-detect nontarget intensity deviants is consistent with Woldorff *et al.* (1991). Näätänen *et al.* (1993) attribute the MMN-like component in the intensity-target difference waves to the *N2b* component generated by target stimuli.

⁵The time intervals used for the early-interval analysis were 150–250 ms for both- and first-position intensity and phoneme deviants, 200–300 ms for second-position intensity deviants, and 220–320 ms for second-position phoneme deviants. All but 4 of the 24 mastoid difference waves (corre-

sponding to the conditions for the 24 Fz difference waves of Figs. 7 and 9) showed a positive average voltage in these early intervals and these four exceptions all showed a positive voltage shift within the interval, but not sufficiently so to result in a positive average voltage over the entire interval.

- Aaltonen, O., Eerola, O., Hellström, Å., Uusipaikka, E., and Lang, A. H. (1997). "Perceptual magnet effect in the light of behavioral and psychophysiological data," *J. Acoust. Soc. Am.* **101**, 1090–1105.
- Alain, C., and Woods, D. L. (1994). "Signal clustering modulates auditory cortical activity in humans," *Percept. Psychophys.* **56**, 501–516.
- Alain, C., and Woods, D. L. (1997). "Attention modulates auditory pattern memory as indexed by event-related brain potentials," *Psychophysiology* **34**, 534–536.
- Alho, K. (1995). "Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes," *Ear Hear.* **16**, 38–51.
- Alho, K., Woods, D. L., and Algazi, A. (1994). "Processing of auditory stimuli during auditory and visual attention as revealed by event-related potentials," *Psychophysiology* **31**, 469–479.
- Alho, K., Woods, D. L., Algazi, A., and Näätänen, R. (1992). "Intermodal selective attention. II. Effects of attentional load on the processing of auditory and visual stimuli in central space," *Electroencephalogr. Clin. Neurophysiol.* **82**, 356–368.
- Bregman A. S. (1990). *Auditory Scene Analysis* (MIT, Cambridge, MA).
- Cowan, N. (1995). *Attention and Memory. An Integrated Framework* (Oxford U. P., Oxford).
- Connolly, J. F., and Phillips, N. A. (1994). "Event-related potential components reflect phonological and semantic processing of terminal word of spoken sentence," *J. Cogn. Neurosci.* **6**, 256–266.
- Connolly, J. F., Phillips, N. A., and Forbes, K. A. K. (1995). "The effects of phonological and semantic features of sentence-ending words on visual event-related potentials," *Electroencephalogr. Clin. Neurophysiol.* **94**, 276–287.
- Czigler, I., and Winkler, I. (1996). "Preattentive auditory change detection relies on unitary sensory memory representations," *NeuroReport* **7**, 2413–2417.
- Dehaene-Lambertz, G. (1997). "Electrophysiological correlates of categorical phoneme perception in adults," *NeuroReport* **8**, 919–924.
- Kraus, N., McGee, T., Sharma, A., Carrell, T. D., and Nicol, T. (1992). "Mismatch negativity event-related potential elicited by speech stimuli," *Ear Hear.* **13**, 158–137.
- Kraus, N., McGee, T. J., Carrell, T. D., and Sharma, A. (1995a). "Neurophysiological bases of speech discrimination," *Ear Hear.* **16**, 19–37.
- Kraus, N., McGee, T. J., Carrell, T. D., King, C., Tremblay, K., and Nicol, T. (1995b). "Central auditory system plasticity associated with speech discrimination training," *J. Cogn. Neurosci.* **7**, 25–32.
- Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., and Koch, D. B. (1996). "Auditory neurophysiologic responses and discrimination deficits in children with learning problems," *Science* **273**, 971–973.
- Levänen, S., Hari, R., McEvoy, L., and Sams, M. (1993). "Responses of the human auditory cortex to changes in one versus two stimulus features," *Exp. Brain Res.* **97**, 177–183.
- Maiste, A. C., Wiens, A. S., Hunt, M. J., Scherg, M., and Picton, T. W. (1995). "Event-related potentials and the categorical perception of speech sounds," *Ear Hear.* **16**, 68–90.
- Näätänen, R. (1990). "The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function," *Behav. Brain Sci.* **13**, 201–288.
- Näätänen, R. (1992). *Attention and Brain Function* (Erlbaum, Hillsdale, NJ).
- Näätänen, R. (1995). "The mismatch negativity: A powerful tool for cognitive neuroscience," *Ear Hear.* **16**, 6–18.
- Näätänen, R., and Alho, K. (1995). "Mismatch negativity—A unique measure of sensory processing in audition," *Int. J. Neurosci.* **80**, 317–337.
- Näätänen, R., Gaillard, A. W. K., and Mäntysalo, S. (1978). "Early selective-attention effect reinterpreted," *Acta Psychol.* **42**, 313–329.
- Näätänen, R., Paavilainen, P., Tittinen, H., Jiang, D., and Alho, K. (1993). "Attention and mismatch negativity," *Psychophysiology* **30**, 436–450.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., and Alho, K. (1997). "Language-specific phoneme representations revealed by electric and magnetic brain responses," *Nature (London)* **385**, 432–434.

- Nordby, H., Roth, W. T., and Pfefferbaum, A. (1988a). "Event-related potentials to time-deviant and pitch-deviant tones," *Psychophysiology* **25**, 249–261.
- Nordby, H., Roth, W. T., and Pfefferbaum, A. (1988b). "Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals," *Psychophysiology* **25**, 262–268.
- Oades, R. D., and Dittmann-Balcar, A. (1995). "Mismatch negativity (MMN) is altered by directing attention," *NeuroReport* **6**, 1187–1190.
- Picton, T. W. (1995). "The neurophysiological evaluation of auditory discrimination," *Ear Hear.* **16**, 1–5.
- Ponton, C. W., and Don, M. (1995). "The mismatch negativity in cochlear implant users," *Ear Hear.* **16**, 131–146.
- Sams, M., Aulanko, R., Aaltonen, O., and Näätänen, R. (1990). "Event-related potentials to infrequent changes in synthesized phonetic stimuli," *J. Cogn. Neurosci.* **2**, 344–357.
- Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). "Auditory frequency discrimination and event-related potentials," *Electroencephalogr. Clin. Neurophysiol.* **62**, 437–448.
- Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T., and Simola, J. (1991). "Seeing speech: Visual information from lip movements modifies activity in the human auditory cortex," *Neurosci. Lett.* **127**, 141–145.
- Sandridge, S., and Boothroyd, A. (1996). "Using naturally produced speech to elicit mismatch negativity," *J. Am. Acad. Audiol.* **7**, 105–112.
- Schröger, E. (1995). "Processing of auditory deviants with changes in one versus two stimulus dimension," *Psychophysiology* **32**, 55–65.
- Schröger, E. (1997). "On the detection of auditory deviations: a pre-attentive activation model," *Psychophysiology* **34**, 245–257.
- Schröger, E., Tervaniemi, M., Wolff, C., and Näätänen, R. (1996). "Pre-attentive periodicity detection in auditory patterns as governed by time and intensity information," *Cogn. Brain Res.* **4**, 145–148.
- 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.
- Sharma, A., Kraus, N., McGee, T., Carrell, T., and Nicol, T. (1993). "Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potentials," *Electroencephalogr. Clin. Neurophysiol.* **88**, 64–71.
- Trejo, L. J., Ryan-Jones, D. L., and Kramer, A. F. (1995). "Attentional modulation of the mismatch negativity elicited by frequency differences between binaurally presented tone bursts," *Psychophysiology* **32**, 319–328.
- Tremblay, K., Kraus, N., Carrell, T. D., and McGee, T. (1997). "Central auditory system plasticity: Generalization to novel stimuli following listening training," *J. Acoust. Soc. Am.* **102**, 3762–3773.
- Winkler, I., Cowan, N., Csépe, V., Czigler, I., and Näätänen, R. (1996). "Interactions between transient and long-term auditory memory as reflected in the mismatch negativity," *J. Cogn. Neurosci.* **8**, 403–415.
- Winkler, I., Czigler, I., Jaramillo, M., Paavilainen, P., and Näätänen, R. (1998). "Temporal constraints of auditory event synthesis: Evidence from ERPs," *NeuroReport* **9**, 495–499.
- Woldorff, M. G., Hackley, S. A., and Hillyard, S. A. (1991). "The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones," *Psychophysiology* **28**, 30–42.
- Woods, D. L., and Clayworth, C. W. (1985). "Click spatial position influences middle latency auditory evoked potentials (MAEPs) in humans," *Electroencephalogr. Clin. Neurophysiol.* **60**, 122–129.
- Woods, D. L., Alho, K., and Algazi, A. (1992). "Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli," *Electroencephalogr. Clin. Neurophysiol.* **82**, 341–355.
- Woods, D. L., Alho, K., and Algazi, A. (1993). "Intermodal selective attention: Evidence for processing in tonotopic fields," *Psychophysiology* **30**, 287–295.
- Woods, D. L., Hillyard, S. A., and Hansen, J. C. (1984). "Event-related brain potentials reveal similar attentional mechanisms during selective listening and shadowing," *J. Exp. Psychol. Hum. Percept. Perform.* **10**, 761–777.

Loudness recalibration as a function of level

Dan Mapes-Riordan^{a)} and William A. Yost

Parmly Hearing Institute, Loyola University Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626

(Received 30 November 1998; revised 19 May 1999; accepted 11 August 1999)

Recent research on loudness has focused on contextual effects on loudness, both assimilation and recalibration. The current experiments examined loudness recalibration [Marks, *J. Exp. Psychol.* **20**, 382–396 (1994)]. In the first experiment, an adaptive tracking procedure was used to measure loudness recalibration as a function of standard- and recalibration-tone level. The standard-tone frequencies were 500 and 2500 Hz and the levels were 80-, 70-, 60-, and 40-dB SPL, and threshold. Seventeen dB of loudness recalibration was obtained (combined over both frequencies) in the 60-dB SPL condition. This amount of loudness recalibration, while substantial, is still less than that obtained by Marks (22 dB), using the method of paired comparisons. The second experiment sought to duplicate Marks' earlier experiment [Marks, *J. Exp. Psychol.* **20**, 382–396 (1994), experiment 2]. The results of this experiment (21 dB) were almost identical to those obtained by Marks. The results of experiment 1 indicate that loudness recalibration is maximum when the recalibration tone is moderately louder than the subsequent standard tones. Relatively little loudness recalibration is exhibited when the standard-tone level equals the recalibration-tone level. In addition, there is no loudness recalibration at threshold. The tracking procedure also identified that the onset of loudness recalibration is very rapid. The difference between the maximum loudness recalibration obtained at each frequency (11 dB at 500 Hz, 6 dB at 2500 Hz) suggests that loudness recalibration is dependent upon the frequency of the standard tone. © 1999 Acoustical Society of America. [S0001-4966(99)01312-0]

PACS numbers: 43.66.Ba, 43.66.Cb, 43.66.Mk [RVS]

INTRODUCTION

The perception of loudness is contextual; that is, the loudness of a sound depends on the temporal vicinity of other sounds. Two types of contextual effects on loudness are assimilation and recalibration. Loudness assimilation occurs when the loudness of a target sound is “pulled toward” the loudness of a neighboring sound, as compared to the loudness of the target sound in isolation. Over the years, two distinct ways of conceptualizing loudness assimilation have emerged. The first are “central tendency” effects. As suggested by the name, this view of assimilation derives from the tendency of psychophysical judgments to regress toward the mean (Hollingworth, 1910). Central-tendency-based loudness assimilation is considered to be a manifestation of response judgment assimilation (e.g., Cross, 1973; Jesteadt, Luce, and Green, 1977; Marks, 1993; Ward, 1979, 1990). The other loudness assimilation effect is “loudness enhancement” (Galambos *et al.*, 1972; Irwin and Zwislowski, 1971). Loudness enhancement effects are thought to be due to memory and/or loudness integration interference (Elmasian, *et al.* 1980; Plack, 1996). Loudness enhancement is created using relatively short-duration tones (typically less than 30 ms) and silent intervals (approximately 100 ms) that are within the region of temporal integration of loudness. In contrast, central-tendency loudness-assimilation experiments use relatively long-duration tones (typically 500 ms) and silent intervals (typically 1–2 s) that exceed the temporal extent of loudness integration. Loudness enhancement is a larger effect (up to 15 dB), and is also frequency dependent (Galambos

et al., 1972; Zwislowski and Sokolich, 1974) while central-tendency loudness assimilation is smaller (approximately 6 dB) and is frequency independent.

Loudness recalibration (Marks, 1994) occurs when loud (recalibration) tones at one frequency (f_1) are followed by relatively quiet tones; one at the same frequency (f_1) and another at a different frequency (f_2). The resulting perception is that the quiet tone at frequency f_1 is relatively quieter than if it had not been preceded by the louder recalibration tone. Loudness recalibration is a one-way, fatigue-like effect similar to loudness adaptation (see Scharf, 1983) in that loud sounds make quiet sounds even quieter, but quiet sounds do not make loud sounds louder.

Loudness recalibration has been studied primarily by Marks (1988, 1992, 1993, 1994, 1996; Marks and Warner, 1991). In this first published work on the phenomenon, Marks (1988) used magnitude estimates of loudness to show that the loudness between two alternating tones can shift by up to 18 dB. In 1991, Marks and Warner showed that loudness recalibration depends on the frequency difference between two tones. Loudness recalibration diminishes as the tones are made less than a critical band apart. A year earlier, Schneider and Parker (1990) were the first to show loudness recalibration using a direct loudness comparison procedure, which is known to be less susceptible to response bias than magnitude estimation procedures. This result was important since, at that time, Algorn and Marks (1990) suggested that loudness recalibration was due to listener response bias, much like central-tendency assimilation. Schneider and Parker suggested that loudness recalibration is due to a gain-control mechanism whose locus is close to the periphery,

^{a)}Electronic mail: dmapes@phi.luc.edu

rather than the result of a more central response bias. Marks (1992) also used direct loudness comparisons producing results in agreement with Schneider and Parker (1990) in that the perception of loudness was directly modified and was not the result of listener response bias. At this time Marks also showed that loudness recalibration is dynamic, varying over time depending on recent stimulus history. In 1993, Marks demonstrated that the contextual shift of loudness is assimilative when tones at only one frequency are presented (using a magnitude estimation procedure) and is contrastive when tones at two different frequencies are compared (using a paired-comparison procedure).

In 1994, using a set of seven experiments consisting of both magnitude estimation and paired comparison procedures, Marks showed that:

- (1) Loudness recalibration depends on the frequency difference between tones with little or no recalibration with close frequencies and maximum effect when the frequency spacing between the tones is beyond a critical bandwidth.
- (2) Given a pair of frequencies spaced just beyond a critical bandwidth (1002 and 1248 Hz), and a set of levels known to generate loudness recalibration, increasing the intensity of the tones eliminated recalibration.
- (3) After increasing the frequency spacing to over 2 octaves (500 and 2500 Hz), increasing the intensity produced little reduction in recalibration.

The above evidence suggests that the effect may derive from changes in the outputs of frequency-specific channels; channels that are subsequently used to derive our perception of loudness (Marks, 1994, 1996). Since at least a portion of the effect is also produced contra laterally (Marks, 1996), its locus must be beyond the point of binaural convergence.

Most of the research on loudness recalibration has been concerned with determining the conditions that produce the effect (e.g., Marks, 1988, 1992, 1993; Marks and Warner, 1991). For example, in most loudness-recalibration studies listeners were presented a range of stimulus levels using the method of paired comparisons (see Stevens, 1951). Although this procedure is useful for detecting the presence of loudness recalibration and has yielded knowledge of frequency dependencies, it is difficult to determine how the magnitude of loudness recalibration varies with stimulus intensity and with time of exposure. The purpose of the present experiments was to determine how the amount of loudness recalibration varies as a function of standard-tone and recalibration-tone level. In addition, these experiments directly compare the amount of loudness recalibration obtained using adaptive tracking and paired-comparison procedures. To accomplish this, a randomized, adaptive two-track procedure (Jesteadt, 1980) was used to make direct loudness comparisons. During a latter portion of this procedure, a steady-state amount of loudness recalibration was created by preceding each set of comparison tones with a louder, recalibration tone. The amount of loudness recalibration was then taken as the difference between the average loudness levels obtained in the baseline and the recalibration phases of the trial sequence. The tracking procedure also has the potential

of measuring the temporal properties of the onset of loudness recalibration.

I. EXPERIMENT 1—LOUDNESS RECALIBRATION VERSUS LEVEL USING THE ADAPTIVE TRACKING METHOD

A. Method

Eleven naïve, paid listeners participated in the experiment. They were all college undergraduate students between the ages of 19 and 25 who self-reported to have normal hearing. Prior to the start of the experiment, each listener was given written and verbal instructions describing the task they were to perform. The entire sequence of 160 trials lasted approximately 9 min. The threshold procedure (80 trials) lasted approximately 6 min. In a typical session, listeners worked in pairs, taking turns running the experiment to allow each to have a break. When a listener could not be paired, they were instructed to take at least a 5 min break before beginning another experimental run. In a typical day, each listener ran three or four randomly assigned conditions over a 90-min period. Each listener ran each of the eight conditions at least ten times over the entire experiment. There were no practice sessions.

The experiment was performed in IAC sound-isolation booths. The test tones were created using Tucker-Davis instrument modules PA4, PF1, and DD1, and presented through Sennheiser HD-520 II headphones. The test tones were digitally generated using a sampling frequency of 16 kHz and antialiasing filters with a cutoff frequency of 7 kHz. The equipment was controlled with a PC-compatible computer (Gateway P5-75). The computer also recorded listener responses and provided visual reinforcement during the experiment.

A randomized adaptive two-track (one-up two-down, two-up one-down) procedure (Jesteadt, 1980) was used to estimate the equal loudness-level point on the psychometric function. Each trial consisted of a fixed-level standard tone at one frequency followed by a comparison tone at another frequency whose level varied from trial-to-trial depending on the response of the listener. For the one-up two-down track, if the listener indicated that the comparison tone was quieter than the standard tone, the level of the comparison tone was raised in the subsequent trial for that track. The level of the comparison tone was lowered if the listener indicated it was louder than the standard tone in two consecutive trials for that track. The action of the two up-one down track was similar to the one up-two down track except that it took two consecutive responses indicating that the comparison tone was quieter in order to increase its level on the next trial, but only one response indicating that the comparison tone was louder to decrease its level. During the experiment, each trial was randomly chosen from either the one up-two down or the two up-one down sequence. The 71-percent point on the psychometric function was estimated by averaging the reversal points (i.e., the points where the comparison tone tracking level changed direction) in the one up-two down sequence (Levitt, 1971). Similarly, averaging the reversal points in the two up-one down sequence produced an esti-

mate of the 29-percent point on the psychometric function. The point of equal loudness (i.e., 50-percent point on the psychometric function) was estimated by averaging the 71-percent and 29-percent points.

Listeners compared the loudness of 500- and 2500-Hz diotic tones presented over headphones. An experimental sequence began with 80 baseline trials (40 for each track). During each baseline trial, the standard tone was presented followed by the comparison tone. The duration of each tone was 500 ms with 10-ms cosine² onset and offset windowing. A 500-ms silent gap was inserted between tones. After the second tone, listeners were asked which tone was louder. Listeners responded by pressing “1” on the keyboard if they thought that the first tone was louder or “2” if they thought the second tone was louder. A 1-s delay was inserted after each response before the next trial began. The 80 baseline trials were immediately followed by 80 additional trials (40 per track) in which a recalibration tone preceded each loudness comparison. The duration of the recalibration tone was 1 s, and was followed by 1 s of silence before the start of the standard tone. Otherwise, the recalibration trials were identical to the baseline trials. The frequency of the recalibration tone was identical to the standard tone.

A total of 12 conditions was tested. In half of the conditions the 500-Hz tone was the fixed standard while the 2500-Hz comparison tone roved, and vice versa in the other half of the conditions. In 10 of the 12 conditions, the level of the recalibration tone was 80 dB SPL and the level of the standard tone was at threshold, 40-, 60-, 70-, or 80-dB SPL. In the other two conditions, the recalibration-tone level was 40 dB SPL and the standard-tone level was 80 dB SPL. Seven of the 11 listeners were tested in the threshold, 40-, and 80-dB SPL standard-tone conditions as well as the 40-dB SPL recalibration-tone conditions. The remaining four listeners completed the 40-, 60-, 70-, and 80-dB SPL standard-tone conditions. The results of the overlapping 40- and 80-dB SPL standard-tone conditions were combined.

At the beginning of each experimental sequence, the level of the comparison tone was set 10 dB above the standard-tone level for the upper level (one up-two down) sequence and 10 dB below the standard-tone level for the lower level (two up-one down) sequence. The step size of the comparison tone was initially set to 4 dB and was then reduced to 2 dB after four reversals. After each experimental run, the average loudness level was separately computed for the baseline sequence and the recalibration sequence. The amount of loudness recalibration for each experimental run was the average baseline sequence level minus the average recalibration sequence level. One loudness-recalibration point was computed for each pair of baseline and recalibration trials. The loudness just-noticeable-difference (jnd) for both the baseline and recalibration sequences was calculated as the difference between the 71-percent and 29-percent points divided by 2 [i.e., the average distance between the mean sequence level and the mean upper (or lower) track level].

A slightly different procedure was used to measure loudness recalibration at threshold. On every trial, two 1-s tone intervals were presented (separated by 1 s of silence) in

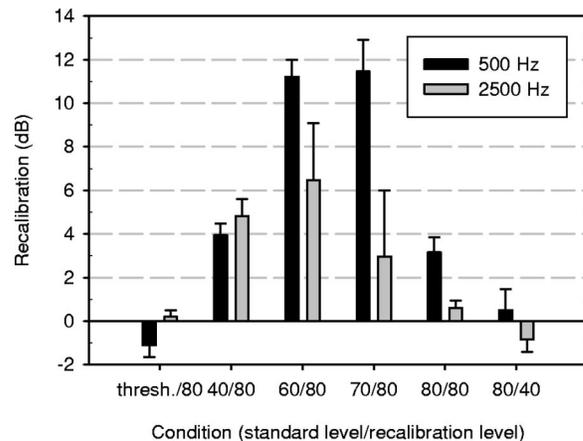


FIG. 1. Average loudness recalibration (dB) and standard error (error bars) for the 12 standard- and recalibration-tone level conditions in experiment 1: threshold with 80-dB SPL recalibration tone (seven listeners), 40-dB SPL standard tone with 80-dB SPL recalibration tone (11 listeners), 60-dB SPL standard tone with 80-dB SPL recalibration tone (four listeners), 70-dB SPL standard tone with 80-dB SPL recalibration tone (11 listeners), and an 80-dB SPL standard tone and 40-dB SPL recalibration tone (seven listeners). In half of the conditions (dark bars), the standard- and recalibration-tone frequency was 500 Hz and the roving comparison-tone frequency was 2500 Hz. In the other four conditions (light bars), the frequencies were reversed.

which only one randomly chosen interval contained a tone. The listeners task was determine which interval contained the tone. If the listener was correct in two consecutive trials, the level of the tone was reduced in the following trial. But, if the listener was incorrect, the level of the tone was increased in the next trial. Forty baseline trials were followed by 40 trials in which a 1-s, 80-dB SPL, recalibration tone followed by 1 s of silence preceded the presentation of the two intervals. Feedback was provided via the computer screen indicating whether their responses were correct or incorrect.

B. Results

The overall mean and standard error of the mean of loudness recalibration in experiment 1 are plotted in Fig. 1. The mean jnd's are plotted in Fig. 2. Individual results are

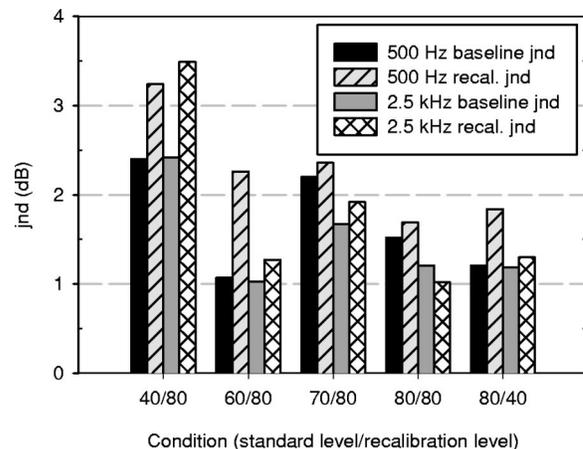


FIG. 2. Average jnd's (dB) for the 10 nonthreshold conditions shown in Fig. 1. In each condition, the average jnd during the baseline and the recalibration portions of the trials are each plotted separately.

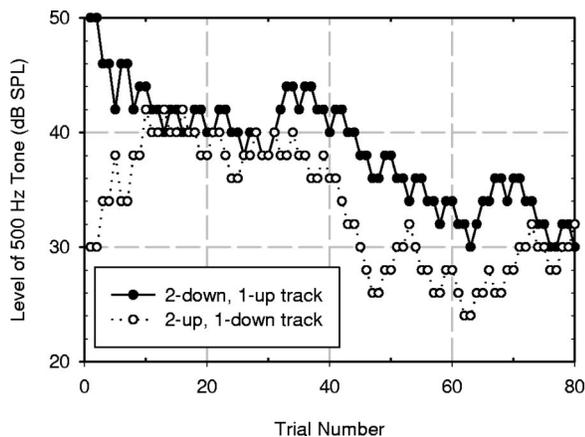


FIG. 3. An example of the results of an individual trial run in experiment 1. For this particular trial, the standard-tone frequency was 2500 Hz and was presented at 40-dB SPL. The roving comparison-tone frequency was 500 Hz. The 1-s duration, 80-dB SPL, 2500-Hz recalibration tone began to precede loudness comparisons starting at trial number 40 in each track.

not presented since the amount of interlistener variation was minimal. Individual trial results in every condition that were more than two standard deviations from the mean for that listener were removed. Fortunately, most of the data were consistent and not many trial results had to be discarded. The amount of data removal for each listener within a particular condition was approximately 10 percent when averaged over all conditions. One particular listener required 25 percent outlier removal; the remaining data was similar to the others.

Figure 1 illustrates that substantial loudness recalibration was produced in the 40-, 60-, and 70-dB SPL standard-tone conditions. For these standard-tone conditions there was an increase in jnd during the recalibration sequence as compared to the preceding baseline sequence (see Fig. 2), indicating that both the mean and jnd of loudness recalibration increased in these conditions. There was relatively little recalibration at the 80-dB SPL standard-tone condition, at threshold, or when the recalibration-tone level was less than the standard-tone level.

C. Discussion

For the range of levels tested, loudness recalibration only occurred when the level of the comparison tones was moderately less than the level of the recalibration tones. This is not too surprising since loudness recalibration would not be expected whenever 500-ms and 1-s duration tones are presented at the same level and frequency as is the case of the 80-dB SPL standard-tone condition. The results of the threshold conditions are also anticipated given that the procedure measuring loudness recalibration at threshold is identical to that used in forward-masking experiments. In relation to forward-masking experiments, the recalibration tone acts as the masking tone and the comparison tone acts as the probe tone. Such experiments do not produce a threshold shift when the temporal gap between the masker and probe tone is 1 s (see Moore, 1989), as is the case in the recalibration experiments.

An example of the temporal onset of loudness recalibration is illustrated in Fig. 3. These data are taken from a

typical trial sequence of the 40-dB SPL standard-tone condition. Note that the onset of loudness recalibration beginning at trial number 40 (as measured by the two-up one-down sequence) is very fast. In fact, it is likely that recalibration is faster than what this experimental procedure is capable of measuring. All of the trial sequences in which loudness recalibration was exhibited display the same rapid onset of loudness recalibration.

The results of this experiment demonstrate that loudness recalibration can be obtained using an adaptive-tracking procedure. Averaging the 60- and 70-dB SPL standard-tone conditions at each frequency, there was an 11-dB reduction in loudness level at 500 Hz and a 4.5-dB reduction at 2500 Hz. Summarizing earlier work on loudness recalibration, Marks (1996) claimed an average loudness recalibration effect of approximately 8 dB at each frequency. The difference between the current results at 500 and 2500 Hz suggests that the amount of loudness recalibration may depend on the frequency of test tones (with a larger portion attributable to lower frequencies) in experiments in which the amount of loudness recalibration is not separable at each frequency.

When averaged over the 60- and 70-dB SPL conditions, experiment 1 produced an *overall contextual shift* (Marks, 1992) in loudness level of 15.5 dB. Although substantial, this amount is still less than that produced by Marks. In one particular experiment (1994, experiment 2), Marks also computed an *overall contextual shift* that averaged 22 dB (combined over both frequencies) using the same frequencies as the current experiment. The next experiment is a replication of Marks' 1994 experiment 2. The purpose of this experiment was to determine if other listeners could achieve similar amounts of loudness recalibration.

II. EXPERIMENT 2—LOUDNESS RECALIBRATION USING THE METHOD OF PAIRED COMPARISONS

A. Method

In this experiment, listeners compared the loudness of 500 and 2500-Hz tones using the method of paired comparisons. In condition A, the level of the 500-Hz tone took on any of the following eight values expressed in dB SPL: 35, 40, 45, 50, 55, 60, 65, 70. The level of the 2500-Hz tone was taken from a similar set of eight levels shifted up 15 dB (50- to 85-dB SPL) relative to the 500-Hz tone levels. In condition B, the set of levels of the 500-Hz tone was increased by 20 dB (55- to 90-dB SPL) and those of the 2500-Hz tone decreased by 20 dB (30- to 65-dB SPL). The amount of loudness recalibration is the difference between the equal loudness-level points (generated from the psychometric functions) of the set of overlapping 500-Hz levels (in dB SPL, 55, 60, 65, 70) in conditions A and B. Presuming that loudness recalibration operates by attenuating lower-level tones with high-level tones within the same frequency channel, it is expected that the equal-loudness levels of the 2500-Hz tones will be relatively higher in condition A since the high-level 2500-Hz tones (70- to 85-dB SPL) will reduce the loudness of the lower-level 2500-Hz tones (50- to 65-dB SPL). In condition B, the high-level 500-Hz tones (75- to 90-dB SPL) will reduce the loudness of the lower-level

TABLE I. The results of loudness matching in experiment 2. The results are shown for the four overlapping 500-Hz conditions as indicated in column one. Columns two and three are the average equal loudness-level points for the 2500-Hz tones. The results when the level of the set of tones at 2500 Hz were high (condition A) are shown in column two. The results of the lower-level set (condition B) are shown in column three. The amount of loudness recalibration at each level of the 500-Hz tone, calculated as the difference between the equal loudness-level points in columns two and three, are shown in column four. The bottom cell in column four is the average loudness recalibration for the four conditions.

Level of 500 Hz tone (dB SPL)	Condition A, matching level of 2500-Hz tone when at high levels (dB SPL)	Condition B, matching level of 2500-Hz tone when at low levels (dB SPL)	Difference between condition A and B (loudness recalibration, dB)
70	74	54	20
65	69	47	22
60	63	43	20
55	60	39	21
			average: 20.75

500-Hz tones (55- to 70-dB SPL) of interest. Thus, the equal-loudness levels of the 2500-Hz tones should be relatively lower in condition B as compared to condition A.

During each experimental run, 384 pairs of tones were compared with each other; each randomly chosen from a set comprised of 6 replicates of the 64 possible pairings. Half of the listeners performed condition A, then condition B, and the other half performed condition B first, then condition A. Each listener had a 15-min break between conditions. There were no practice sessions.

During each trial, the order of presentation of the 500- and 2500-Hz tones was randomly chosen. The duration of each tone was 1 s with 10-ms cosine² onset and offset and 500 ms of silence between tones. After the second tone was presented, the listener was asked which tone was louder. A 1-s delay was inserted between each trial. The entire sequence of trials lasted approximately 30 min. Eight naïve, paid listeners participated in the experiment. Seven of the eight listeners in this experiment also participated in experiment 1. The new listener was also a college undergraduate student who self-reported to have normal hearing. (The results for this new listener were similar to the other, experienced listeners.) The equipment used in this experiment was identical to that used in experiment 1.

B. Results

The psychometric functions were created for conditions A and B, each data point representing the proportion of trials in which the 2500-Hz tone was judged louder than the 500-Hz tone over all of the listeners for that condition. The 50-percent point on the psychometric function for the overlapping conditions was obtained by using the fitted logistic equation to solve for the level of the 2500-Hz tone required to make the logistic equation equal 0.5. The amount of loudness recalibration was then found by taking the difference between the 50-percent points in the corresponding A and B conditions. The amounts of loudness recalibration for the four overlapping conditions are shown in Table I. Since the

interlistener variation was relatively small, only the mean values across all listeners are presented. The differences between the results of this experiment and those of obtained by Marks are small. Overall, the average loudness recalibration across the four levels of the 500-Hz tone (20.75 dB) is only 1 dB less than that obtained by Marks (21.75 dB).

C. Discussion

Experiment 2 produced more loudness recalibration than experiment 1. What factors could account for this difference? There are three primary differences in the stimuli used in experiment 1 and experiment 2:

- (1) The duration of tones in experiment 2 was longer (1000 vs 500 ms).
- (2) Experiment 2 used a larger range of levels from 30- to 90-dB SPL.
- (3) The interval between louder (recalibration) and softer (comparison) tones was longer in experiment 2. The interval also varied in duration and contained other intervening tones.

The first difference is unlikely to produce more loudness recalibration given that the loudness difference between 500- and 1000-ms tones is negligible. The second difference would not likely produce additional loudness recalibration since experiment 1 showed that the amount of loudness recalibration decreased at lower and higher standard-tone levels. However, it is possible that the higher maximum levels in experiment 2 (90- vs 80-dB SPL) could have been a factor. The third difference suggests that temporal factors (and/or the presence of intervening tones between the recalibration and standard tones) are responsible for the extra loudness recalibration in the paired-comparison experiment. Loudness recalibration could be a global effect that increases with the duration between louder recalibration tones and softer comparison tones.

III. GENERAL DISCUSSION

For the range of levels tested, loudness recalibration only occurred when the levels of the comparison tones were moderately less than the level of the recalibration tones. Relatively little recalibration was exhibited when the standard-tone level equaled the recalibration-tone level or at threshold. The fact that loudness recalibration was absent at threshold further distinguishes it from loudness adaptation.

The absence of loudness recalibration at threshold may also provide evidence as to the locus of loudness recalibration in the auditory system. For example, assume that loudness recalibration consists of a gain control-like mechanism and the noise floor of the auditory system can be modeled as additive noise. Given these assumptions, the lack of loudness recalibration at threshold implies that the dominant source of noise in the auditory system precedes the gain mechanism responsible for loudness recalibration. If the noise source followed the gain control, a gain reduction would have attenuated the signal but not the noise, thereby producing a higher threshold.

Experiment 1 demonstrated that an adaptive-tracking procedure cannot generate as much loudness recalibration as a paired-comparison procedure, except perhaps at the 500-Hz, 60- and 70-dB SPL standard-tone conditions. It is possible that the larger levels in experiment 2 (90-, 95-dB SPL) created more loudness recalibration than the 80-dB SPL recalibration-tone level used in experiments 1. The higher levels may have also produced a small amount of loudness fatigue, thereby producing more apparent recalibration. This may also explain the small (yet nonzero) amount of loudness recalibration at the 500-Hz, 80-dB SPL standard-tone condition.

Previous experiments (Marks, 1994) have shown that the amount of loudness recalibration is dependent upon the frequency difference between the comparison tones. The results of experiment 1 indicate that loudness recalibration is also dependent upon the frequency of the standard tone, since more loudness recalibration was produced in the 500-Hz conditions as compared to the corresponding 2500-Hz conditions. This suggests that the amount of loudness recalibration may not be equally attributable to each frequency as previously assumed (e.g., see Marks, 1996) in experiments like experiment 2 in which the amount of loudness recalibration at each frequency is not separable.

Although loudness recalibration and central-tendency-based loudness assimilation are manifested using tones of similar duration, they are opposite effects that may be due to separate mechanisms. Loudness recalibration may be a gain control-type mechanism that directly affects the activity in frequency-specific neural channels. Loudness assimilation, on the other hand, may be a listener response bias effect caused by memory interference or integration mechanisms; effects that are strongest when tones at the same frequency and only moderately different in level are competing for attention. The magnitude of central-tendency-based loudness assimilation may depend on how well listeners can ignore prior stimuli when making loudness comparisons. If loudness recalibration and assimilation are due to separate mechanisms, it may be difficult to measure one phenomenon independent of the other since they could be operating simultaneously. This would explain the relatively wide variability of loudness comparisons when the comparison-tone level was relatively close to the recalibration-tone level as shown for the 70-dB SPL standard-tone condition in Fig. 1. Perhaps some listeners are better than others at ignoring the interfering recalibration tone (i.e., they are less susceptible to assimilation effects) when making loudness judgments of sub-sequent comparison tones.

The adaptive-tracking procedure revealed that the temporal onset of loudness recalibration is very rapid; so rapid that the tracking procedure may not have been able to measure it. Recovery from loudness recalibration was not measured in this study. Measurements of recovery from loudness recalibration might allow one to determine if loudness recalibration is a temporally local or global effect. In the present experiments, the recalibration tone occurred on every trial, leaving the possibility that loudness recalibration may be a temporally local phenomenon. As such, the sequential order of the stimuli that occur on a trial-by-trial basis could influ-

ence the amount of loudness recalibration at a local level. It was not possible to measure such sequential effects in this study. Although we have some data (Mapes-Riordan and Yost, 1998) that supports the view that loudness recalibration is a global effect, additional research is needed to determine whether loudness recalibration is local or global.

ACKNOWLEDGMENTS

This research was supported by a program project grant from the National Institutes of Deafness and other Communication Disorders (DCP01). We thank Bertram Scharf and another, anonymous reviewer, as well as our colleagues at the Parnly Hearing Institute, for their helpful comments on this work.

- Algom, D., and Marks, L. E. (1990). "Range and regression, loudness scales and loudness processing: Toward a context-bound psychophysics," *J. Exp. Psychol.* **16**, 706–721.
- Cross, D. V. (1973). "Sequential dependencies and regression in psychophysical judgments," *Percept. Psychophys.* **14**, 547–552.
- Elmasian, R., Galambos, R., and Bernheim, A. (1980). "Loudness enhancement and decrement in four paradigms," *J. Acoust. Soc. Am.* **67**, 601–607.
- Galambos, R., Bauer, J., Picton, T., Squires, K., and Squires, N. (1972). "Loudness enhancement following contralateral stimulation," *J. Acoust. Soc. Am.* **52**, 1127–1130.
- Hollingworth, H. L. (1910). "The central tendency of judgment," *J. Phil. Psychol.* **7**, 461–469.
- Irwin, R. J., and Zwislocki, J. J. (1971). "Loudness effects in pairs of tone bursts," *Percept. Psychophys.* **10**, 189–192.
- Jesteadt, W. (1980). "An adaptive procedure for subjective judgements," *Percept. Psychophys.* **28**, 85–88.
- Jesteadt, W., Luce, R. D., and Green, D. M. (1977). "Sequential effects in judgements of loudness," *J. Exp. Psychol.* **3**, 92–104.
- Levitt, H. (1971). "Transformed up-down procedures in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Mapes-Riordan, D., and Yost, W. A. (1998). "Temporal properties of loudness recalibration," *J. Acoust. Soc. Am.* **103**, 3020.
- Marks, L. E. (1988). "Magnitude estimation and sensory matching," *Percept. Psychophys.* **43**, 511–525.
- Marks, L. E. (1992). "The contingency of perceptual processing: Context modifies equal-loudness relations," *Psychol. Sci.* **3**, 285–291.
- Marks, L. E. (1993). "Contextual processing of multidimensional and unidimensional auditory stimuli," *J. Exp. Psychol.* **19**, 227–249.
- Marks, L. E. (1994). "Recalibrating' the auditory system: The perception of loudness," *J. Exp. Psychol.* **20**, 382–396.
- Marks, L. E. (1996). "Recalibrating the perception of loudness: Interaural transfer," *J. Acoust. Soc. Am.* **100**, 473–480.
- Marks, L. E., and Warner, E. (1991). "Slippery context effect and critical bands," *J. Exp. Psychol.* **17**, 986–996.
- Moore, B. C. J. (1989). *An Introduction of the Psychology of Hearing* (Academic, London), 3rd ed., pp. 119–121.
- Plack, C. J. (1996). "Loudness enhancement and intensity discrimination under forward and backward masking," *J. Acoust. Soc. Am.* **100**, 1024–1030.
- Scharf, B. (1983). "Loudness Adaptation," in *Hearing Research and Theory*, edited by J. V. Tobias and E. D. Schubert (Academic, New York), Vol. 2, pp. 1–56.
- Schneider, B., and Parker, S. (1990). "Does stimulus context affect loudness or only loudness judgements?" *Percept. Psychophys.* **43**, 511–525.
- Stevens, S. S. (1951). "Mathematics, measurement, and psychophysics," in *Handbook of Experimental Psychology*, edited by S. S. Stevens (Wiley, New York), p. 43.
- Ward, L. M. (1979). "Stimulus information and sequential dependencies in magnitude estimation and cross-modality matching," *J. Exp. Psychol.* **5**, 444–459.
- Ward, L. M. (1990). "Critical bands and mixed-frequency scaling: Sequential dependencies, equal loudness contours, and power function exponents," *Percept. Psychophys.* **47**, 551–562.
- Zwislocki, J. J., and Sokolich, W. G. (1974). "On loudness enhancement of a tone burst by a preceding tone burst," *Percept. Psychophys.* **16**, 87–90.

Psychometric functions for gap detection in a yes–no procedure

Mary Florentine

Institute for Hearing, Speech, and Language and Department of Speech-Language Pathology and Audiology (133 FR), Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115

Søren Buus and Wei Geng

Communication and Digital Signal Processing Center, Department of Electrical Engineering (442 DA), Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115

(Received 17 August 1998; revised 2 July 1999; accepted 18 August 1999)

To examine models of temporal resolution and to investigate the decision processes underlying the detection of a brief pause in a bandpass noise, psychometric functions for gap detection were measured at octave frequencies from 0.25 to 8 kHz. Three normal listeners were tested using a constant-stimulus procedure with a cued Yes–No paradigm. The Minimum Detectable Gap (MDG) estimated from the midpoint of the psychometric functions decreased systematically with increasing frequency. The slopes of the psychometric functions generally increased as the test frequency increased up to 2 kHz, but remained constant at the higher frequencies. Two models were investigated: an energy-detector model and a loudness-detector model. Both consisted of auditory filtering, a nonlinearity, and short-term integration. In the energy-detector model, the nonlinearity was a square law. In the loudness-detector model, it was a compressive power law. Using the usual Gaussian approximations, the energy-detector model fails at low frequencies because the probability distributions of short-term energy differ from Gaussian distributions. The probability distributions of short-term loudness closely follow Gaussian distributions. The loudness-detector model predicts the frequency dependence of the MDG quite accurately, except at 0.25 kHz. It also predicts psychometric functions that resemble the data at low frequencies, but the predicted slopes increase much less with frequency than the measured slopes. This result may indicate that the onset response to the trailing marker of the gap provides an important cue for detection of gaps with durations exceeding the MDG. © 1999 Acoustical Society of America. [S0001-4966(99)02212-2]

PACS numbers: 43.66.Ba, 43.66.Mk [RVS]

INTRODUCTION

Detecting a gap in a bandpass noise is a relatively simple task that provides a frequency-selective assessment of temporal resolution in the auditory system (e.g., Zwicker and Feldtkeller, 1967; Fitzgibbons and Wightman, 1982; Fitzgibbons, 1983; Florentine and Buus, 1983; Shailer and Moore, 1983; Buus and Florentine, 1985; Eddins *et al.*, 1992). Understanding how the properties of bandpass noises and auditory processing affect this simple task is likely to be important. First, gap detection in bandpass noises may be related to speech perception in noise. Tyler *et al.* (1982) showed that detection thresholds for gaps in narrow-band noises correlated significantly with impaired listeners' ability to understand speech in noise, even after partialing out the effects of the listeners' elevated thresholds. Second, the variability of bandpass noises makes them somewhat similar to the time-varying sounds encountered in everyday listening situations. Therefore, understanding factors that limit gap detection in bandpass noise may yield insight into auditory processes that play a significant role in the perception of everyday sounds, including speech.

Although considerable knowledge has been gained by measurements of gap-detection thresholds, such measurements tell relatively little about how the cue used to detect a gap (i.e., the listeners' decision variable) depends on the duration of the gap. Psychometric functions measure this dependency and describe important aspects of the transforma-

tion from stimulus to decision variable. Therefore, measurements of psychometric functions for gap detection should provide insight into the auditory processes underlying gap detection and temporal resolution in the auditory system.

Psychometric functions also provide a thorough test of models of gap detection by examining whether they predict how the decision variable changes with increasing duration of the gap. Unfortunately, psychometric functions do not appear to be available for gap detection in bandpass noises, although they have been measured for gap detection in broadband noises (Green and Forrest, 1989; He *et al.*, 1999) and in sinusoids (Irwin and Kemp, 1976; Shailer and Moore, 1987; Moore *et al.*, 1989). Therefore, the present study investigates psychometric functions for gap detection in bandpass noises.

I. METHOD

A. Stimuli and apparatus

The gaps were carried by 85-dB-SPL bandpass noises with center frequencies (CFs) between 0.25 and 8 kHz. The bandwidth was three auditory-filter bandwidths [i.e., Equivalent Rectangular Bandwidths (ERBs) calculated as $0.11 (CF + 165 \text{ Hz})$; see Buus, 1997] and the filter slopes were essentially infinitely steep. The 786-ms bandpass noises had 20-ms rise and fall times. The gap started 250 ms after the onset of the noise and was produced by setting samples of

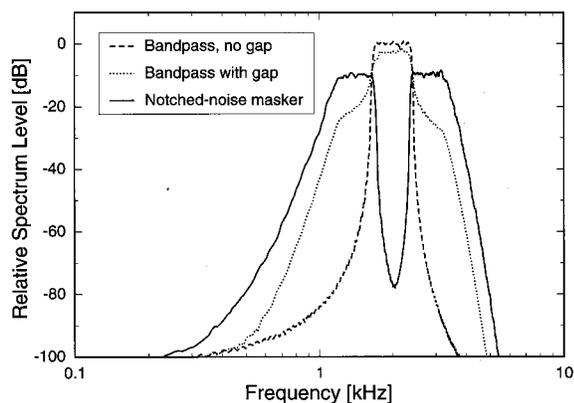


FIG. 1. Averaged spectrum of the stimulus ensemble for the 2-kHz center frequency. The spectrum was measured within a 102.4-ms Hanning window temporally centered on the gap. (The effective duration of the analysis window is 38.4 ms.) The dashed line shows the spectrum of the three-ERB-wide bandpass noise without a gap. The dotted line shows the spectrum of the bandpass noise with a 10-ms gap. The solid line shows the spectrum of the notched-noise masker.

the bandpass noise to zero for the duration of the gap. In the following, these bandpass noises are called gap noises.

To prevent listeners from hearing spectral splatter caused by abrupt gating of the gap, the gap noise was presented with a complementary notched-noise masker, whose spectrum level was 10 dB lower than that of the gap noise. The combined noises were filtered with a nine-ERB-wide filter to limit the bandwidth and loudness of the stimulus ensemble while keeping the spectral splatter inaudible. The filter characteristic ensured that the rise and fall times of the gap were fast enough not to increase the listeners' gap-detection thresholds appreciably. The spectral properties of the stimuli are illustrated in Fig. 1, which shows the spectra of stimuli centered at 2 kHz. As shown by the dotted line, the gap introduces considerable spectral splatter, but its spectrum level is more than 10 dB below that of the notched-noise masker, which renders it inaudible. (The 2.5-dB decrease of the level in the passband of the noise with the gap results because the gap reduces the overall energy within the analysis window and because turning the noise off and on spreads part of the energy to frequencies outside the passband.)

Each listener was tested separately in a sound-attenuating room. A PC-compatible computer with a signal processor (TDT AP2) generated the stimuli, sampled the listeners' responses, and executed the psychophysical procedure. The gap noise and the notched-noise masker were reproduced by a D/A converter (TDT DD1; sample rate=41.67 kHz), whose output was attenuated (TDT PA4) and anti-alias filtered (TDT FT5, $f_c = 20$ kHz, 135 dB/octave). The output of the filter was led through a summation amplifier (TDT SM3) to the headphone amplifier (TDT HB6), which fed one earpiece of the Sony MDR-V6 headphones.

The gap noise was produced in the frequency domain by setting the real and imaginary parts of all spectral lines within the passband to random values chosen from a Gaussian distribution. This spectrum was inverse-Fourier transformed to produce a bandpass noise to which the gap was

added in the time domain. The resulting waveform was Fourier transformed and added to the spectrum of the notched-noise masker, which was produced by setting the real and imaginary parts of all spectral lines outside the passband of the bandpass noise to random values chosen from a Gaussian distribution and scaled by $1/\sqrt{3}$. The spectrum of the combined gap noise and notched-noise masker was multiplied by the nine-ERB-wide filter function. This filter function had unity gain from four ERBs below the CF to four ERBs above it. At frequencies outside this region, the gain decreased according to the RoEx auditory-filter functions centered at the end points of the unity-gain range. Finally, the filtered stimulus was inverse-Fourier transformed and shaped with the 20-ms rise and fall.

B. Listeners and procedure

Three young adult listeners with normal audiometric thresholds and no history of hearing difficulties participated in this study. Their ages ranged from 23 to 25 years.

Psychometric functions for gap detection were measured using a constant-stimulus procedure with a cued Yes–No paradigm. Before the onset of data collection, an adaptive procedure was used to practice the listeners and to aid in determining the range of gap durations to be tested for each listener and center frequency. Five (six at 0.25 kHz) gap durations were chosen to encompass the range of the psychometric function. In addition, a 0.5-ms gap was tested to provide an estimate of the false-alarm rate.

Each trial consisted of two 786-ms intervals separated by 500 ms. The first interval (the “cue”) never contained a gap. The second interval always contained a gap. The listener responded “Gap” if it was heard and “No Gap” if it was not. No feedback was given because the correct answer depended on the listener's perception and could not be defined objectively. Each point on an individual listener's psychometric function was based on 100 trials.

The cue in the first interval was chosen not to have a gap to avoid biasing the listener by encouraging a comparison of the signal gap in the second interval to an arbitrarily chosen cue gap. If the cue had been chosen to be a barely audible gap, it might have encouraged the listener to respond “Gap” for signal gaps far below threshold because both would sound uninterrupted. If the cue had been chosen to be a gap well above threshold, it might have encouraged the listener to respond “No Gap” to signal gaps that were audible but were heard less clearly than the cue. Accordingly, the most appropriate cue appeared to be a stimulus without a gap. Using a cue without a gap makes the Yes–No task similar to a same–different task and allows the listener to base detection on any reliable difference between the first and second interval.

C. Data analysis

To summarize the psychometric functions, the data for each listener and CF were fit by a logistic function

$$P_{\text{yes}} = a + \frac{1-a}{1 + \exp(-k[\log\langle t \rangle - \log\langle m \rangle])}, \quad (1)$$

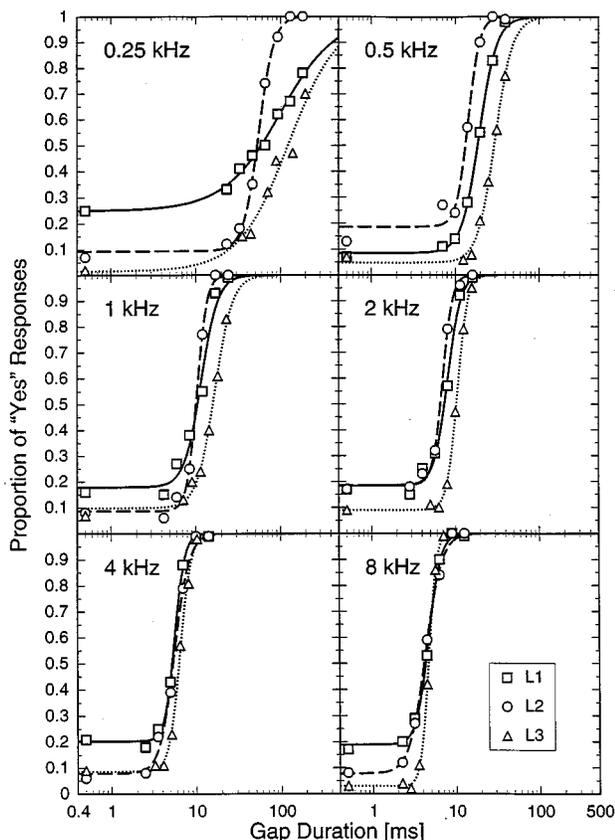


FIG. 2. Individual listeners' psychometric functions for gap detection. The proportion of "Yes" responses is plotted as a function of gap duration. Each panel shows data for a different center frequency. Within each panel, each listener's data are shown by a different symbol as indicated in the legend. The lines show the best-fitting logistic functions.

where a is the false-alarm rate, t is the gap duration, m is the midpoint of the psychometric function, and k is a free parameter that determines the slope. A logarithmic scale of time is used because the form of the psychometric function is likely to be roughly invariant when it is plotted on a logarithmic scale of gap duration. Certainly, data for the Minimum Detectable Gap duration (MDG) generally have a standard deviation that is approximately proportional to the MDG (e.g., Florentine and Buus, 1983). This finding indicates that the psychometric functions for various stimulus conditions should be roughly parallel when plotted on a logarithmic scale of gap duration. The utility of a logarithmic scale of gap duration also follows from Green and Forrest's (1989) demonstration that several different psychometric functions could reasonably be summarized by a single function when the gap durations were normalized by dividing them by the MDG. Such normalization corresponds to shifting psychometric functions horizontally along a logarithmic scale of gap duration.

II. RESULTS

Figure 2 shows psychometric functions for each of the three listeners at all center frequencies. As shown by the various lines, the logistic functions describe the data well. The average rms error is only around 0.005. Except at 0.25 kHz, the psychometric functions appear relatively similar

TABLE I. Parameters for the logistic functions used to describe the psychometric functions. The first column shows the center frequency, CF. The next three columns show the false-alarm rate, a , the slope parameter, k , and the gap duration at the midpoint of the psychometric function, m . These values are averages of parameters used to fit the psychometric functions for individual listeners. The standard deviations are shown in parentheses. The range of plus and minus one standard deviation for m is the mean multiplied and divided by the error factor shown in parenthesis. The fifth and sixth columns show the slope parameter, k' , and the midpoint, m' , for logistic functions fitted to the predictions by the loudness-detector model (see Fig. 9) with the false-alarm rate fixed at 0.1.

CF	a	k	m [ms]	k'	m' [ms]
0.25 kHz	0.12 (0.12)	5.8(4.6)	88.1(1.54)	6.6	27.7
0.5 kHz	0.11 (0.07)	11.1(2.7)	20.3(1.45)	7.4	18.6
1 kHz	0.12 (0.05)	12.4(4.2)	12.7(1.28)	7.9	12.5
2 kHz	0.15 (0.06)	15.2(2.4)	8.4(1.23)	8.2	8.6
4 kHz	0.12 (0.07)	14.4(2.7)	5.9(1.09)	8.1	6.2
8 kHz	0.10 (0.08)	15.1(5.8)	4.6(1.05)	7.7	4.6

across listeners, although the false-alarm rates vary. The MDGs—defined as the midpoint of the fitted psychometric function—decrease systematically as CF increases and are in excellent quantitative agreement with those obtained in our earlier studies (i.e., Florentine and Buus, 1983; Buus and Florentine, 1985). The present MDGs also agree qualitatively with those obtained by Fitzgibbons (1983, 1984) and by Shailer and Moore (1983), but the present MDGs are somewhat larger due to differences in stimuli and threshold criteria.

To summarize the data in Fig. 2, the best-fitting values of the false-alarm rate, a , the slope parameter, k , and the gap duration at the midpoint of the psychometric function, m , for each listener and center frequency were averaged across the three listeners. Table I shows the average parameters together with slope parameters and midpoints for psychometric functions predicted by a model that is discussed later.

The general trends in the data are illustrated by Fig. 3, which shows psychometric functions obtained with the average parameters shown in Table I. Because the MDGs decrease with increasing frequency, the psychometric functions are ordered from left to right according to decreasing fre-

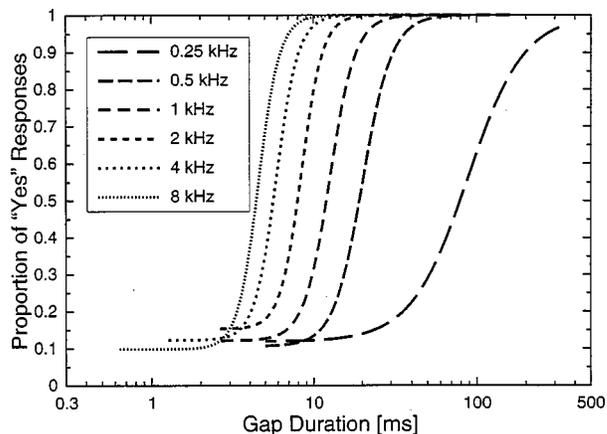


FIG. 3. Average psychometric functions in a cued Yes-No paradigm. Each curve shows the logistic function whose parameters are the averages of parameters obtained for the three listeners.

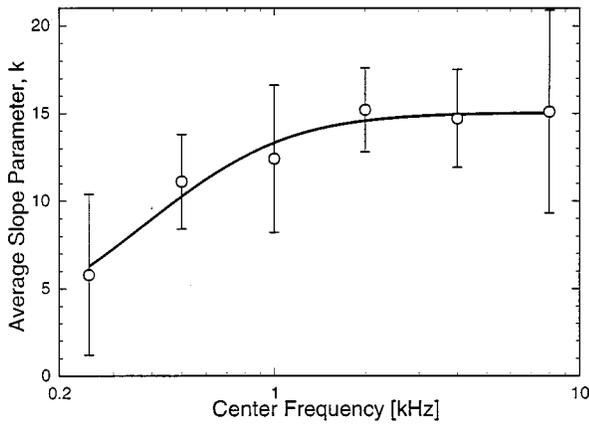


FIG. 4. Average slopes of the psychometric functions plotted as a function of center frequency. The error bars show plus and minus one standard deviation calculated across the slopes for the three individual listeners. The solid curve shows the function used to summarize the data.

quency. The psychometric functions in Fig. 3 also show that the slope is shallower at low than at high frequencies.

To illustrate the effect of frequency on the slope of the psychometric functions, the average slopes are plotted as a function of frequency in Fig. 4. It shows that the slope increases as the frequency increases from 0.25 to 2 kHz, but is relatively constant between 2 and 8 kHz. On the average, the slope as a function of frequency, $k(f)$, can be approximated by the following formula:

$$k(f) = G \left(\frac{(f/F_c)^n}{1 + (f/F_c)^n} \right)^{1/n}, \quad (2)$$

where G is the asymptotic slope at high frequencies, F_c is the nominal cutoff frequency of this high-pass function, and n determines the slope of $k(f)$ below F_c . The constants G , F_c , and n are free parameters. The best fit to the present data is obtained when $F_c = 550$ Hz, $G = 15.05$, and $n = 2.07$, which yields an rms error of 0.61. As shown in Fig. 4, these parameters produce a good approximation to the data.

III. DISCUSSION

A. Comparison with other data in the literature

Although numerous studies have measured gap detection, there are very few data on the form of the psychometric functions for gap detection in general and apparently no data on the form of the psychometric function for gap detection in bandpass noise. Shailer and Moore (1987) and Moore *et al.* (1989) showed that psychometric functions for detection of gaps in tones at low frequencies are nonmonotonic due to phase effects. Therefore, they are difficult to compare with the present data. At high frequencies the nonmonotonicity is not apparent, but the lack of randomness in the tonal stimuli is likely to make the task somewhat different from the detection of a gap in a narrow-band noise.

Data for broadband noise ought to be more comparable to the present data, because earlier work indicates that gap detection in broadband noise depends primarily on temporal resolution in the high-frequency auditory channels (Florentine and Buus, 1983; Buus and Florentine, 1985). In fact, He

et al.'s (1999) psychometric functions for broadband noise show that the proportion of "Yes" responses typically increases from 20% to 80% in about 2 ms, which is comparable to the present data of 3.1 ms at 4 kHz and 2.2 ms at 8 kHz. This agreement is also evident if He *et al.*'s (1999) data for individual young listeners are fitted on our coordinates. The k 's that yield the best fit of Eq. (1) to their data for detection of gaps in 400-ms noise bursts averaged 18.7 (± 2.8 std. dev.). This value agrees reasonably with the asymptotic high-frequency value of 15 for the function that summarizes the slopes in Fig. 4.

Green and Forrest (1989) also obtained psychometric functions for gap detection in broadband noise, but they used a two-alternative forced-choice paradigm. Their data show that the percentage of correct responses increased from 60% to 90% over about 1.1 ms, or about 46% of their 2.4-ms MDG. The present data show that the proportion of "Yes" responses increases from 20% to 80% (corresponding to 60% and 90% correct in a 2AFC task) in 3.1 ms at 4 kHz and 2.2 ms at 8 kHz. These values correspond to about 50% of the MDGs. In addition, converting Green and Forrest's (1989) data to percent "Yes" [$P_{\text{Yes}} = 2P_C - 100$] and fitting them on our coordinates yields a slope parameter of 17, which agrees reasonably with our value of 15 for the asymptotic high-frequency slope parameter (see Fig. 4 and Table I).

B. Modeling gap detection

1. Previous model: Energy-detector model

Models of gap detection often assume that a gap can be heard if the short-term energy at the output of a short-term integrator at the end of the gap is sufficiently below the output during the uninterrupted signal (e.g., Plomp, 1964; Buunen and van Valkenburg, 1979; Buus and Florentine, 1985). Applying standard Signal Detection Theory to this energy-detector model, Buus and Florentine (1985) showed that the sensitivity, d' , in a single auditory channel can be calculated as

$$d' = \sqrt{\frac{W\tau}{1 + W\tau\sigma_I^2}} \left(1 - \frac{2W\tau}{2W\tau - 1} \exp\left(\frac{-t}{\tau}\right) + \frac{1}{2W\tau - 1} \exp(-2Wt) \right), \quad (3)$$

where W is the bandwidth of the auditory filter, τ is the time constant of an exponential-window short-term integrator, σ_I is a proportionality constant that characterizes the amount of internal noise added by the auditory system, and t is the duration of the gap. Buus and Florentine (1985) set W equal to the critical bandwidth and τ equal to 30 ms. This energy-detector model was quite successful in predicting how the MDG for octave-band noises varied with center frequency, except that it predicted better-than-obtained performance for center frequencies below 1 kHz.

2. Effects of auditory filtering

According to the energy-detector model, auditory-filtering affects temporal resolution as measured by the

MDGs in two ways. First, auditory filtering affects the variance within an auditory channel when the bandwidth of the gap noise exceeds the bandwidth of the auditory filters. Because the variance increases as the auditory-filter bandwidth decreases at low frequencies, temporal resolution is poorer at low than at high frequencies. Second, the model assumes that ringing of the auditory filters diminishes temporal resolution at low frequencies, because the ringing increases as the auditory-filter bandwidth decreases toward low frequencies.

Eddins *et al.* (1992) argued that the latter effect is not important. They found that MDGs were independent of center frequency when the absolute bandwidth of the gap noise was fixed, but decreased as the absolute bandwidth increased. These results were consistent with a simple energy-detector model that did not include auditory filtering. Therefore, they concluded that the effect of center frequency on MDGs obtained with gap noises of constant relative bandwidth (e.g., octave-band noises) was not due to ringing of the auditory filters, but simply resulted because the bandwidth increased with center frequency.

As discussed below, Eddins *et al.*'s (1992) conclusion appears incorrect for several reasons. First, when the bandwidth of the gap noise is narrow, the variability of the noise is very large and the variance of the decision variable in the energy-detector model is determined almost exclusively by the stimulus. Accordingly, the variance is independent of center frequency, but decreases as the bandwidth of the noise increases up to the auditory-filter bandwidth. The latter effect causes the energy-detector model to predict that the MDGs decrease with increasing bandwidth of noises with bandwidths less than the auditory-filter bandwidth. Second, Eddins *et al.* (1992) filtered their stimuli after introducing the gap. Therefore, the rise and fall time of the gap increased with decreasing bandwidth. Although the rise-fall time always was shorter than the MDG, it was longer than the auditory-filter ringing when the gap noises were narrow. Accordingly, one would expect little or no effect of auditory-filter ringing on MDGs for narrow-band gap noises. Taken together, these two effects allow the energy-detector model to account for the independence of center frequency and effect of bandwidth that Eddins *et al.* (1992) observed for narrow-band gap noises.

When the gap noises were wide, on the other hand, the rise-fall time of the filtered stimulus was short and one might expect auditory-filter ringing to affect the MDGs. Likewise, the variance of the decision variable produced in an auditory channel by a wideband gap noise ought to depend on the auditory-filter bandwidth and not on the stimulus bandwidth. Therefore, one might expect that gap noises with fixed absolute bandwidths that are relatively wide should yield larger MDGs at low than at high frequencies. However, this expectation ignores the number of auditory channels activated by a noise with fixed absolute bandwidth, which is larger at low than at high frequencies. Florentine and Buus (1996) found that the increased number of channels largely counteracts the reduced temporal resolution caused by the increased variance and prolonged ringing of the auditory filters at low frequencies. Accordingly, the energy-detector model also predicts the effect of bandwidth and indepen-

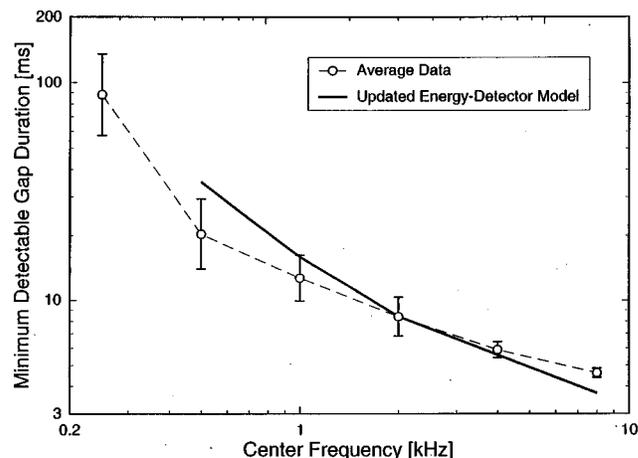


FIG. 5. The average MDGs calculated from the midpoints of the psychometric functions for the three listeners in the present experiment. The error bars show the range of plus and minus one standard deviation. The solid line shows predictions by the updated energy-detector model.

dence of center frequency for the wider gap noises used by Eddins *et al.* (1992). Finally, the independence of MDGs on center frequency holds only over a limited frequency range. Measurements of MDGs over a wider range of frequencies show complex interactions between bandwidth and upper cutoff frequency (Snell *et al.*, 1994). The energy-detector model agrees with these findings (Florentine and Buus, 1996).

These considerations lead to the conclusion that Eddins *et al.*'s (1992) data should not be taken to indicate that temporal resolution in the auditory system is frequency independent. According to the energy-detector model, the effects of stimulus variance after auditory filtering, filter ringing, and multiple observations in independent auditory channels interact in a complex manner that is consistent with a wide variety of data on gap-detection thresholds.

3. Updated energy-detector model

Although the energy-detector model developed by Buus and Florentine (1985) predicts many aspects of gap detection, some of its parameters may need to be modified in light of recent data. Modifications include (1) using the Equivalent Rectangular Bandwidth (ERB, see Glasberg and Moore, 1990; Buus, 1997) instead of the critical bandwidth and (2) shortening the integration time from 30 to 10 ms. These modifications improve the low-frequency predictions and bring the energy-detector model into agreement with modern estimates of auditory frequency selectivity (e.g., Glasberg and Moore, 1990) and short-term temporal integration (e.g., Moore *et al.*, 1988).

As shown in Fig. 5, the modified parameters of the updated energy-detector model provide a reasonable fit to the average MDGs for center frequencies of 0.5 kHz and above. However, if the usual Gaussian approximations are used, the updated energy-detector model predicts (erroneously) that gap-detection thresholds should not be measurable at 0.25 kHz with the threshold criterion (d' equal to about 1.3) imposed by using m as an estimate of the MDG (Geng, 1997). This prediction arises in large part because it was derived

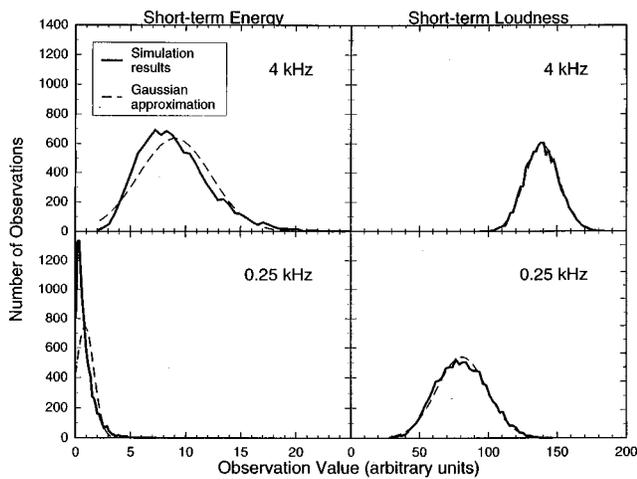


FIG. 6. Frequency distributions of short-term energy (left) and short-term loudness (right) at 0.25 (bottom) and 4 kHz (top).

under the assumption that the output of the integrator can be approximated by Gaussian distributions, which is valid only when the bandwidth-integration-time product ($W\tau$) is larger than about 10. Given the short integration time and the relatively narrow bandwidth of the auditory filter (about 45 Hz at 0.25 kHz), the Gaussian assumption is clearly violated. Even at 8 kHz, $W\tau$ is only about 9 when τ is 10 ms. Thus, it appears useful to investigate if decision variables other than energy will allow reasonably accurate approximations with Gaussian distributions.

4. Loudness-detector model

The true probability distribution of short-term energy is the Chi-square distribution, which is strongly skewed and has a long tail toward high values when $W\tau$ is small. To obtain distributions that are more symmetric and closer to Gaussian distributions, a compressive transformation is needed. This indicates that it may be advantageous to use a decision variable derived from integration of some neural activity that can be thought to represent the output of inner hair cells after the acoustic signal has been compressed by the cochlear mechanics. The present work investigated a model in which the output of an auditory filter was transformed by a power function (of intensity) with an exponent of 0.23, which corresponds to the transformation to specific loudness in Zwicker's (1958) excitation-pattern model of loudness (see also Moore *et al.*, 1997). This loudness-detector model is similar to the model used by Oxenham, Moore, and colleagues to describe forward masking, detection of a tone pip in a gap, and detection of increments and decrements (e.g., Oxenham and Moore, 1994; Peters *et al.*, 1995; Oxenham, 1997; for review see Moore and Oxenham, 1998).

To investigate the statistical properties of Gaussian noise processed by the loudness-detector model, simulations were performed. The results are shown in Fig. 6. Each panel shows the frequency distribution of 10 000 observations of the output of a simulated auditory channel together with Gaussian distributions that have the same mean and standard deviation as the observed variable. To illustrate the effect of

the auditory-filter bandwidth, data are shown for auditory channels tuned to 0.25 and 4 kHz. As shown on the left, the short-term energy is clearly non-Gaussian. As shown on the right, integration of the compressed output of the filter yields a short-term loudness, whose frequency distribution is well approximated by a Gaussian distribution, even at the lowest frequencies where the distributions of short-term energy are clearly skewed.

Attempts to approximate the distributions of short-term loudness by applying simple transformations to the expected variance ($1/W\tau$) of the short-term energy revealed that the variance was smaller than the expected value calculated from the ERBs of the auditory filters. The discrepancy results from the finite slope of the Rounded Exponential (RoEx) form of the auditory filters. The expressions for the mean and variance of short-term energy commonly used in Signal Detection Theory are valid only for ideal, rectangular filters. The ERB of a real filter with finite slopes is the bandwidth of an ideal rectangular filter whose output has the same average power as the output of the real filter. That is, the ERB is the proper bandwidth for estimating the mean of the short-term energy distribution, but it is not appropriate for estimating the variance. To obtain proper estimates of the variance of the short-term energy, one must use the bandwidth of an ideal filter whose output has the same variance as the output of the real filter. This bandwidth is called the Equivalent Statistical Bandwidth (ESB), which is wider than the ERB for typical real filters. If the input to a filter is white noise, the ESB can be calculated as

$$\text{ESB} = \frac{(\int_0^\infty H^2(f) df)^2}{\int_0^\infty H^4(f) df}, \quad (4)$$

where $H(f)$ is the amplitude gain as a function of frequency. Evaluation of this expression using the RoEx transfer function shows that the ESB is 1.6 times wider than the ERB. The normalized variance of a broadband noise filtered by an auditory filter is $1/(1.6W\tau)$, where W is the ERB, which is the normally quoted auditory-filter bandwidth.

To arrive at useful approximations for variance of the output of the loudness-detector model, it was necessary to account for the compression and—in the present experiment—the slight increase in variance caused by using a bandpass noise rather than a white noise as the input to the auditory filter. To account for the compression, the standard deviation of the normalized energy may be reduced by the small-signal “gain” (i.e., the derivative) of the compressive power function for inputs near unity (gain=0.23). To account for the effects of using bandpass noise as a stimulus, the statistics of the short-term loudness in simulated auditory channels tuned to frequencies between 0.25 and 14 kHz was examined. The normalized standard deviations of the short-term loudness at various center frequencies are shown in Fig. 7. These results showed that the normalized standard deviation could be well approximated by assuming an ESB (corresponding to the combined characteristic of the auditory filter and the stimulus filtering) of about 1.4 ERBs and reducing the normalized standard deviation of the short-term energy $[= \sqrt{1/(1.4W\tau)}]$ by the compressive small-signal gain of 0.23. The solid line in Fig. 7 shows this approxima-

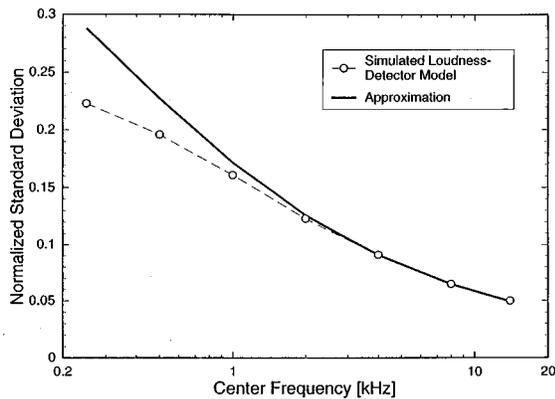


FIG. 7. Normalized standard deviation of simulated output of the short-term integrator after auditory filtering and compression of the three-ERB-wide noise band used in the gap-detection experiments. The normalized standard deviation is plotted as a function of center frequency. The solid line shows an approximation calculated as $0.23/\sqrt{1.4W\tau}$.

tion. It is excellent for frequencies at and above 1 kHz, but overestimates somewhat the stimulus variability at low frequencies.

In their derivation of d' for an energy-detector model [Eq. (3)], Buus and Florentine (1985) assumed that the ringing of the auditory filter could be approximated by the output amplitude following an exponential decay with a time constant equal to the reciprocal of the critical bandwidth. This assumption together with an exponential short-term integrator is responsible for the second term of Eq. (3). Following similar assumptions, it can be shown that the short-term loudness (i.e., the output of the integrator acting on the compressed output of the filter) will decay as

$$N'(t) = \left(\frac{\alpha W \tau}{\alpha W \tau - 1} \exp\left(\frac{-t}{\tau}\right) - \frac{1}{\alpha W \tau - 1} \exp(-\alpha W t) \right), \quad (5)$$

where α is the exponent (relative to amplitude; =0.46) of the power-function nonlinearity following the filter, and W is the bandwidth of the filter (or more precisely, the reciprocal of the time constant for the exponential amplitude decay of the filter output), τ is the time constant for the short-term integrator (=10 ms), and t is the time since the beginning of the gap. In the following, W is assumed equal to the ERB.

It follows that it should be possible to estimate d' for a single auditory channel in the loudness-detector model as

$$d' = \sqrt{\frac{1.4W\tau}{(0.5\alpha)^2 + 1.4W\tau\sigma_1^2}} \left(1 - \frac{\alpha W \tau}{\alpha W \tau - 1} \exp\left(\frac{-t}{\tau}\right) + \frac{1}{\alpha W \tau - 1} \exp(-\alpha W t) \right). \quad (6)$$

To apply this expression to the present data, it can be assumed that our three-ERB-wide noises encompass three independent auditory channels such that the listeners' d' can be estimated as $\sqrt{3}$ times the d' calculated for a single channel.

As shown in Fig. 8, the MDGs predicted by the loudness-detector model fit the data better than those predicted by the energy-detector model (compare to Fig. 5). The data and the loudness-detector model agree between 0.5 and

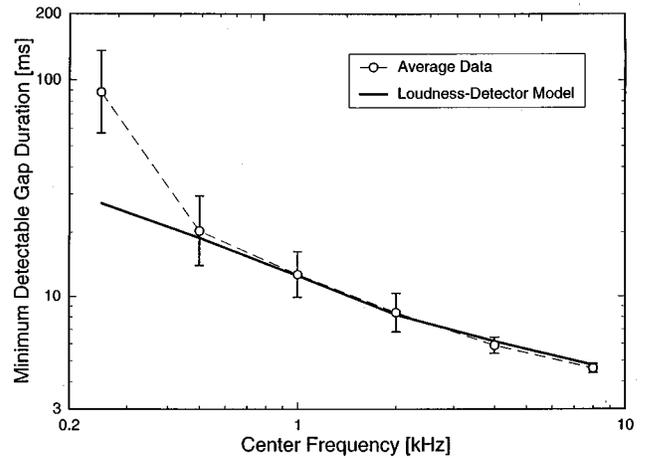


FIG. 8. Predictions for gap detection by the loudness-detector model. The solid line shows predictions obtained from the loudness-detector model when threshold corresponds to the d' calculated from the midpoint and false-alarm rate obtained at each center frequency and the internal-noise constant, σ_I , is 0.27. The circles show the average MDGs for three listeners and are replotted from Fig. 5.

8 kHz. However, the predicted MDG at 0.25 kHz is only about one-third that obtained. The loudness-detector model's prediction of better-than-obtained performance at 0.25 kHz may reflect that the exponential-decay approximation to the filter ringing is inaccurate over the relatively long durations required to describe gap-detection performance at 0.25 kHz. It may also reflect possible inefficiency of the listeners caused by the pronounced, relatively slow fluctuations of noise filtered by the low-frequency auditory filters and/or that the time constant for the short-term integrator may increase at low frequencies (Peters *et al.*, 1995).

Psychometric functions predicted by the loudness-detector model also resemble the data, but clear discrepancies are apparent. As shown in Fig. 9, the probability for a "Yes" response increases from the false-alarm rate (chosen to be 0.1, which is a typical value) to near 100% as the gap duration increases about ten-fold. This agrees quite well with data for low frequencies presented in Fig. 3. However, as shown in the fifth column of Table I, the loudness-detector model shows only a very small effect of frequency on the predicted slope of the psychometric functions. This is contrary to the data in the third column of Table I, which show

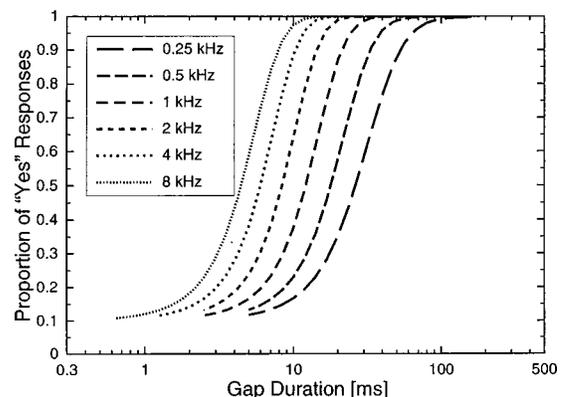


FIG. 9. Psychometric functions predicted by the loudness-detector model.

a clear increase in the slope of the psychometric functions as the frequency increases from 0.25 to about 2 kHz (see also Fig. 4). The measured psychometric functions at and above 2 kHz are about twice as steep as those predicted by the loudness-detector model.

The discrepancy between the predicted and measured psychometric functions may indicate that gap detection is not based on a simple stationary transformation of the output of the auditory filter. Certainly, it is reasonable to expect that gap detection may be aided by the onset response elicited in the auditory nerve by the abrupt rise of the trailing noise at the end of a sufficiently long gap. Zhang *et al.* (1990) investigated the auditory-nerve response to gaps in broadband noise and found a vigorous response to the onset of the noise at the end of the gap when the gap duration was somewhat longer than the MDG. Because the onset response at the end of the gap is absent for brief gaps but strong when the gap duration exceeds the MDG, its effect on gap detection ought to vary with the gap duration. The detection contribution by the onset response must be small or absent for brief gaps, but may be considerable for long gaps. Therefore, the onset response is likely to make the psychometric functions steeper than they would be if the onset response were absent.

The loudness-detector model makes no attempt to include the onset response. It is quite possible that the onset response at the end of the gap is more vigorous at high than at low frequencies. The auditory filters make the effective rise time of the filtered stimulus faster as their absolute bandwidths increase with increasing frequency and the onset response increases as the rise time decreases (Smith and Brachman, 1982). Given that the loudness-detector model does not include this effect, it is not surprising that it fails to predict the steep psychometric functions obtained at high frequencies. A more realistic model should take into account how the onset response contributes to detection of gaps with various durations. To our knowledge, comprehensive parametric data that would allow quantitative modeling of this important factor do not exist. Therefore, we believe that attempting to include these effects in our model is not warranted at present.

IV. SUMMARY

The present study measured three normal listeners' psychometric functions for detection of temporal gaps in bandpass noises using a cued Yes–No procedure. The data were compared to predictions by a loudness-detector model, which based its decision on the output of a short-term integrator acting on the compressed output of the auditory filters. The results show that:

- (1) The slope of the psychometric function for gap detection increases with increasing center frequency up to about 2 kHz and remains approximately constant at higher frequencies.
- (2) Gap-detection thresholds—as determined by the midpoint of the psychometric functions—decrease systematically with increasing center frequency.
- (3) An energy-detector model employing the usual Gaussian approximations fails to predict the data. Part of the fail-

ure results because the bandwidth-duration product at low frequencies is too low for the approximation to be valid when modern estimates of the auditory-filter bandwidth and the short-term integration time are used.

- (4) A loudness-detector model, which employs a compressive nonlinearity at the output of the auditory filters, produces decision variables that closely follow Gaussian distributions.
- (5) The loudness-detector model produces reasonable predictions of gap-detection thresholds, except at 0.25 kHz where the predicted threshold is considerably lower than that obtained.
- (6) The loudness-detector model also predicts the form of the psychometric functions at low frequencies, but the predicted increase in slope with increasing frequency is much smaller than that obtained. This failure may indicate that a realistic model of gap detection must take into account how the onset response elicited in the auditory nerve at the end of the gap grows with increasing duration of the gap.

ACKNOWLEDGMENTS

The data are from the third author's Master's thesis. Work supported by NIH/NIDCD Grant No. R01-DC00187 and GN Danavox. The authors would like to thank Craig Formby and an anonymous reviewer for their helpful comments. Parts of this paper were presented at the joint meeting of the 16th International Congress on Acoustics and the 135th meeting of the Acoustical Society of America in Seattle, WA [M. Florentine *et al.*, *J. Acoust. Soc. Am.* **103**, 2847 (1998)].

- Buunen, T. J., and van Valkenburg, D. A. (1979). "Auditory detection of a single gap in noise," *J. Acoust. Soc. Am.* **65**, 534–536.
- Buus, S. (1997). "Auditory masking," in *Encyclopedia of Acoustics*, edited by M. J. Crocker (Wiley, New York), pp. 1427–1445.
- Buus, S., and Florentine, M. (1985). "Gap detection in normal and impaired listeners: The effect of level and frequency," in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, London), pp. 159–179.
- Eddins, D., Hall, J. W., and Grose, J. H. (1992). "The detection of temporal gaps as a function of frequency region and absolute noise bandwidth," *J. Acoust. Soc. Am.* **91**, 1069–1077.
- Fitzgibbons, P. J. (1983). "Temporal gap detection in noise as a function of frequency, bandwidth, and level," *J. Acoust. Soc. Am.* **74**, 67–72.
- Fitzgibbons, P. J. (1984). "Tracking a temporal gap in bandpass noise: Frequency and level effects," *Percept. Psychophys.* **35**, 446–450.
- Fitzgibbons, P. J., and Wightman, F. L. (1982). "Gap detection in normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **72**, 761–765.
- Florentine, M., and Buus, S. (1983). "Temporal resolution as a function of level and frequency," *Proc. 11th Int. Cong. Acoust.* **3**, 103–106.
- Florentine, M., and Buus, S. (1996). "Some current trends in temporal processing," in *Recent Trends in Hearing Research*, edited by H. Fastl, S. Kuwano, and A. Schick (Oldenburg University, Germany), pp. 161–191.
- Geng, W. (1997). "A method for clinical assessment of temporal resolution," Master's thesis, Northeastern University, Boston, MA.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- Green, D. M., and Forrest, T. G. (1989). "Temporal gaps in noise and sinusoids," *J. Acoust. Soc. Am.* **86**, 961–970.
- He, N., Horowitz, A. R., Dubno, J. R., and Mills, J. H. (1999). "Psychometric functions for gap detection in noise measured from young and aged subjects," *J. Acoust. Soc. Am.* **106**, 966–978.

- Irwin, R. J., and Kemp, S. (1976). "Temporal summation and decay in hearing," *J. Acoust. Soc. Am.* **59**, 920–925.
- Moore, B. C. J., Glasberg, B. R., Donaldson, E., McPherson, T., and Plack, C. J. (1989). "Detection of temporal gaps in sinusoids by normally hearing and hearing-impaired subjects," *J. Acoust. Soc. Am.* **85**, 1266–1275.
- Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (1988). "The shape of the ear's temporal window," *J. Acoust. Soc. Am.* **83**, 1102–1116.
- Moore, B. C. J., Glasberg, B. R., and Baer, T. (1997). "A model for the prediction of thresholds, loudness, and partial loudness," *J. Aud. Eng. Soc.* **45**, 224–240.
- Moore, B. C. J., and Oxenham, A. J. (1998). "Psychoacoustic consequences of compression in the peripheral auditory system," *Psychol. Rev.* **105**, 108–124.
- Oxenham, A. J. (1997). "Increment and decrement detection in sinusoids as a measure of temporal resolution," *J. Acoust. Soc. Am.* **102**, 1779–1790.
- Oxenham, A. J., and Moore, B. C. J. (1994). "Modeling the additivity of nonsimultaneous masking," *Hear. Res.* **80**, 105–118.
- Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (1995). "Effects of level and frequency on the detection of decrements and increments in sinusoids," *J. Acoust. Soc. Am.* **97**, 3791–3799.
- Plomp, R. (1964). "Rate of decay of auditory sensation," *J. Acoust. Soc. Am.* **36**, 277–282.
- Shailer, M. J., and Moore, B. C. J. (1983). "Gap detection as a function of frequency, bandwidth, and level," *J. Acoust. Soc. Am.* **74**, 467–473.
- Shailer, M. J., and Moore, B. C. J. (1987). "Gap detection and the auditory filter: Phase effects using sinusoidal stimuli," *J. Acoust. Soc. Am.* **81**, 1110–1117.
- Smith, R. L., and Brachman, M. L. (1982). "Adaptation in auditory-nerve fibers: A revised model," *Biol. Cybern.* **44**, 107–120.
- Snell, K. B., Ison, J. R., and Frisina, D. R. (1994). "The effects of signal frequency and absolute bandwidth on gap detection in noise," *J. Acoust. Soc. Am.* **96**, 1458–1464.
- Tyler, R. S., Summerfield, Q., Wood, E. J., and Fernandes, M. A. (1982). "Psychoacoustic and phonetic temporal processing in normal and hearing-impaired," *J. Acoust. Soc. Am.* **72**, 740–752.
- Zhang, W., Salvi, R. J., Saunders, S. S. (1990). "Neural correlates of gap detection in auditory nerve fibers of the chinchilla," *Hear. Res.* **46**, 181–200.
- Zwicker, E. (1958). "Über die psychologische und methodische Grundlagen der Lautheit," *Acustica* **8**, 237–258.
- Zwicker, E., and Feldtkeller, R. (1967). *Das Ohr als Nachrichtenempfänger* (Hirzel, Stuttgart). Available in English (translation by H. Müsch, S. Buus, and M. Florentine): *The Ear as a Communication Receiver* (Acoustical Society of America, Woodbury, NY, 1999).

Informational masking by everyday sounds

Eunmi L. Oh

Waisman Center, University of Wisconsin, Madison, Wisconsin 53705

Robert A. Lutfi

Department of Communicative Disorders, and Waisman Center, University of Wisconsin, Madison, Wisconsin 53705

(Received 20 January 1999; revised 26 August 1999; accepted 27 August 1999)

Tone-in-noise detection is severely degraded when only a few spectral components of the noise are presented at random on each trial [Neff and Green, *Percept. Psychophys.* **41**, 409–415 (1987)]. The elevations in threshold are attributed to uncertainty regarding the noise caused by sparse sampling of noise components—informational masking is the term used to describe the result. The present experiment was undertaken to determine how informational masking is affected when sparse sampling is from a set of common everyday sounds instead of noise. On each presentation a different masker waveform of constant total power was synthesized from the magnitude and phase of a fixed number, m , of spectral components ($m = 2–921$ across blocks of trials). The components were selected at random from 1 of 50 common environmental sounds (e.g., baby crying, door slamming, phone ringing), or 1 of 50 samples of Gaussian noise. Masked thresholds for a 1.0-kHz signal in the presence of both types of maskers were obtained using an adaptive, two-interval, forced-choice procedure. Results with noise replicated earlier, results showing largest elevations in threshold for 10–20 sampled components. Results with everyday sounds showed a similar pattern with thresholds elevated above those for noise by as much as 10 dB at the larger values of m . The differences in masked thresholds were systematically related to differences in the ensemble variance of masker spectra, as predicted by a model previously applied to noise [Lutfi, *J. Acoust. Soc. Am.* **94**, 748–758 (1993)]. Not predicted was the result of a subsequent trial-by-trial analysis in which 9–11 dB less masking was observed for samples from everyday sounds rated as easily recognized by listeners. The results suggest that listeners fail to exploit lawful dependencies among spectral components of everyday sounds to aid detection, unless enough information is available for the sound to be easily recognized. © 1999 Acoustical Society of America. [S0001-4966(99)04412-4]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Fe [SPB]

INTRODUCTION

Tone detection in noise can be severely degraded when there is uncertainty regarding the noise waveform to be presented from trial to trial. The resultant elevation in threshold is referred to as *informational* masking (Pollack, 1975). In a typical experiment, high levels of noise uncertainty are created by presenting only partial information regarding the noise in the form of a small set of spectral components sampled at random on each trial. Frequencies are often constrained to fall outside a band immediately surrounding the signal so as to minimize energy-based masking of the signal (cf. Patterson, 1976). The number of sampled components is critical. For 10–20 components, informational masking is at a maximum, amounting to as much as 40 dB whether components are presented simultaneously (Neff and Green, 1987) or sequentially (Watson and Kelly, 1981). For larger numbers of components, the amount of informational masking is significantly less. In this case, enough information is presumably available regarding the noise that there is little uncertainty associated with the small differences that remain among the different noise samples played from trial to trial (Green, 1988; Oh and Lutfi, 1998).

The results of past studies using noise have often led to speculation on the role of uncertainty in everyday listening. Everyday sounds, after all, can also vary quite unpredictably

from one moment to the next, and often yield only partial information because of masking by other sounds. The popular view, however, is that results with noise are not likely to generalize to everyday sounds (e.g., Watson, 1976). Everyday sounds, unlike arbitrary samples of noise, are highly familiar; they also have a high degree of regularity or structure imposed on them by common physical laws governing the resonance of the source. Both factors are expected to reduce uncertainty associated with these sounds in the absence of complete information. Consider, for example, the spectrograms of the three everyday sounds shown in Fig. 1. The examples are representative of three types of structure that are common to everyday sounds: harmonicity as in the case of the car horn, amplitude modulation as in the case of the door creak, and frequency modulation as in the case of the baby cry.¹ In all cases, the spectrograms are characterized by a high level of redundancy or repetition which distinguishes them from Gaussian noise (first panel). This suggests that very little information (perhaps a few spectral components) would be needed to infer the overall pattern and avoid confusing the signal with a spectral component of the masker. Reinforcing this view is that fact that each of these types of structure has already been shown to reduce masking of a tonal signal in noise (Moore *et al.*, 1986; Hall *et al.*, 1984; Grose and Hall, 1990).

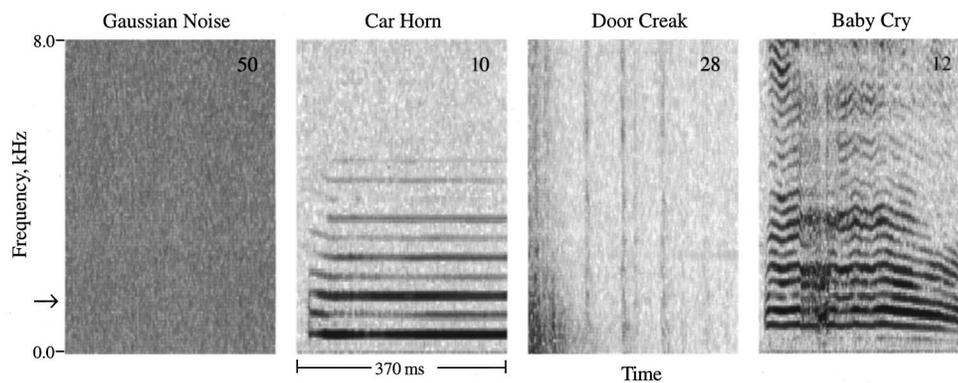


FIG. 1. Reading from left to right. Spectrogram of Gaussian noise, and examples of three forms of temporal-spectral structure common to the set of everyday sounds: Harmonicity (fire horn), amplitude modulation across frequency (door creak), and frequency modulation (baby cry). The number in the upper left-hand corner of each panel indicates the number exemplars falling in each category for the set of sounds used in this study, as determined by visual inspection of the spectrograms. The arrow on the left indicates signal frequency.

While such arguments are compelling, there are also reasons to expect that everyday sounds would be more effective than noise as informational maskers. First, it is important to note that in past studies imposing structure on noise, the structure itself was never uncertain, that is, it did not vary across the different samples of noise. This is not true for different samples of everyday sounds, as Fig. 1 makes clear. Uncertainty regarding structure, like uncertainty regarding any other aspect of the sound, might be expected to increase the potential for informational masking. Second, for everyday sounds increasing the number of sampled components does not reduce gross differences among potential maskers in the same way it does for noise; consequently, uncertainty remains high even for large numbers of sampled components. If listeners are uncertain regarding structure, and if increasing the number of sampled components does not appreciably reduce uncertainty, then one may reasonably expect that everyday sounds would produce more, not less, informational masking, particularly when the number of sampled components is large. This, in fact, is the prediction of a model that has been successfully applied to the results with noise (Lutfi, 1993). In the component relative entropy (CoRE) model differences among maskers are quantified in terms of the ensemble variance of masker spectra across trials. Increasing the number of sampled components reduces the ensemble variance for Gaussian noise, but has much less effect for everyday sounds. Hence, the model predicts that components sampled from everyday sounds should eventually produce more masking than components sampled from noise as the number of components is increased. The goal of the present study was to test these two very different predictions regarding the effectiveness of everyday sounds as informational maskers.

I. METHODS

A. Listeners

Five normal-hearing listeners, three males and two females, were paid at an hourly rate for their participation. The ages of the listeners ranged from 20 to 37 years. The listeners' pure-tone air conduction thresholds were less than 15 dB HL (ANSI, 1989) between 150 and 8000 Hz in both ears. All listeners participated in a previous study in our lab (Oh and Lutfi, 1998) involving a very similar detection task.

B. Stimuli

The signal was a 1000-Hz tone. Any single masker was a collection of m spectral components selected at random from 1 of 50 samples of Gaussian noise or 1 of 50 samples of common everyday sounds. The everyday sounds were taken from the "Living Sound Effects" CD set produced by Bainbridge Entertainment Co., Inc. (1983). The labels given to these sounds are listed in Table I. The everyday sounds were selected to be brief in duration but recognizable. This was confirmed in a pilot study. In that study, the listeners were first allowed to read the list of labels for 50 sounds. After studying the list, the sounds were presented in random order and the listeners were asked to identify each without using the list. All five listeners correctly identified all 50 sounds.

The procedure for synthesizing maskers was the same as in our previous study (Oh and Lutfi, 1998). The everyday sounds and the samples of Gaussian noise were analyzed into 921 spectral components ranging from 0.1 to 10 kHz with a resolution of 11 Hz. On each presentation a different masker waveform of constant total power was synthesized from the magnitude and phase of a fixed number, m , of these spectral components selected at random ($m=2-921$ in different blocks of trials). In different conditions components were selected from the set of 50 common environmental sounds or the 50 samples of Gaussian noise. Also, in different conditions components within a critical band (920-1080 Hz) immediately surrounding the signal were included or excluded. It is common practice to exclude masker components falling within a critical band surrounding the signal so as to reduce energy-based masking of the signal not related to the effects of masker uncertainty (cf. Neff and Green, 1987; Neff and Callaghan, 1988). Finally, the effect of phase was examined by synthesizing masker components using their own phase (preserved-phase condition) or the phase of an unrelated set of randomly selected components (scrambled-phase condition). As phase affects temporal structure it might be expected to influence the amount of informational masking obtained using the everyday sounds. Altogether there were 8 conditions of the study (2 masker types \times 2 critical band conditions \times 2 phase conditions). A masking function of m was obtained for each listener in each condition using the same set 50 Gaussian noise samples and 50 everyday sounds from which to sample the m components.

In all conditions, the signal and maskers were gated on and off together with 10-ms, \cos^2 onset/offset ramps for a

TABLE I. Labels and categories for 50 common environmental sounds for $m=906$ (h=harmonic, am=amplitude modulated, fm=frequency modulated). Note: sounds were categorized according to their most distinct feature as judged by the authors through visual inspection of the spectrograms; many of these sounds have more than one feature.

Sounds rated as easy to identify (5–6)		Sound rated as hard to identify (1–2)		Other sounds (3–4)	
phone ringing	h	zipper closing	fm	bottles falling	am
ray gun	fm	supermarket checkout	am	clock striking	h
pop	am	door creak	am	cloth tear	fm
timer	h	cash register	am	dog barking	am
car horn	h	metal door slamming	am	electric drill	h
gas station bell	h	pans falling	am	glass breaking	am
native drums	am	door slamming	am	golf swings	am
pouring liquid	fm	parking meter	am	gunshots	am
birds chirping	fm	car door closing	am	struck nail	am
siren	h	firecracker	am	pinball machine	am
train warning bell	h	boxes falling	am	sanding wood	am
water in brook	am	paper tear	fm	scooter revving	fm
helicopter hovering	h	baby crying	fm	winding up music box	am
train whistle	h	paper cutter	fm	wood dropping	am
ice tinkling in glass	am	zipper opening	fm	camera shutter	fm
				applause	am
				bowling ball	am
				small waterfall	am
				mailbox squeak	fm
				windshield wipers	am

total duration of 370 ms. Average total power of maskers was kept constant at 60 dB SPL, regardless of the number of masker components. The signal and maskers were computer generated and played over a 16-bit, digital-to-analog converter (Tucker Davis Technologies DD1) at a sampling rate of 44.1 kHz. All stimuli were presented monaurally through Sennheiser model HD-520 headphones to the ear having lower threshold at 1000 Hz. The Sennheiser headphones were calibrated using a loudness balancing procedure with a TDH-49 earphone. All stimuli were presented to individual listeners who were seated in a double-walled, IAC sound attenuation chamber.

C. Procedure

On each trial, 2 of the 50 sound samples were selected at random. For each sample, a fixed number of frequency components was drawn at random on a linear frequency scale. The phases and amplitudes of the selected frequency components were then used to synthesize the maskers played on each trial. The two maskers were presented to listeners separated by a 500-ms silent interval. The signal occurred with equal probability in the first or second interval. Listeners were asked to indicate by key press which one of the two intervals contained the signal. Visual feedback was presented on a CRT after each response indicating whether the response was correct or incorrect. Listeners were instructed to use the feedback to aid their performance. In a further attempt to aid listeners' performance, the signal was presented in quiet preceding each block.

An adaptive, two-interval, forced-choice procedure was used to measure signal threshold in quiet and in the presence of the maskers. The decision rule estimated the 70.7% point on the psychometric function (Levitt, 1971). The level of the signal was controlled by a programmable attenuator (Tucker

Davis Technologies PA4) with an initial step size of 4 dB reduced to 2 dB after the third reversal. A trial block consisted of a total of 12 reversals, and threshold for that block was designated as the average of the last 8 reversal levels. Within an experimental condition, the number of masker components was fixed. Threshold for each condition was the average of five estimates obtained from five consecutive blocks of trials. The order in which experimental conditions were run was randomized. Ten blocks of trials were collected within a single experimental session lasting approximately 1 h with breaks.

II. RESULTS

A. Individual masking functions

Figure 2 shows for each listener the masking functions produced by the masker components sampled from noise (open circles) and those sampled from everyday sounds (filled circles). Each circle gives the total masking for a fixed number of masker components averaged over five blocks of trials with standard errors. The masking functions on the left are for the condition in which masker components within the critical band centered about the signal (920–1080 Hz) were excluded, and those on the right are for the condition in which they were included. The masking functions for noise with critical band components excluded were previously obtained in the study by Oh and Lutfi (1998). There are large individual differences in total masking as is typical for this type of detection task. In each case, however, the masking functions for noise replicate the nonmonotonic relation between amount of masking and the number of masker components. The functions for noise also show a broad peak at

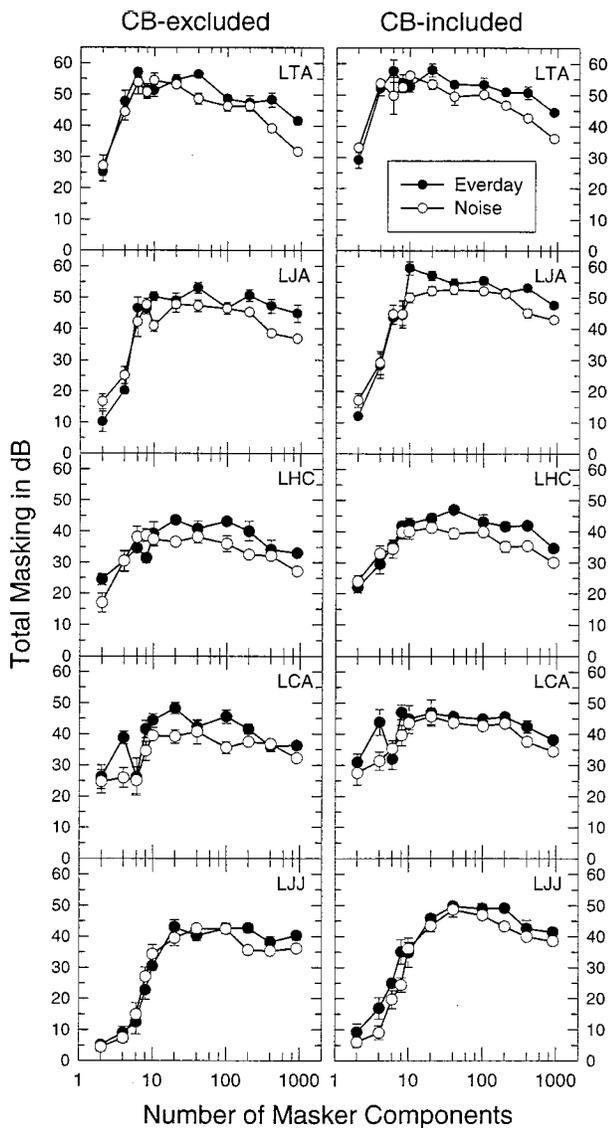


FIG. 2. Individual masking functions with standard errors for the maskers sampled from everyday sounds (filled circles) and Gaussian noise (open circles), and for the conditions in which critical-band components were excluded (left), and included (right).

20–40 components similar to the masking functions obtained in nearly identical conditions by Neff and Callaghan (1988) and Neff and Green (1987).

The masking functions obtained with everyday sounds are similar in shape to those for noise and also peak at between 20–40 components. However, where the number of components is greater than 40 these functions are consistently elevated above those for noise. The difference in masked thresholds for the two types of maskers is greatest where the number of components is the largest $m = 906, 921$). At this point, the maskers sampled from everyday sounds produce as much as 10 dB more masking than those sampled from noise for some listeners. As for the noise, overall less masking is obtained when critical-band components are excluded. However, the general pattern of results is the same whether critical-band components are excluded or included.

Figure 3 shows the effect of phase for the maskers

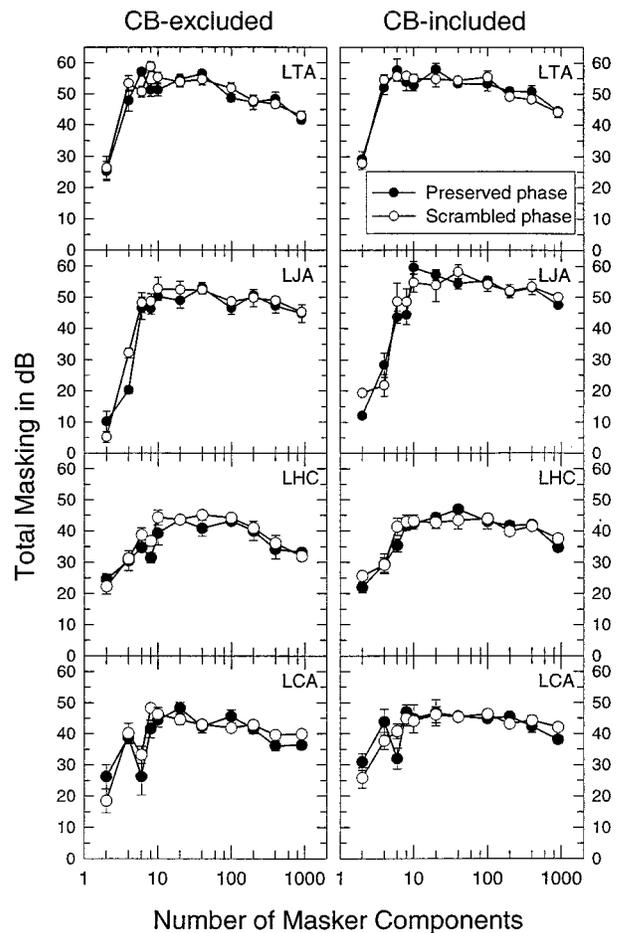


FIG. 3. Individual masking functions with standard errors when the phase of each masker component was preserved (filled circles) and when the phase was scrambled (open circles); critical-band components excluded (left), and included (right).

sampled from everyday sounds. Filled circles give total masking when the phase of each masker component is preserved, and open circles give total masking when the phase is scrambled. Listener LJJ was not able to participate in this part of the study. Scrambling the phase of spectral components of the everyday sounds has little or no influence on listeners' performance. This is true whether critical-band components are excluded or included (right panels).

B. Predictions of the CoRE model

Lutfi (1993) has shown that many results from studies of informational masking conducted over the last 25 years are well described by a model that makes relatively few assumptions about the detection process. According to the component relative entropy (CoRE) model listeners adopt a maximum-likelihood decision rule for any given detection task, but implement this rule imperfectly because of an inability to ignore irrelevant information that varies from one presentation to the next. Specifically, listeners are assumed to integrate information in the maximum-likelihood statistic within an "attentional" bandwidth (W) that far exceeds the physical bandwidth of the signal. For masker components sampled from everyday sounds it is not easy to identify what a maximum-likelihood decision rule might be. Nonetheless,

it is possible to use the model to make predictions under the assumption that the listener's decision rule is not different from that for noise. In this case, individual spectral components of the everyday sounds would be treated as statistically independent when, of course, they are not. However, to the extent that listeners actually ignore these statistical dependencies, the model's predictions should agree with the data obtained for everyday sounds. In this way the role of such statistical dependencies, in particular the role of lawful structure on detection, can be evaluated.

Predictions of the CoRE model for the noise-sampled maskers of the present experiment have been developed by Oh and Lutfi (1998). We review the outcome of this derivation here and refer the reader to the previous paper for details. Under the assumption that individual spectral components of maskers are random and statistically independent the prediction for total masking T is given by

$$T = E + d' \sqrt{n} \sigma_L, \quad (1)$$

where E is the amount of energy-based masking in dB, n is a free parameter representing the number of independent auditory filters over which listeners integrate information within W , $d' \sqrt{n} \sigma_L$ is the predicted amount of informational masking (I), and $d' = 0.78$ is the index of detectability corresponding to 70.7% correct in the adaptive, two-interval, forced-choice task (Levitt, 1971). The amount of energy-based masking E increases monotonically with the number m of masker components. It is estimated, as in the previous paper, using a three-parameter roex filter for a typical normal-hearing adult (Patterson *et al.*, 1982). The predicted amount of informational masking I varies proportionally with σ_L , a measure of the effective spectral variation across maskers within the attentional band. Formally, σ_L is the expected value of the standard deviation in masker level (in dB) at the output of any auditory filter within the bandwidth W . The value of σ_L is computed directly from the power spectra of maskers and varies nonmonotonically with m . The point at which σ_L reaches its maximum value in this case is largely determined by the bandwidth of the auditory filter and the size of the attentional band W .

The top panel of Fig. 4 shows the predictions of the CoRE model for the data averaged across the five listeners when masker components are sampled from noise (open circles, dashed curves) and when they are sampled from everyday sounds (filled symbols, continuous curves). The data and predictions are from the condition where critical-band components are excluded. For comparison to the previous results of Oh and Lutfi (1998), the values of the two free parameters of the model are the same as those used in that study ($W = 0.1 - 2.5$ kHz and $n = 1.64$). The fits to the data are good given that these listeners represent only a subset of the listeners of the previous study. In particular, the model accounts rather nicely for the increasing difference in masked thresholds for the two types of maskers as the number of masker components is increased. The bottom panel of Fig. 4 shows that the predicted difference is due predominately to the variation in σ_L for each type of masker. The bottom two curves in this panel give the predicted amount of informational masking for each type of masker proportional to σ_L .

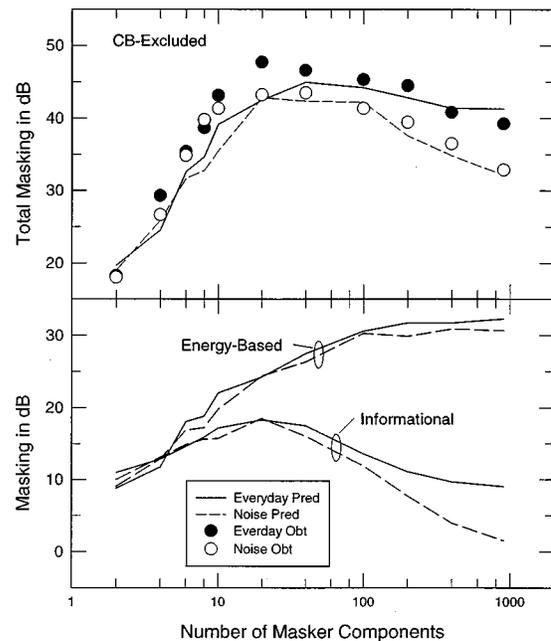


FIG. 4. Top panel: Predictions of the CoRE model for the data averaged across the five listeners when masker components are sampled from noise (open circles=data, dashed curves=prediction) and from everyday sounds (filled symbols=data, continuous curves=predictions) with critical band components excluded. The estimated values of the two free parameters of the model are $W = 0.1 - 2.5$ kHz and $n = 1.64$ (see Oh and Lutfi, 1998 for details regarding the model). Bottom panel: The predictions of the CoRE model in the top panel are broken down into contributing estimates of informational and energy-based masking.

For noise, the spectra of maskers become more alike as the number of masker components m is increased beyond 20. In the limit as m is increased all noise-sampled maskers have essentially flat spectra, hence σ_L and the amount of informational masking decrease and converge to zero (dashed curve). For everyday sounds, the differences among masker spectra are less affected by the number of components sampled, and σ_L does not converge to zero. Hence, the amount of informational masking decreases with m at a slower rate and never reaches zero (continuous curve). These results are to be compared with estimates for energy-based masking which are essentially the same for the two types of maskers (top pair of curves). The prediction for total masking in this case is simply the dB sum of the energy-based and informational masking [cf. Eq. (1)].

To test the models predictions for the effects of energy-based masking, Fig. 5 shows the data and predictions for the maskers sampled from noise with critical-band components included (filled circles, continuous curves) and excluded (open circles, dashed curves). Eliminating critical-band components in this case results in approximately 4–6 dB reduction in masking roughly independent of the number of masker components. The model accounts reasonably well for this result. The model also predicts that the reduction in total masking should be attributed entirely to a reduction in energy-based masking, inasmuch as the estimates of σ_L are not significantly affected by the inclusion or exclusion of a small number of critical-band components (bottom panel of Fig. 5). Similar effects are obtained for the maskers sampled from everyday sounds. Finally, it should be noted that be-

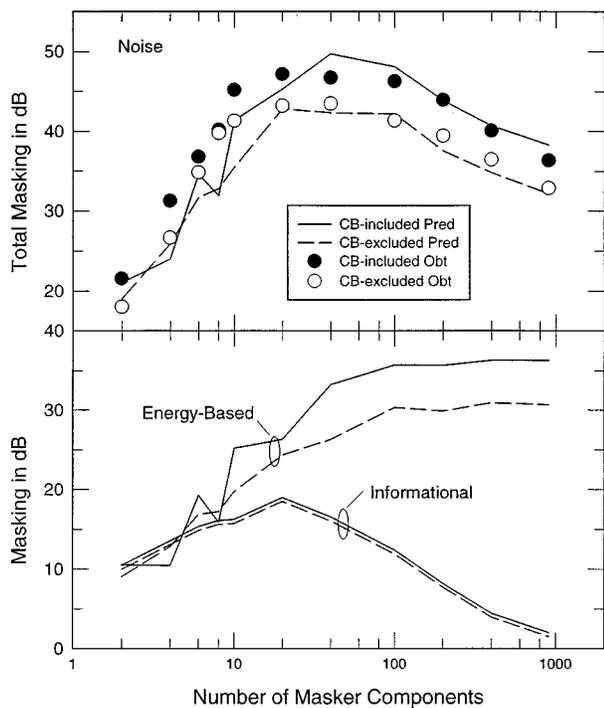


FIG. 5. Same as Fig. 4 except predictions of the CoRE model are given for maskers sampled from noise with critical-band components included (filled symbols) and excluded (open symbols).

cause σ_L is computed exclusively from the magnitude spectra of maskers, the CoRE model predicts no effect of phase for the maskers sampled from everyday sounds (cf. Fig. 2).

C. Effect of familiarity

So far our analysis has focused on differences in the acoustic properties of noise and everyday sounds, their structure and variability. We have yet to consider the possible role of differences in listeners' familiarity with these sounds as maskers. It is well known that masker familiarity can significantly reduce informational masking (Spiegel and Watson, 1981). One might have expected, therefore, that the maskers sampled from everyday sounds would have produced less, not more, masking than those sampled from noise. One would only expect this to be true, however, to the extent that the maskers continued to be recognized as everyday sounds. As fewer components made up these maskers the sounds must have become unrecognizable at some point. Hence, to properly evaluate the effect of familiarity it is necessary to obtain some estimate of how well the maskers were recognized by listeners.

In a followup experiment listeners were asked to rate their ability to recognize the maskers sampled from the everyday sounds (phase preserved). On each trial the listener was presented a single masker sample with its correct label presented on the CRT. The listener was asked to judge on a six-point scale how well they could identify the sound as the given label—1 being “very poor,” and 6 being “very well.” In the early stages of this experiment it appeared that only maskers comprised of more than 100 components would ever be judged as having a measure of recognition better than 1—very poor. We decided, therefore, to concentrate

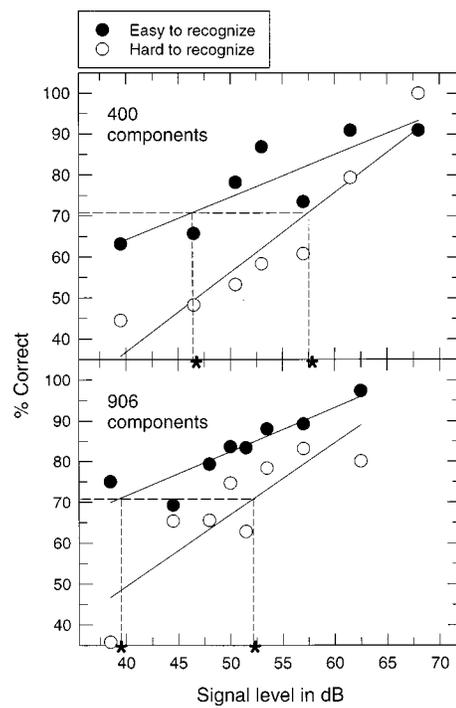


FIG. 6. Psychometric functions estimated from the trial-by-trial data for sounds rated as easy to recognize (filled circles) and for sounds rated as hard to recognize (open circles). The upper panel shows psychometric functions for maskers with 400 components (10 maskers for each category), and the lower panel shows psychometric functions for maskers with 906 components (15 maskers for each category). Asterisks give estimates of signal threshold in each case.

data collection on the two cases, $m=400$ and 906, where listeners more often gave a rating greater than 1. In both of these cases the listeners judged three masker samples of each of the 50 everyday sounds. We then chose for $m=906$ the 15 everyday sounds rated as most easy to identify (rating 5–6) and the 15 rated as least easy to identify (rating 1–2). These are shown in Table I along with an identification of the type of temporal-spectral structure in each case. For $m=400$ we chose only ten sounds in each category; however, the ratings were much the same as for $m=906$. The previously obtained data were then analyzed on a trial-by-trial basis for the maskers sampled from these sounds.

Figure 6 shows the results of this analysis where percent correct versus signal level is plotted for each sound in each category. The figure shows a clear detection advantage for the sounds reported as most easily recognized by listeners (filled circles) compared to those rated as least easily recognized (open circles). Estimates of signal threshold in each case (asterisks) indicate an 11-dB detection advantage for the easily recognized sounds with 400 components, and a 12-dB advantage for the easily recognized sounds with 906 components. The results are suggestive of an effect of familiarity, but are not conclusive. Note, in particular, from Table I that a disproportionate number of the sounds rated as easily recognized also had harmonic structure. It is not possible to say, therefore, whether it was the harmonic structure or the ability to recognize these sounds that was responsible for the detection advantage in these conditions.² We return to this question in the discussion.

III. DISCUSSION

The interest in everyday sounds in this study largely derives from the intrinsic properties that distinguish these sounds from random samples of Gaussian noise—their temporal-spectral structure and to a lesser extent their familiarity. For the unstructured, unfamiliar sound uncertainty is likely to be high, particularly when, as in the case of sparse sampling, only partial information is available. For an everyday sound, however, a familiar pattern of temporal-spectral variation provides a means by which a listener might resolve uncertainty by inference—that is, by using the partial information to draw reasonable expectations regarding the information that is missing. In this way the listener may avoid often confusing spectral components of the signal with those of the noise. The question for the present study is how much information reduction is possible before this process breaks down. The answer based on these results appears to be very little. Even with nearly complete information regarding the sound ($m > 200$), listener performance is consistent with that of an observer that treats the everyday sounds as having the same statistical properties as Gaussian noise. Hence, at least for these conditions, listeners do not appear able to take advantage of lawful or known relations among individual spectral components of the maskers to segregate those components from the signal. This conclusion is further reinforced by the failure to show any effect of scrambling phase in this study.

The present results contrast with those from past studies suggesting that listeners require very little information in a natural sound to reduce uncertainty regarding the source. Freed (1990), for example, indicates that listeners are able to judge the hardness of mallets used to strike metal cooking pans based solely on the relative ratio of high- to low-frequency energy in the sound. Similarly, Repp (1987) reports that listeners are able to distinguish among different hand positions based on the presence and extent of low-frequency peaks in hand-clap spectra. Li *et al.* (1991) suggest that listeners use the shape of the spectral peaks to identify the gender of a walker, and Warren and Verbrugge (1984) claim that listeners can distinguish between bottles breaking or bouncing based solely the time pattern of the sounds. Given such reports, one might have expected that fewer spectral components of the everyday sounds would be required to reduce masker uncertainty in this study.

Why then were the listeners of the present experiment not better able to exploit the partial information from the everyday sounds? It seems unlikely that insufficient information was the reason. Consider, for example, that for $m = 921$ the frequency spacing among components is 11 Hz—too small to expect that individual spectral components would be separately resolved and that the temporal-spectral structure would, therefore, not be available. There is an alternative explanation. In all past studies showing an effect of structure, the structure itself was never in doubt to the listener. The common practice has been to present one and the same structure across an entire block of trials. Such is not the case in the present study for the everyday sounds. We have, for example, identified three types of structure, each very different from the next, that have different probabilities of

occurrence on any given trial. Even within these types, the differences in structure are much greater than have existed in the past studies. It seems likely then that our failure to show a substantial detection advantage with everyday sounds reflects the listener's continued uncertainty regarding structure, rendering such structure ineffective. The obvious followup study would be a replication wherein masker components are randomly selected from a single everyday sound with clearly recognized structure. Uncertainty regarding noise structure has not been adequately addressed in past studies of tone-in-noise masking. This would seem an important next step as everyday listening must involve varying degrees of knowledge regarding not only the individual waveforms, but also the general temporal-spectral structure of the ambient noise.

Finally, there is one result of this study that deserves special consideration. For the subset of the everyday sounds rated as easy to recognize by these listeners ($m = 400, 906$) substantially less informational masking was found than for the subset rated as difficult to recognize. There at least two ways to interpret this result. The first, and most obvious, is to conclude that recognition is a necessary condition for an effect of structure. The notion is appealing, but is not supported by studies using noise. Hall *et al.* (1984), for example, have shown that the release from tone masking resulting from comodulation of noise components increases as more components are added to the noise. What is being "recognized" in this case is clearly not individual noise waveforms, but rather some aspect of the noise structure that allows the listener to segregate the signal from the noise as more noise components are made available. So too we might expect that a listener could take advantage of structure in a collection of everyday sounds at some point without having to recognize the individual waveforms. This leads to the second interpretation which is that structure is responsible for the release from masking and recognition is merely a by-product of an easily resolved structure. While this later interpretation appears more likely, the present data do not weigh heavily in favor of one or the other interpretation with regard to everyday sounds.

IV. SUMMARY

The results of the present study replicate earlier results showing a nonmonotonic relation between the amount of masking of a tonal signal and the number m of masker components randomly sampled from Gaussian noise. The new result is that for $m > 40$ the amount of masking is 4–10 dB greater when maskers, equal in power to the noise, are sampled from a small set of recognizable everyday sounds. A follow-up experiment also showed that the noise advantage is essentially eliminated for everyday maskers rated by listeners as easy to recognize. Except perhaps for this later result, the data are well described by a model in which the everyday sounds are treated as having the same statistical properties as the Gaussian noise. The results suggest that uncertainty regarding stimulus structure may play at least an important role in informational masking as uncertainty regarding waveforms with common structure.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Sid Bacon and an anonymous reviewer for helpful comments on an earlier version of this manuscript. This research was supported by a grant from the NIDCD (Grant No. R01 CD01262-09).

¹Note, we use the term “structure” rather loosely to refer to any spectral-temporal regularity in the sound whether it be imposed by a real or synthesized source. It is not our intent to provide a rigorous classification of structures. We chose to focus on these three types of structures inasmuch as they are most commonly used in the literature to characterize everyday sounds, and were indeed representative of the sample used in the present study as determined from visual inspection of the spectrograms.

²Note that because of these differences in spectra, it is also possible that the CoRE model would predict a detection advantage for the easy to recognize sounds. Predictions in this case would be derived by applying the decision rule [Eq. (6) of Oh and Lutfi, 1998] to the trial-by-trial data, and comparing percent correct for the easy and hard to recognize sounds. We have not undertaken this analysis here, but have chosen instead to pursue the potential effect of harmonicity in a subsequent study (see Oh and Lutfi, 1997).

ANSI (1989). ANSI S3.6-1989, “Specifications for audiometers” (ANSI, New York).

Freed, D. J. (1990). “Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events,” *J. Acoust. Soc. Am.* **87**, 311–322.

Green, D. M. (1988). *Profile Analysis: Intensity Discrimination* (Oxford U.P., New York).

Grose, J. H., and Hall, J. W. (1990). “The effect of modulation coherence on signal threshold in frequency-modulated noise bands,” *J. Acoust. Soc. Am.* **88**, 703–710.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (1984). “Detection in noise by spectro-temporal pattern analysis,” *J. Acoust. Soc. Am.* **76**, 50–56.

Levitt, H. (1971). “Transformed up-down methods in psychoacoustics,” *J. Acoust. Soc. Am.* **49**, 467–477.

Li, X., Logan, R. J., and Pastore, R. E. (1991). “Perception of acoustic source characteristics: Walking sounds,” *J. Acoust. Soc. Am.* **90**, 3036–3049.

Lutfi, R. A. (1993). “A model of auditory pattern analysis based on component-relative-entropy,” *J. Acoust. Soc. Am.* **94**, 748–758.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1986). “Thresholds for hearing mistuned partials as separate tones in harmonic complexes,” *J. Acoust. Soc. Am.* **80**, 479–483.

Neff, D. L., and Callaghan, B. P. (1988). “Effective properties of multicomponent simultaneous maskers under conditions of uncertainty,” *J. Acoust. Soc. Am.* **83**, 1833–1838.

Neff, D. L., and Green, D. M. (1987). “Masking produced by spectral uncertainty with multicomponent maskers,” *Percept. Psychophys.* **41**(5), 409–415.

Oh, E., and Lutfi, R. A. (1997). “Reducing the effects of masker uncertainty with harmonicity and onset/offset synchrony,” *J. Acoust. Soc. Am.* **102**, 3160.

Oh, E. L., and Lutfi, R. A. (1998). “Nonmonotonicity of informational masking,” *J. Acoust. Soc. Am.* **104**, 3489–3499.

Patterson, R. D. (1976). “Auditory filter shapes derived with noise stimuli,” *J. Acoust. Soc. Am.* **59**, 640–654.

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). “The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold,” *J. Acoust. Soc. Am.* **72**, 1788–1803.

Pollack, I. (1975). “Auditory informational masking,” *J. Acoust. Soc. Am.* **57**, S5.

Repp, B. H. (1987). “The sound of two hands clapping: An exploratory study,” *J. Acoust. Soc. Am.* **81**, 1100–1109.

Spiegel, M. F., and Watson, C. S. (1981). “Factors in the discrimination of tonal patterns. III: Frequency discrimination with components of well-learned patterns,” *J. Acoust. Soc. Am.* **69**, 223–230.

Warren, W. H., and Verbrugge, R. R. (1984). “Auditory perception of breaking and bouncing events: Psychophysics,” *J. Acoust. Soc. Am.* **10**, 704–712.

Watson, C. S., and Kelly, W. J. (1981). “The role of stimulus uncertainty in the discrimination of auditory patterns,” in *Auditory and Visual Pattern Recognition*, edited by D. J. Getty and J. H. Howard, Jr. (Erlbaum, Hillsdale, NJ).

Watson, C. S., Kelly, W. J., and Wroton, H. W. (1976). “Factors in the discrimination of tonal patterns. II: Selective attention and learning under various levels of stimulus uncertainty,” *J. Acoust. Soc. Am.* **60**, 1176–1186.

The auditory continuity phenomenon: Role of temporal sequence structure^{a)}

Carolyn Drake^{b)}

Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, EPHE, 28 rue Serpente, F-75006, Paris, France

Stephen McAdams

Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, EPHE, 28 rue Serpente, F-75006, Paris, France and Institut de Recherche et de Coordination Acoustique/Musique (IRCAM), 1 place Igor-Stravinsky, F-75004, Paris, France

(Received 19 February 1999; accepted for publication 4 August 1999)

The auditory continuity phenomenon occurs when listening to a series of alternating high- and low-level tones: instead of perceiving this intermittence, listeners often report hearing a continuous tone upon which is superimposed a series of intermittent tones. The temporal limits to the perception of this phenomenon are investigated, as well as effects of task instructions on it. A loudness-matching paradigm developed previously [McAdams *et al.*, *J. Acoust. Soc. Am.* **130**, 1580–1591 (1998)] provided both an objective indication of this phenomenon and an indication of its strength. In the studies reported here, the phenomenon was observed reliably when the low-level tones were at least half the duration of the high-level tones: the greater the duration ratio between low-level and high-level tones, the stronger the phenomenon. Duty-cycle duration did not affect the strength of this phenomenon. It proved to be particularly robust as its strength was unaffected by task instructions, level of expertise, and repetition. A model is proposed in which adjusted levels depend on the relative levels of the high- and low-level tones weighted by relative duration and attentional focusing. © 1999 Acoustical Society of America. [S0001-4966(99)00412-9]

PACS numbers: 43.66.Cb, 43.66.Mk, 43.66 Lj [SPB]

INTRODUCTION

The phenomenon of auditory continuity or auditory induction may occur when listening to a series of alternating high- and low-level tones: instead of perceiving this alternation, listeners often report hearing a continuous low-level tone upon which is superimposed a series of intermittent tones (Thurlow and Elfner, 1959; Warren, Obusek, and Ackroff, 1972). Traditional loudness models (Zwicker, 1977) have difficulty accounting for this continuity phenomenon (henceforth referred to simply as “the phenomenon”). However, Warren (1982) proposed an explanation in terms of a subtractive process derived from his auditory induction theory which proposes that a portion of the internal representation of the louder sound is subtracted or reallocated to serve as a substrate for synthesis of the apparently continuous fainter sound. Bregman (1990) has also proposed a theoretical framework which can interpret these findings: the auditory system organizes the auditory scene before computing perceptual attributes of events. Thus, the phenomenon occurs because of a bottom-up perceptual mechanism which he has called the “old-plus-new” listening strategy: the “old” information is composed of the low-level tones and

an equivalent part of the high-level tones, and the “new” information is composed of the remaining part of the high-level tones. The auditory system considers that the old information does not stop, but continues throughout the time period occupied by the new information. An interpolation is performed between the properties of the low-level sounds occurring before and after the interrupting sound (Ciocca and Bregman, 1987). Thus, the high-level tones are in fact separated into two parts: one attributed to the continuous tone and one attributed to the intermittent tones.

The old-plus-new strategy allows two predictions concerning the perception of the phenomenon. First, it should only be perceived if the old events (the low-level tones) are well established. Thus, the longer the old tones are in relation to the new tones, the stronger should be the phenomenon. Second, it should be very robust because it results from a relatively low-level process, occurring before loudness attribution. Thus, it should be: (1) observed in all subjects, (2) unaffected by task constraints and instructions, and (3) unaffected by training and practice.

Previous studies have examined some limiting factors to the perception of this phenomenon: it does not occur if the high- and low-level tones are separated from each other by short pauses and nonabrupt transitions, or are presented to separate ears. It only occurs if the two sets of tones differ by certain level differences and are in the same spectral range (Bregman and Rousseau, 1991; Houtgast, 1972; McAdams, Botte, and Drake, 1998; Thurlow, 1957; van Noorden, 1977; Verschuure, Rodenburg, and Maas, 1976; Warren, 1984; Warren *et al.*, 1994; Warren, 1999). To our knowledge, few

^{a)}Preliminary work leading to this study was reported at the 4th French Congress of Acoustics, Marseille (C. Drake and S. McAdams, “The Continuity illusion: Role of temporal sequence structure,” *Actes du He Congrès Français d’Acoustique*, Marseille, 1997, Vol. 1, pp. 519–522).

^{b)}Address correspondence to C. Drake, Laboratoire de Psychologie Expérimentale, 28 rue Serpente, 75006, Paris, France. Electronic mail: drake@idf.ext.jussieu.fr

studies have systematically studied the role of temporal factors in the continuity phenomenon. McAdams *et al.* (1994, 1998), for example, found that the phenomenon was observed under all temporal conditions examined: the high-level tones were equal to or shorter in duration than the low-level tones (high/low duration ratios of 100/100, 200/200, 100/300, and 100/700 ms). In the present study, we examine two temporal factors that may play a determining role in whether or not the continuity phenomenon is observed: the relative duration of high- and low-level tones and the total duty-cycle duration (duration of high-level tone plus duration of low-level tone). We therefore examine instances where the high-level tones are both longer *and* shorter than the low-level tones and where the duty cycle ranges from short (150 ms) to long (2100 ms). We also examine the robustness of the phenomenon by comparing the role of instructions given to the subjects (instructions explicitly concerning the continuity phenomenon and instructions emphasizing the rhythmic structure without reference to the phenomenon). In addition, we also investigated the role of experience in this sort of task (by comparing experts and novices), and eventual improvements over repeated measures.

We establish whether or not listeners perceived the phenomenon by using a paradigm developed previously (McAdams *et al.*, 1998). A similar paradigm was developed independently by Warren *et al.* (1994), except that they used a modified method-of-limits procedure rather than an adjustment paradigm. In our paradigm, we measured the perceived loudness of various components of the stimulus: listeners were asked to adjust the loudness of (a) a continuous sound to equal that of the perceived continuous tone in the reference sequence, and (b) the loudness of an intermittent sequence to equal that of the perceived intermittent sequence in the reference sequence. Two distinct adjustment patterns were observed depending on the percept. When listeners perceived the phenomenon, the level of the continuous tone was adjusted very near the level of the low tones, but the level of the intermittent tone was systematically adjusted well below that of the high tones, as predicted by the old-plus-new hypothesis. When listeners did not perceive the phenomenon, the level of both components of the reference sequence was adjusted near the physical values. In the present study, we use this method to identify the temporal conditions under which listeners perceive the continuity phenomenon.

I. EXPERIMENT 1: RELATIVE DURATION OF HIGH- AND LOW-LEVEL TONES

Previous studies have shown that the phenomenon is perceived when the low-level tones (old events alone) have the same duration as or are longer than the high-level tones (old and new events together). We predict that the phenomenon will disappear in the opposite configuration when the low-level tones are shorter than the high-level tones, because the old events are not temporally well established. We therefore measure listeners' loudness matches to parts of the reference sequence in order to establish the conditions under which they perceive the phenomenon. We identify whether or not listeners perceive the phenomenon by measuring the precision with which they match the levels: previous work

has established two distinct level-matching patterns when subjects perceive and do not perceive the phenomenon. Two groups of subjects were examined (novices and experts) to test the role of expertise, and six repeated measures were made to test the role of practice.

A. Method

1. Subjects

All eight subjects reported normal hearing. They were all psychology students at the Université René Descartes. Subjects were divided into two equal-sized groups: the "novices" had not previously participated in psychoacoustic experiments, and the "experts" had previously participated in auditory experiments (at least 10 h, but not on the continuity phenomenon) and were musically experienced (had been playing an instrument for at least 5 years).

2. Apparatus

Sinusoidal signals were synthesized at a sampling rate of 20 kHz with 16-bit resolution on an Oros DSP card controlled by a Compaq computer. In the case of alternating-level signals, high and low bursts were processed in different channels. The signals were then filtered with a Kemo VBF/24 bandpass filter with cutoff frequencies of 930 and 1070 Hz and -48 dB/oct slopes. The filtering of the signals served for antialiasing. The filtered signals were then routed through Charybis D programmable attenuators with 0.25-dB resolution that were controlled by the computer. The final signals were mixed and sent to both earpieces of a TDH-49 headset. Experimental sessions took place in an IAC single-walled sound-isolation chamber. Subjects adjusted levels for comparison stimuli with a single-turn potentiometer and signaled their satisfaction with the adjustment by pressing a button on the response box. At this point the computer presented the next trial. Levels at each earpiece were verified using a Bruel & Kjaer 4153 artificial ear and 2230 sound-pressure meter.

3. Stimuli

Sequences were composed of 1-kHz pure tones with overlapping 5-ms linear onset and offset ramps. Reference sequences with alternating levels were composed of eight low-level tones interleaved with seven high-level tones. The duration of the short tones was 50 or 100 ms, and the duration of the long tones was 300, 700, or 1500 ms. All combinations of short and long tones were tested, giving 12 duration conditions illustrated in Fig. 1. Six of these duration conditions corresponded to the "classic" configuration where the high-level tones were shorter than the low-level tones (in which continuity is usually reported), and the other six duration conditions corresponded to an "inverted" configuration where the high-level tones were longer than the low-level tones. The total duration of the sequences depended on the duration combinations, varying between 2.5 and 11.2 s. The sound level of the low-level part of the reference stimulus was varied randomly from trial to trial within the set $60 \pm \{1,3,5\}$ dB. In all analyses and plots, lev-

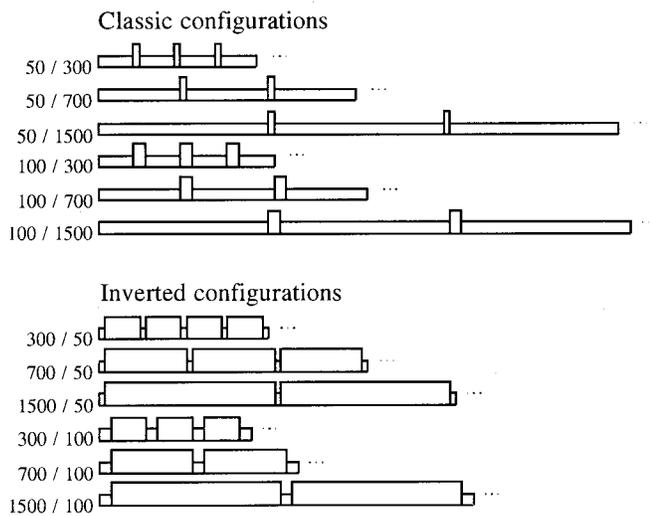


FIG. 1. Experiments 1 and 3: Stimuli used in the classic and inverted configurations (high- and low-level tone durations in ms).

els are presented relative to the mean of the roving range (60 dB). The high-level tones were 6 dB higher than the low-level tones.

There were two types of comparison sequences: (1) a continuous tone of the same total duration as the reference sequence, or (2) an intermittent sequence with tones having the same duration as the high-level tones of the reference sequence and being separated by pauses of the same duration as the low-level tones of that sequence. They were presented at an initial level of $60 \pm \{7, 8, 9, 10\}$ dB to avoid response bias. The reference and comparison sequences were separated by 800 ms, and the next reference sequence followed 1500 ms later.

4. Procedure

Half the subjects (two experts and two novices) completed experiment 1 the first day, and then experiment 3 about a week later. The other half of the subjects completed the two experiments in the opposite order. The first experimental session was preceded by a *familiarization phase* in which all the stimuli were presented to the subjects. They were questioned as to what they heard in order to verify whether or not auditory continuity was perceived for each duration condition. The familiarization phase was followed by a *test phase* composed of 72 trials (6 duration conditions, 2 stimulus configurations, 6 repetitions). Each trial consisted of the repeated alternation between the reference and comparison sequences. During this alternation, the level of the comparison sequence could be adjusted with a single-turn potentiometer. Subjects could listen to the alternation as many times as necessary to make a satisfactory match, at which point they signaled the computer to record the level of the comparison sequence by pressing a button. The listener aligned the turnpot to a fixed reference point at the beginning of each trial. Subjects were instructed to (a) match the loudness of the intermittent comparison sequence to that of what they perceived as intermittent in the reference sequence, or (b) match the loudness of the continuous comparison sequence to that of what they perceived as continuous in the

reference sequence. The stimuli were presented in a random order. Six measures were made in each experimental condition. The dependent measure was the adjustment "error" (in dB) between the adjusted level of the comparison sequence (intermittent level = L_I ; continuous level = L_C) and that of the targeted part of the reference sequence (low level = L_L ; high level = L_H) (error = $L_L - L_C$ or $L_H - L_I$). The experiment lasted about 2 h per subject, with a pause halfway through the session.

B. Results

1. Subjective descriptions

In the familiarization phase at the beginning of the first experimental session, subjects were asked to describe the 12 sequences. No reference was made to the continuity phenomenon. In the classic configurations, 94% of the responses indicated the phenomenon, whereas only 42% indicated the phenomenon in the inverted configurations. The percentage reports of the phenomenon were more divergent across the two classes of stimuli for the expert group than for the novice group (classic: novices = 88% and experts = 100%, inverted: novices = 63% and experts = 21%), and were independent of the temporal combination.

2. Classic configuration

Figure 2 shows the mean adjusted levels for the classic and inverted duration conditions. For the *continuous* adjustable tones, matches were extremely accurate (mean adjustment = 59.9 dB; mean adjustment error = $59.9 - 60.0 = -0.1$ dB). A single-sample *t*-test revealed that the mean adjustment error did not differ significantly from zero. The matched level did not vary with the duration conditions: a mixed analysis of variance (ANOVA) on the mean adjustment errors by group (novices/experts), duration of high-level tones (50 or 100 ms), duration of low-level tones (300, 700, or 1500 ms), and repetition (6) revealed no significant differences.

However, for the *intermittent* adjustable sequence, matches were quite different from the physical levels (mean adjustment = 60.8 dB; mean adjustment error = $60.8 - 66.0 = -5.2$ dB), indicating that subjects systematically *underestimated* the level of the intermittent tones. A single-sample *t*-test revealed that the mean adjustment error differed significantly from zero [$t(7) = 5.6$; $p < 0.001$]. A mixed ANOVA on the mean adjustment errors by group (novices/experts), duration of high-level tones (50 or 100 ms), duration of low-level tones (300, 700, or 1500 ms), and repetition (6) revealed no significant differences between the two groups of subjects or over the six repetitions. The duration of the high-level tones did not influence the matches but the duration of the low-level tones did [$F(2, 12) = 9.17$, $p < 0.01$]; Fig. 2 shows that the longer the low-level tones, the greater the adjustment error (difference from L_H).

Thus, in these classic conditions, adjustment errors corresponded to the patterns expected if subjects were experiencing the continuity phenomenon. The phenomenon was greater when the low-level tones were longer.

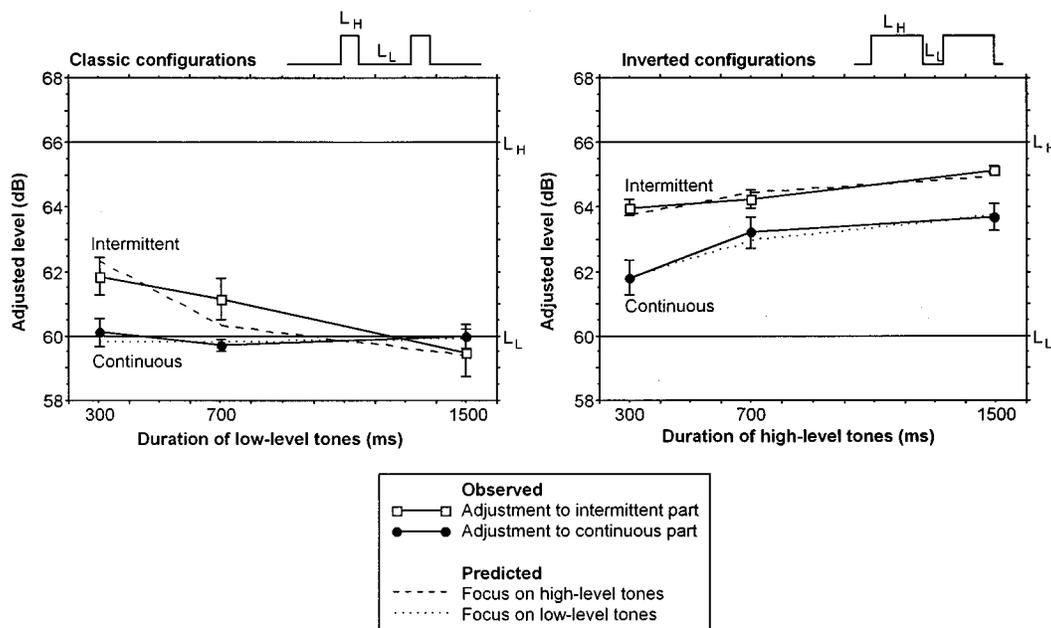


FIG. 2. Experiment 1: Solid-line curves represent mean levels adjusted to match the loudness of the continuous and intermittent comparison sequences for the six duration conditions. The data are averaged over groups, repetitions, and the short-tone durations (50 and 100 ms). Note that the abscissa corresponds to the durations of the longer tones (low-level for the classic configurations and high-level for the inverted configurations). Horizontal lines indicate physical levels of low- and high-level tones in the reference sequences. Error bars (where visible) indicate ± 1 standard error. Dashed and dotted curves represent predictions based on subject means weighted for duration with attentional focus on high-level (dashed) and low-level (dotted) parts of the reference sequence (see Sec. V).

3. Inverted configuration

For the *continuous* adjustable tones, systematic adjustment errors were observed (mean adjustment = 62.9 dB; mean adjustment error = 62.9 - 60.0 = +2.9 dB), indicating that contrary to the classic conditions, subjects systematically *overestimated* the level of the continuous tones. A single-sample *t*-test revealed that the mean adjustment error differed significantly from zero [$t(7) = 7.4$, $p < 0.001$]. A mixed ANOVA on the mean adjustment errors by group (novices/experts), duration of high-level tones (300, 700, or 1500 ms), duration of low-level tones (50 or 100 ms), and repetition (6) revealed only a significant effect of high-level tone duration [$F(2,12) = 10.24$, $p < 0.01$]: the longer the high-level tone, the higher the level match of the continuous tones.

For the *intermittent* adjustable sequence, systematic matches were also observed (mean adjustment = 64.5 dB; mean adjustment error = 64.5 - 66.0 = -1.5 dB), indicating that subjects systematically *underestimated* the level of the intermittent tones. A single-sample *t*-test revealed that the mean adjustment error differed significantly from zero [$t(7) = 8.7$, $p < 0.001$]. A mixed ANOVA on the mean adjustment errors by group (novices/experts), duration of low-level tones (50 or 100 ms), duration of high-level tones (300, 700, or 1500 ms), and repetition (6) revealed no significant differences between the two groups of subjects or over the six repetitions. As for the continuous adjustable tones, the duration of the low-level tones did not influence the matches but the duration of the high-level tones did [$F(2,12) = 22.96$, $p < 0.001$]: the longer the high-level tone, the higher the adjusted level of the intermittent tones.

Thus, in these inverted conditions, adjustment errors did

not correspond to the patterns expected if subjects were experiencing the continuity phenomenon. Instead, the matches suggest that subjects were adjusting to an intermediate level, the estimation of which was affected by the duration of the high-level tones.

C. Discussion

We have identified some temporal constraints on the continuity phenomenon: it is only observed if the old information temporally dominates the new information, that is, the low-level tones must be equal to or longer than the high-level tones. The strength of the phenomenon varied with the temporal sequence structure. Greater underestimations of the level of the intermittent sequence were observed when the low-level tones were longer, suggesting a stronger phenomenon. This could be due either to the ratio between the durations of the high- and low-level tones, or to the increased duration of the duty cycle. Experiment 2 will separate these two factors. However, when the phenomenon was not perceived, listeners adjusted both the high- and low-level tones towards an intermediate level which increased with longer durations. The adjustment behaviors are thus both quantitatively and qualitatively different for the two classes of stimuli (those that induce the phenomenon and those that do not).

In order to investigate the relation between subjective reports of the illusion and adjustment errors, two adjustment data sets were created: for each of the 12 temporal conditions, subjects were classified according to whether they reported the phenomenon or not. Very similar patterns of over- and underestimation were observed in the two data sets, once

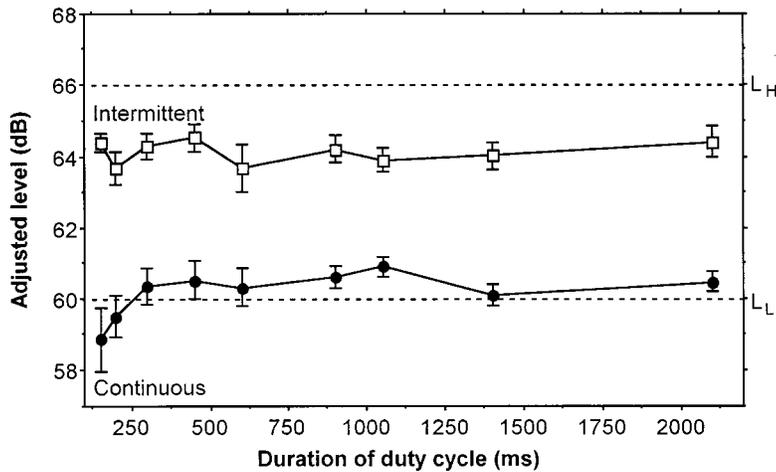


FIG. 3. Experiment 2: Mean levels adjusted for the continuous and intermittent comparison sequences adjusted to the classic, equal, and inverted configurations, where the duty-cycle duration was varied systematically. The data are averaged over high/low duration ratios and repetitions. Error bars indicate ± 1 standard error.

again emphasizing the fundamental nature of this phenomenon.

For the classic conditions, the phenomenon was surprisingly robust because it was not influenced by listeners' experience with listening to sequences (novice/expert), nor by repeated exposure to the phenomenon (repetitions). However, the instructions in this experiment may have facilitated the phenomenon because subjects were asked to adjust a continuous tone, thus emphasizing the phenomenon. In fact, this may have led to confusion for the listeners in the inverted conditions where they reported hearing the continuity phenomenon only 42% of the time. Instructions which do not orient the listener toward the phenomenon will be used in experiment 3.

II. EXPERIMENT 2: DUTY-CYCLE DURATION OR DURATION RATIO?

In experiment 1, we examined two extreme temporal conditions: in the classic configuration the high-level tones were very short (50 or 100 ms) and in the inverted configuration the low-level tones were very short (50 or 100 ms). We observed two distinct percepts (phenomenon or no phenomenon). But what happens between these two extremes? Does the phenomenon become progressively less strong as the temporal conditions change or is there an abrupt passage from the presence to the absence of the phenomenon? We tested these two possibilities by examining a range of sequences which varies systematically along two continua: the total duty-cycle duration and the duration ratio between high- and low-level tones.

A. Method

Eight new novice subjects with normal hearing participated in the experiment. They were all psychology students at the Université René Descartes. The same apparatus and procedure was used as in experiment 1. Fifteen sequences were constructed which varied in the duty-cycle duration and ratio between the duration of high- and low-level tones. Three duration ratios were examined: "classic," where the high-level tones were half the duration of the low-level tones (ratio 1:2), "equal," where the high- and low-level tones were the same duration (ratio 1:1), and "inverted," where

the high-level tones were twice as long as the low-level tones (ratio 2:1). For each duration ratio, the duty cycle varied in total duration (ratios 1:2 and 2:1—50/100, 100/200, 150/300, 300/600, 350/700, 700/1400 ms, ratio 1:1—100/100, 300/300, 700/700 ms). All other experimental details were identical to experiment 1.

B. Results

1. Subjective descriptions

Subjects were asked to describe each temporal combination. Contrary to expectations (the phenomenon should only be perceived if old temporally dominates the new, i.e., only in classic and equal conditions), all subjects described perceiving the phenomenon in all 15 duration conditions, with the exception of two subjects in the 50/100-ms condition who were unable to describe the sequence.

2. Duty-cycle duration

Figure 3 shows the mean adjusted levels averaged across the three duration ratios for each of the duty-cycle durations. In accordance with the subjective descriptions, adjusted levels for all 15 duration conditions followed the pattern associated with the perceived phenomenon (continuous correctly matched, intermittent underestimated). In all cases, for the continuous comparison sequences, a single-sample *t*-test revealed that the mean adjustment errors (difference from L_L) did not differ significantly from zero, and for the intermittent comparison sequence, a single-sample *t*-test revealed that the mean adjustment errors (difference from L_H) differed significantly from zero [$t(7) = 6.7$, $p < 0.001$]. Noteworthy is the fact that the size of the phenomenon (degree of underestimation) is constant across the different duration conditions, that is, it is unaffected by the duty-cycle duration and duration ratios examined here. A repeated-measures ANOVA on the mean adjustment errors for the continuous and intermittent comparison sequences by condition (15) revealed no significant difference between the conditions. Whereas only a limited range of ratios was examined, limiting conclusions for the moment, the duty-cycle durations covered the wide range of durations which could feasibly induce the phenomenon. Duty-cycle duration did not therefore appear to influence the perception of the phenomenon.

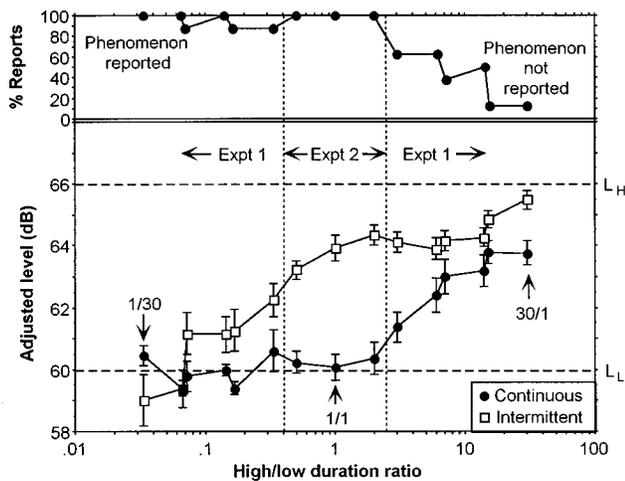


FIG. 4. Experiments 1 and 2: Percentage reports of the continuity phenomenon (upper panel) and mean adjusted levels for the continuous and intermittent comparison sequences (lower panel) as a function of the ratio between the durations of high- and low-level tones. The data are averaged over three duty-cycle durations, groups, and repetitions. Error bars indicate ± 1 standard error.

3. Duration ratios

In order to compare a wider range of duration ratios, Fig. 4 presents the mean adjusted levels and subjective descriptions for the 12 ratios used in experiment 1 and the three ratios used in experiment 2. The ratios ranged here from 1:30 (classic—50/1500 ms) to 30:1 (inverted—1500/50 ms), the wide range of duration ratios which could feasibly induce the phenomenon. The two predicted matching patterns were observed which corresponded with the subjective descriptions: that for the continuity phenomenon on the left of the figure, and that without the phenomenon on the right of the figure. Three points are noteworthy. First, as predicted, the phenomenon was observed for all ratios in which the old was longer than the new, but also for the inverted configuration where the new was twice as long as the old (2:1). Above this ratio, the phenomenon was less frequently observed. Second, when the phenomenon was perceived, the smaller the high/low duration ratio, the greater the underestimation of the level of the intermittent sequence. Underestimation was greatest for the highest ratios (-6 dB for 1:30 and 1:14), decreasing progressively to -2 dB for the 1:2, 1:1, and 2:1 ratios. Separate repeated-measures ANOVAs for the continuous and intermittent sequences on the mean adjustment errors for the seven ratios yielding reports of the phenomenon revealed a significant main effect of ratio for the intermittent comparison sequences [$F(1,6) = 14.6, p < 0.001$] but not the continuous sequences. Third, when the phenomenon was not perceived, the adjustments appeared to correspond to a global mean level: the adjustments increased gradually as the relative duration of the high tones increased, thus corresponding to an increase in mean level. A repeated-measures ANOVA on the mean adjustment errors for the four largest high/low duration ratios yielding fewer reports of the phenomenon revealed significant increases for both the continuous and intermittent comparison sequences.

C. Discussion

The results from experiments 1 and 2 indicate that the continuity phenomenon only occurred when the duration of the high-level tones was not too long in relation to the low-level tones: they must be in a ratio of at most 2:1. In these conditions, the smaller the ratio, the stronger the phenomenon. The duty-cycle duration does not influence the strength of the phenomenon. There was not an abrupt transition from the perception of the phenomenon to the absence of the phenomenon, but rather a gradual change from one state to the other, with a gradual decrease in the strength of the phenomenon as indicated by the decreases in the underestimation of the level of the intermittent sequence.

III. EXPERIMENT 3: INSTRUCTIONS EMPHASIZING THE RHYTHMIC STRUCTURE

In experiments 1 and 2, the instructions could be considered as facilitating the phenomenon because subjects were asked to adjust the intensity of a continuous comparison tone, thus emphasizing the phenomenon. Here, we test its robustness by comparing the previous results with those obtained with instructions which draw listeners' attention to the rhythmic structure rather than to an eventual continuity: listeners were required to adjust two intermittent comparison sequences, each corresponding to the duration of either the high- or low-level tones in the reference sequence. The rhythmic task might be considered to require listeners to focus on either the high- or the low-level tones independently in the reference sequence in order to match levels with the comparison sequence. However, if the phenomenon is compelling and unavoidable, then adjusting the comparison sequence with the low-level rhythm should correspond to the same pattern of results as in experiment 1, otherwise one might expect listeners to correctly match the physical levels in the reference sequence.

A. Method

The same eight subjects (four novices and four experts) as in experiment 1 participated in this experiment. Half the subjects (two experts and two novices) completed experiment 1 the first day, and then experiment 3 about a week later. The other half of the subjects did the two experiments in the opposite order. The same apparatus was used as in experiments 1 and 2. The reference sequences were identical to those in experiment 1. The comparison sequences were intermittent tones with the same onset and offset durations as either (a) the high-level tones or (b) the low-level tones in the reference sequences. They thus emphasized the rhythmic structure rather than the continuity phenomenon. All other details were as described previously. Subjects were instructed to adjust the level of the intermittent sequence to match that of the tones with the same rhythm in the reference sequence.

B. Results and discussion

1. Subjective descriptions

As for experiment 1, subjects were asked to describe the 12 sequences at the beginning of the first experimental ses-

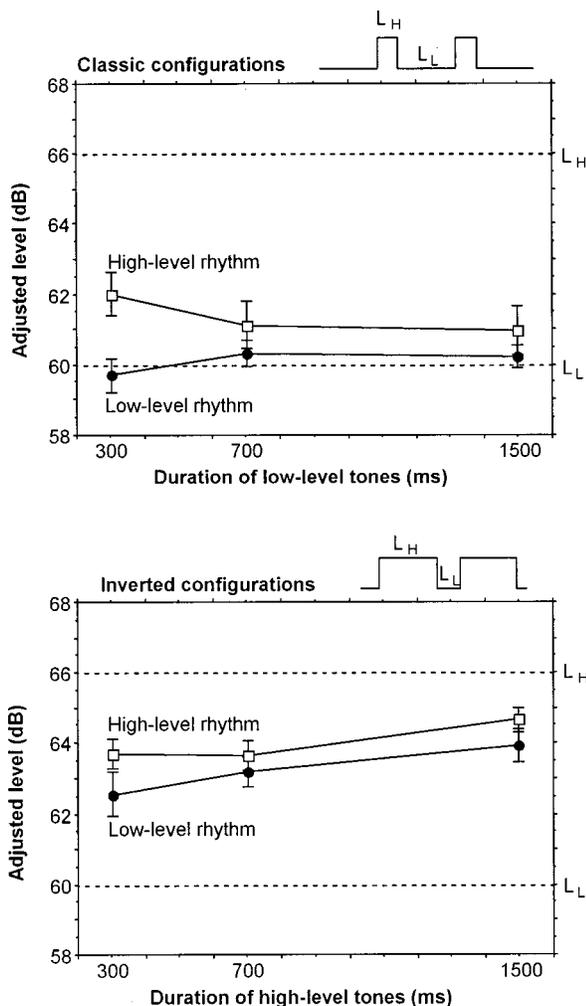


FIG. 5. Experiment 3: Mean adjusted levels for the intermittent comparison sequences for the 12 duration conditions averaged over groups, repetitions, and short-tone durations. Error bars indicate ± 1 standard error.

sion. No mention was made to the continuity phenomenon. In the classic configuration, 90% of the responses indicated the phenomenon, whereas only 45% indicated the phenomenon in the inverted configuration. The percentage reports of the phenomenon were more divergent across the two classes of stimuli for the expert group than for the novice group and were independent of the duration conditions.

2. Classic configuration

For the rhythm corresponding to the *low-level* (long-duration) tones, matches were extremely accurate (mean adjustment = 59.9 dB, mean adjustment error = 59.9 - 60.0 = 0.1 dB) (see Fig. 5). A single-sample *t*-test revealed that the mean adjustment error did not differ significantly from zero. A mixed ANOVA on the mean adjustment errors by group (novices/experts), duration of high-level tones (50 or 100 ms), duration of low-level tones (300, 700, or 1500 ms), and repetition (6) revealed no significant differences.

However, for the rhythm corresponding to the *high-level* (short-duration) tones, adjustment were far from accurate (mean adjustment = 61.4 dB; mean adjustment error = 61.4 - 66.0 = -4.6 dB), indicating that subjects systematically *underestimated* the level of these tones. A mixed ANOVA

on the mean adjustment errors by group (novices/experts), duration of high-level tones (50 or 100 ms), duration of low-level tones (300, 700, or 1500 ms), and repetition (6) revealed no significant differences between the two groups of subjects or over the six repetitions. The duration of the high-level tones did not influence the matches but the duration of the low-level tones did [$F(2,12) = 3.17$, $p < 0.04$]: Fig. 5 shows that the longer the low-level tone, the greater the adjustment error.

Thus, as for experiment 1, adjustment errors corresponded to the patterns expected if subjects were experiencing the continuity phenomenon. The phenomenon was greater when the low-level tones were longer.

3. Inverted configuration

For the rhythm corresponding to the *low-level* (short-duration) tones, systematic adjustment errors were observed (mean adjustment = 63.2 dB; mean adjustment error = 63.2 - 60.0 = +3.2 dB), indicating that subject systematically *overestimated* their level. A single-sample *t*-test revealed that the mean adjustment error differed significantly from zero [$t(7) = 8.4$, $p < 0.001$]. A mixed ANOVA on the mean adjustment errors by group (novices/experts), duration of high-level tones (300, 700, or 1500 ms), duration of low-level tones (50 or 100 ms), and repetition (6) revealed no significant differences, although the same tendency described above was observed here: the longer the high-level tone, the higher the adjustment of the low-level tones.

For the rhythm corresponding to the *high-level* (long-duration) tones, systematic adjustments were also observed (mean adjustment = 66 - 2 = 64 dB, mean adjustment error = -2.0 dB), indicating that subjects systematically *underestimated* their level. A single-sample *t*-test revealed that the mean adjustment error differed significantly from zero [$t(7) = 9.2$, $p < 0.001$]. An ANOVA on the mean adjustment errors by group (novices/experts), duration of low-level tones (50 or 100 ms), duration of high-level tones (300, 700, or 1500 ms), and repetition (6) revealed no significant differences between the two groups of subjects or over the six repetitions. The duration of the low-level tones did not influence the adjustments but the duration of the high-level tones did [$F(2,12) = 5.35$, $p < 0.02$]: the longer the high-level tone, the higher the adjustment of those tones.

Thus, as for experiment 1, in these inverted conditions, adjustment errors did not correspond to the patterns expected if subjects were experiencing the continuity phenomenon. Instead, the matches suggest that subjects were adjusting to a duration-weighted mean level: the adjusted levels increased as the relative duration of the high-level tones increased.

4. Comparison of experiments 1 and 3

The results obtained here are almost identical to those obtained in experiment 1. To test for an order effect reflecting the possibility that instructions emphasizing the phenomenon may influence its perception in later conditions, the results of experiments 1 and 3 were compared in a mixed ANOVA on the mean adjustment errors by group (novices/experts), duration of high-level tones (50 or 100 ms), dura-

tion of low-level tones (300, 700, or 1500 ms), repetition (6), and order (experiment 1/3; 3/1). Particularly worthy of note here is that the order of presentation of the two experiments and all of its interactions are not statistically significant, ruling out the possibility that the similarity in the data is due to contamination from the instructions orienting listeners toward the continuity phenomenon in experiment 1. The similarity of results from experiments 1 and 3 also occurs despite a different matching stimulus for the low-level tone, a continuous tone in experiment 1, and an intermittent tone in experiment 3. This suggests that the overestimation of the fainter sound's loudness was not a consequence of the mismatch between standard and comparison stimuli (continuous/intermittent or intermittent/continuous). Thus, the phenomenon is surprisingly robust because it occurs even when instructions should not favor perception of the phenomenon and should indeed create a bias toward perceptually decomposing the sequence.

IV. GENERAL DISCUSSION

According to the Auditory Scene Analysis framework proposed by Bregman (1990), the continuity phenomenon results from a primitive organizational process by which the auditory system is predisposed to expect existing sounds to continue throughout new events if no evidence exists to indicate that they have stopped. In this way, listeners organize the sequence into two streams. In the present case, the continuous tone is one stream (the old information) and the series of intermittent tones another (the new information). We have provided evidence of the presence of this phenomenon by asking listeners to adjust the perceived level of each stream separately. When the sequence is organized into two streams, the continuity phenomenon is perceived: the level of the low-level tones is correctly adjusted but there is a systematic underestimation of the level of the high-level tones. This underestimation suggests that the perceived level of the intermittent part (the new information) is derived from the high level by subtracting out the part assigned to the old stream.

Evidence in favor of the primitive nature of this phenomenon has been provided in three ways. First, the strength of the phenomenon is unaffected by the task: extremely similar underestimations of the level of the intermittent tones were obtained in both experiment 1 when the instructions emphasized the phenomenon and in experiment 3 when the instructions emphasized a different organizational principle, namely the interleaved rhythmic structure. The phenomenon occurs even when the listeners were asked to think about the stimuli in a different way. Second, it is thought that musical expertise often results in an enhanced ability to listen selectively to particular aspects of a sequence, and so we could have expected that musicians would be more resistant to the phenomenon: their adjustments might have been more veridical. However, this was not the case: no differences were observed between the musician and nonmusician listeners. Finally, repeated listening to the stimuli in our sequences could have led to more veridical adjustments. This was not the case either, as no improvements were observed over rep-

etitions. The phenomenon therefore appears to be particularly robust.

The main aim of this series of experiments was to establish whether the temporal structure of the old and new information influences the continuity phenomenon. Previous studies have tended to use only conditions in which the time period occupied by the old information alone is equal to or longer than that occupied by the combination of the old and new information. Is this just a random sampling error, or have researchers implicitly used the duration conditions under which the phenomenon is most clearly demonstrated? The latter is probably the case because we have demonstrated that the phenomenon only occurs under particular duration conditions: the old sounds must be relatively well established in relation to the new sounds. In experiments 1 and 3 we have shown that the phenomenon only occurs when the low-level tones (the old information) are almost equal to or longer than the high-level tones. In both these experiments we used extremely long (300, 700, and 1500 ms) and short (50 and 100 ms) tones. In experiment 2, we used the whole range of durations in order to examine the importance of the duty-cycle duration and duration ratios between the old and new information. Whereas the duty-cycle duration did not affect the degree of under- and overestimation, the relative duration is revealed to be the determining factor of whether or not the phenomenon is observed. More precisely, the high/low duration ratio must be at most 2:1. Warren (1994) states that continuity is reported when intermittent tones do not exceed 300 ms. Our results are coherent with this statement (the phenomenon is not observed in the inverted conditions where the intermittent/loud tones were 700 and 1500 ms). Contrary to previous assumptions (Houtgast, 1972; Warren *et al.*, 1972), and in agreement with data provided by Warren *et al.* (1994), the phenomenon is not all or nothing: its strength varies with the duration conditions, the degree of underestimation becoming greater as the high/low ratio decreases. This suggests that the more the old information dominates the new information, the stronger the tendency to perceptually organize the sequence according to this strategy.

Another issue concerns how well listeners are able to adjust loudness under various duration conditions, either when the continuity phenomenon is perceived or not. We have shown previously (McAdams *et al.*, 1998) that traditional loudness models positing precise loudness perception predict that tones should be adjusted close to the physical levels (66 dB for the long/high-level tones and 60 dB for the short/low-level tones). This is not what is observed for either the intermittent stream in the classic configurations nor for the inverted configurations. For the classic configurations, adjusted levels are those predicted by a pressure subtraction model (McAdams *et al.*, 1998) if the long/low-level tone duration is 1500 ms, but this predicted level is progressively overestimated as the duration of the tones decreases. In the inverted configurations, the intermittent/high-level tones were systematically underestimated (64.7 and 64.0 dB in experiments 1 and 3, respectively) and the continuous/low-level tones were systematically overestimated (62.9 and 63.2 dB, respectively).

Various explanations can be proposed for these systematic adjustment errors. A first possibility may be that listeners adjust towards a *mean level* (63.0 dB). Averaged over the different duration conditions and experiments, these predictions are quite close to the data for the inverted configuration (long/high-level=63.6 dB, short/low-level=63.0 dB). However, this explains neither the effect of duration condition (adjusted levels were higher for greater long/short ratios), nor the different adjustments for the high- and low-level tones. Further, the predictions are nowhere near the data for the classic configuration.

Second, a *duration-weighted mean*, taking into consideration the relative long-term contributions of the long- and short-duration levels (computed over the duration of the stimulus sequence), could explain the duration condition effect because the level of the longer tones would affect the mean more than the level of the shorter tones. However, for the inverted conditions the theoretical duration-weighted means (64.8, 65.4, and 65.7 dB for 300, 700, and 1500 ms, respectively) were considerably higher than the adjusted levels, whereas for the classic conditions, the duration-weighted means were in between the adjusted levels for the 300- and 700-ms long-tone durations (61.2 and 60.6 dB, respectively) and above the adjusted levels for the 1500-ms duration (60.3 dB). In all of these cases, calculations based on acoustic pressure units are nearly identical to results obtained when loudness units (sones) are used and then reconverted to dB. While the duration-weighted mean varies with the duration of the long-duration tones, it cannot account for the difference between adjustments for the long- and short-duration tones in either of the two stimulus configurations, since there are no free parameters in the equation related to focusing on a given subset of tones in the stimulus sequence.

A third possibility is that the relative weights accorded to the long-duration and short-duration tones may be modified by *attentional focusing* on them. In developing a model for this effect, we assume that listeners focus on high- or low-level tones (or long- or short-duration tones, according to the task instructions) and that this focusing is not perfectly selective, along a sensory dimension that we assume for simplicity's sake to be isomorphic to level, thus allowing influence from the other type of tone due to both tones stimulating largely the same population of auditory fibers. We hypothesize that the relative weighting of high- and low-level tones in their contribution to the perceived loudness would be different when attention is directed to one or the other (in an attempt to match their loudness) and that differences should also occur between classic and inverted configurations due to the different perceptual representations upon which the focusing operates. Our model for weighting the levels with respect to relative duration and attentional focus is given in Eq. (1)

$$L_{\text{adj}} = \alpha \times \frac{D_L}{D_L + D_H} \times L_L + \beta \times \frac{D_H}{D_L + D_H} \times L_H, \quad (1)$$

where L_{adj} is the adjusted level, D_L is the duration of the low-level tone, D_H is the duration of the high-level tone, L_L is the level (in dB) of the low-level tone, L_H is the level (in dB) of the high-level tone, and α and β are the weights for

low-level and high-level tones, respectively, for a given attentional focus. Since levels, durations, and mean adjusted levels (combined across 50-ms and 100-ms short-duration tones and using 75 ms as the duration) are known, we solve the simultaneous equations for the three pairs of long-duration tones (300, 700, 1500 ms) to determine α and β . For example, in the inverted configuration the three equations would be:

$$L_{\text{adj}}(300) = 0.2 \times 60\alpha + 0.8 \times 66\beta = 12.0\alpha + 52.8\beta$$

$$[D_L = 75 \text{ ms}, D_H = 300 \text{ ms}], \quad (2)$$

$$L_{\text{adj}}(700) = 0.097 \times 60\alpha + 0.903 \times 66\beta = 5.64\alpha + 59.60\beta$$

$$[D_L = 75 \text{ ms}, D_H = 700 \text{ ms}], \quad (3)$$

$$L_{\text{adj}}(1500) = 0.048 \times 60\alpha + 0.952 \times 66\beta = 2.88\alpha + 62.83\beta$$

$$[D_L = 75 \text{ ms}, D_H = 1500 \text{ ms}]. \quad (4)$$

Each equation can be rewritten in terms of α , e.g., for 300 and 700 ms

$$\alpha = \frac{L_{\text{adj}}(300) - 52.80\beta}{12.0}, \quad (5)$$

$$\alpha = \frac{L_{\text{adj}}(700) - 59.60\beta}{5.64}. \quad (6)$$

Assuming α equal in both cases, we solve for β

$$\beta = \frac{12.0 \times L_{\text{adj}}(700) - 5.64 \times L_{\text{adj}}(300)}{417.41}, \quad (7)$$

by inserting the mean adjusted levels for a given subject in Eq. (7). Doing this for the three pairs (300–700, 300–1500, 700–1500) gives three values of β which are averaged for each subject, since we assume that relative duration does not affect the attentional focus on high-level or low-level tones in either configuration. The mean β is reinserted into Eqs. (2)–(4) to obtain three values of α which are then averaged. The aim is to estimate a single set of coefficients for all temporal relations of a given stimulus configuration (classic, inverted) and focusing strategy (high-level tones, low-level tones). It is important to note that differences in the pattern of α and β across stimulus configurations and focusing strategies include differences due to the percept, i.e., no continuity in inverted and continuity (and thus the old-plus-new subtractive mechanism) in classic configurations. Plugging the mean coefficients averaged across subjects back into Eq. (1) gives the model predictions plotted in Fig. 2. A close correspondence exists between observed and predicted values when means are weighted both for duration and attentional focus. The mean values of α and β and their standard errors across subjects are plotted in Fig. 6 for the two stimulus configurations and two focusing conditions. If both coefficients have values close to 1, the duration-weighted mean accounts for the data. It is clear from this figure that both stimulus configuration and attentional focus affect the relative weights of the component tones in the sequence as they contribute to loudness.

These pairs of weighting coefficients vary in an interesting way across stimulus configurations and focusing strate-

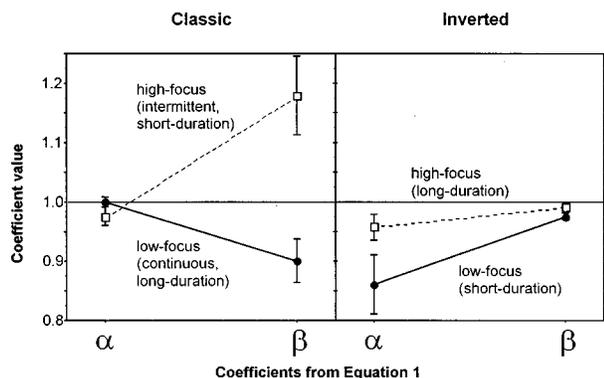


FIG. 6. Coefficients for the model that weight the levels according to relative duration and attentional focusing in loudness matches. Classic stimulus configuration in the left panel and inverted configuration in the right panel.

gies (Fig. 6). First, short- and long-duration tones do not have the same pattern of coefficients across stimulus configurations: focusing on low-level tones in the classic configuration is not the same as focusing on high-level tones in the inverted configuration and vice versa, suggesting that the differences between configurations are not due simply to a reversal of the high and low levels on the short and long durations. Second, high- and low-level tones do not have the same pattern across stimulus configurations: low-level classic is not the same as low-level inverted and high-level classic is not the same as high-level inverted, suggesting that the duration differences across configurations are critical to the perceived loudnesses. Both of these results are consistent with the differences in reports of the continuity phenomenon given by our subjects. Third, α and β are affected both by the stimulus configuration and by the focusing strategy. For the classic configuration, the weight on the low-level tones is determined entirely by the relative duration ($\alpha = 1$), whereas the weight on the high-level tones increases ($\beta > 1$) when they are the object of focus and decreases ($\beta < 1$) when they are not. A nearly inverse pattern exists for the inverted configuration: the weight on the high-level tones is largely determined by relative duration ($\beta \cong 1$), whereas the weight on the low-level tones is determined by relative duration ($\alpha \cong 1$) when they are not the object of focus and paradoxically decreases ($\alpha < 1$) when they are, indicating an overbearing influence of the high-level tones when trying to focus on the loudness of the short-duration, low-level tones in the inverted configuration.

In summary, the homophonic continuity phenomenon in the classic stimulus configuration is quite robust with respect to task instructions, listener expertise or training, and repeated exposure to the task. The phenomenon depends on the relative durations of high- and low-level tones and appears to require that the low-level portions temporally dominate the high-level portions, or at least that the ratio of their durations (high/low) be at most 2:1, in order to be perceived systematically. In stimulus conditions in which short-duration tones

alternate with long-duration tones, the effect of longer tones on loudness matches cannot be explained only by a simple mechanism involving the global level accumulated over an entire duty cycle, but also seems to involve a relative attentional weighting of the high- and low-level portions in inverted conditions and of the continuous and intermittent streams in the classic conditions. This attentional focusing is not perfectly selective and part of the unfocused portion appears to influence the perceived loudness. The nature of this relative weighting is clearly different for classic and inverted conditions: it involves an inversion of the high- and low-level weights suggesting quite different underlying sensory representations for the two stimulus configurations, consistent with subjective reports concerning the presence of the continuity phenomenon.

ACKNOWLEDGMENTS

This research was funded in part by a grant from the French Ministry of the Environment. Lucette Cysique and Sophie Savel helped run subjects and Renaud Brochard helped with DSP programming.

- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).
- Bregman, A. S., and Rousseau, L. (1991). "Auditory intensity changes can cue perception of transformation, accompaniment or replacement," *Bull. Psychonom. Soc.* **29**, 476.
- Ciocca, V., and Bregman, A. S. (1987). "Perceived continuity of gliding and steady-state tones through interrupting noise," *Percept. Psychophys.* **42**, 476–484.
- Drake, C., and McAdams, S. (1997). "The continuity illusion: Role of temporal sequence structure," *Actes du 4e Congrès Français d'Acoustique, Marseille, Vol. 1*, pp. 519–522.
- Houtgast, T. (1972). "Psychophysical evidence of lateral inhibition in hearing," *J. Acoust. Soc. Am.* **51**, 1885–1894.
- McAdams, S., Botte, M.-C., and Drake, C. (1994). "Le phénomène de la continuité auditive et la répartition de la sonie entre flux auditifs," *J. Phys. IV (C5)*, 383–386.
- McAdams, S., Botte, M.-C., and Drake, C. (1998). "Auditory continuity and loudness computation," *J. Acoust. Soc. Am.* **103**, 1580–1591.
- Thurlow, W. R. (1957). "An auditory figure-ground effect," *Am. J. Psychol.* **70**, 653–654.
- Thurlow, W. R., and Elfner, L. F. (1959). "Continuity effects with alternately sounding tones," *J. Acoust. Soc. Am.* **31**, 1337–1339.
- van Noorden, L. P. A. S. (1977). "Minimum differences of level and frequency for perceptual fission of tones sequences ABAB," *J. Acoust. Soc. Am.* **61**, 1041–1045.
- Verschuure, J., Rodenburg, M., and Maas, A. J. J. (1976). "Presentation conditions of the pulsation threshold method," *Acustica* **35**, 47–54.
- Warren, R. M. (1982). *Auditory Perception: A New Synthesis* (Pergamon, New York).
- Warren, R. M. (1984). "Perceptual restoration of obliterated sounds," *Psychol. Bull.* **96**, 371–383.
- Warren, R. M. (1999). *Auditory Perception: A New Analysis and Synthesis* (Cambridge University Press, Cambridge).
- Warren, R. M., Bashford, J. A., Healy, E. W., and Brubaker, B. S. (1994). "Auditory induction: Reciprocal changes in alternating sounds," *Percept. Psychophys.* **55**, 313–322.
- Warren, R. M., Obusek, C. J., and Ackroff, J. M. (1972). "Auditory induction: Perceptual synthesis of absent sounds," *Science* **176**, 1149–1151.
- Zwicker, E. (1977). "Procedure for calculating loudness of temporally variable sounds," *J. Acoust. Soc. Am.* **62**, 675–682.

Effects of frequency and duration on psychometric functions for detection of increments and decrements in sinusoids in noise

Brian C. J. Moore^{a)}

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

Robert W. Peters

Division of Speech and Hearing Sciences, Department of Medical Allied Health Professions, The School of Medicine, and Department of Psychology, The University of North Carolina at Chapel Hill, North Carolina 27599-7190

Brian R. Glasberg

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

(Received 11 March 1998; revised 29 July 1999; accepted 18 August 1999)

Psychometric functions for detecting increments or decrements in level of sinusoidal pedestals were measured for increment and decrement durations of 5, 10, 20, 50, 100, and 200 ms and for frequencies of 250, 1000, and 4000 Hz. The sinusoids were presented in background noise intended to mask spectral splatter. A three-interval, three-alternative procedure was used. The results indicated that, for increments, the detectability index d' was approximately proportional to $\Delta I/I$. For decrements, d' was approximately proportional to ΔL . The slopes of the psychometric functions increased (indicating better performance) with increasing frequency for both increments and decrements. For increments, the slopes increased with increasing increment duration up to 200 ms at 250 and 1000 Hz, but at 4000 Hz they increased only up to 50 ms. For decrements, the slopes increased for durations up to 50 ms, and then remained roughly constant, for all frequencies. For a center frequency of 250 Hz, the slopes of the psychometric functions for increment detection increased with duration more rapidly than predicted by a "multiple-looks" hypothesis, i.e., more rapidly than the square root of duration, for durations up to 50 ms. For center frequencies of 1000 and 4000 Hz, the slopes increased less rapidly than predicted by a multiple-looks hypothesis, for durations greater than about 20 ms. The slopes of the psychometric functions for decrement detection increased with decrement duration at a rate slightly greater than the square root of duration, for durations up to 50 ms, at all three frequencies. For greater durations, the increase in slope was less than proportional to the square root of duration. The results were analyzed using a model incorporating a simulated auditory filter, a compressive nonlinearity, a sliding temporal integrator, and a decision device based on a template mechanism. The model took into account the effects of both the external noise and an assumed internal noise. The model was able to account for the major features of the data for both increment and decrement detection. © 1999 Acoustical Society of America. [S0001-4966(99)02112-8]

PACS numbers: 43.66.Fe, 43.66.Mk [RVS]

INTRODUCTION

How intensity changes are represented in the auditory system continues to be a fundamental problem (Green, 1988; Plack and Carlyon, 1995). This paper is concerned with two aspects of this problem. The first is: what measure or measures can best represent the perceptual response to the stimulus in intensity discrimination experiments? Grantham and Yost (1982), Green (1988, 1993), and Plack and Carlyon (1995) have presented detailed discussions of the various measures, including the change in intensity divided by the intensity of the standard, $\Delta I/I$; that ratio expressed in decibels, $10 \log_{10}(\Delta I/I)$; the ratio of the higher intensity to the lower intensity, $(I + \Delta I)/I$; that ratio expressed as the change in level in decibels, $\Delta L = 10 \log_{10}((I + \Delta I)/I)$; the change in

sound pressure divided by the sound pressure of the standard, $\Delta p/p$; and the decibel equivalent, $20 \log_{10}(\Delta p/p)$. It has been proposed that a useful measure is one to which the detectability, d' , of a change in intensity is linearly related (Buus and Florentine, 1991; Green, 1993).

Psychometric functions have often been fitted with equations of the form $d' = mX^k$, where X is ΔL , $\Delta I/I$, or $\Delta p/p$ (Egan *et al.*, 1969). The exponent k is often referred to as defining the slope of the psychometric function, but in this paper we refer to it as describing the *shape* of the function. A number of investigators (Rabinowitz *et al.*, 1976; Florentine, 1986; Viemeister, 1988; Buus and Florentine, 1991; Buus *et al.*, 1995; Dai *et al.*, 1999) have concluded that d' for the level discrimination of pulsed tones is approximately proportional to ΔL , i.e., when $X = \Delta L$, $k \approx 1$. Green *et al.* (1979) concluded that, when $X = \Delta I$, k was between 0.5 and 0.7, but

^{a)}Electronic mail: bcjm@cus.cam.ac.uk

this is quite similar to proportionality to ΔL for values of ΔL up to about 4 dB. For the detection of increments imposed on a continuous pedestal, Jeffress (1975) and Green *et al.* (1979) reported that, when $X = \Delta I$, k was greater than one. However, neither of these two papers actually presented the raw data, so it is not possible to determine if some other function would have fitted better. Dai *et al.* (1999) reported that, for $X = \Delta L$, k was greater than one, but decreased with increasing increment duration, and was close to one for durations of 100 and 300 ms.

Psychometric functions are usually measured with stimulus values giving percent correct values between about 55% and 95% correct. Buus and Florentine (1991) noted that many studies used stimuli which gave high sensitivity, i.e., relatively small changes in level resulted in good performance. When changes in level are small it is difficult to determine which of the measures ΔL , $\Delta I/I$, and $\Delta p/p$ gives the most linear relationship to d' (Buus *et al.*, 1995).

We wanted to extend earlier results by measuring psychometric functions for the detection of increments and decrements in level imposed on a longer-duration pedestal. Stimuli of this type have been used in several previous studies (Bacon and Viemeister, 1985; Bacon and Moore, 1986; de Boer, 1986; Forrest and Green, 1987; Plack and Moore, 1991; Moore *et al.*, 1993; Peters *et al.*, 1995; Moore *et al.*, 1996; Moore and Peters, 1997; Oxenham, 1997), but we are not aware of any previous work measuring psychometric functions for such stimuli. We took several steps in order to ensure that we could measure psychometric functions over a range of increment and decrement magnitudes sufficient to distinguish the various measures. First, we included stimuli with very short durations. Second, we measured psychometric functions over a wide range, including increments and decrements giving scores in the range 90%–100%. Third, we added masking noise to make the task more difficult, so as to achieve bigger increment and decrement magnitudes. This noise also served the purpose of masking spectral splatter associated with the rapid onsets and offsets of the increments and decrements.

One problem with using noise is that it may affect the form of the psychometric function. Green (1960) has proposed that large values of k may occur because the listener is uncertain about the parameters of the signal when it is heard in noise; his comments were made in relation to the detection of a bandpass noise signal presented in a broadband background, but they might also apply when noise is added to a sinusoidal pedestal. Buus *et al.* (1995), however, reported that the proportional relationship between d' and ΔL for the intensity discrimination of pulsed tones is not changed by masking noise.

The second issue addressed in this paper is the effect of duration on the ability to detect a change in intensity in a stimulus. It has usually been found that intensity discrimination improves with increasing duration (Henning, 1970; Jeffress, 1975; Green *et al.*, 1979; Florentine, 1986; Eddins and Green, 1995; Moore and Peters, 1997). This could be interpreted as reflecting some form of integration in the auditory system, such as energy integration. However, another possibility is that performance improves with increasing duration

because longer stimuli provide more detection opportunities through repeated sampling of the stimuli. This idea is sometimes called “multiple looks” (Viemeister and Wakefield, 1991). If the improvement in detection of intensity changes with increasing duration does reflect multiple looks, and if the information from each “look” is weighted equally and combined optimally, then the detectability, d' , of a change in intensity should be linearly related to the square root of the number of “looks.” Assuming that the “looks” involve sampling over short durations, this implies that d' should be linearly related to the square root of duration. If d' changes with duration more slowly than this, the implication is either that the information from individual looks is not combined optimally, or that it is not weighted equally. For example, in an increment-detection task, the information from the onsets of the increments might be much more salient than information from the ongoing parts of the increments.

Our interest in this issue was partly kindled by an earlier study (Moore and Peters, 1997) in which we found that thresholds for increment and decrement detection in a gated sinusoidal pedestal presented in a moderate level of continuous background noise changed relatively little with increment or decrement duration especially for high-frequency (4000 Hz) pedestals. Oxenham (1997) has reported a similar finding. The effects of duration under these conditions appeared to be smaller than the effects found using a continuous pedestal (Leshowitz and Raab, 1967; Green *et al.*, 1979; Dai *et al.*, 1999); see later for a review (Table V). A possible interpretation of the small effects of duration is in terms of the “change detection” hypothesis proposed by Macmillan (1973). He suggested that the auditory system contains a change detector which is insensitive to the direction of change (i.e., whether it is an increment or a decrement) and whose sensitivity does not change with signal duration. Evidence supporting this idea has been provided by Carlyon and Moore (1986) and Hafter *et al.* (1996).

One problem with our earlier study (Moore and Peters, 1997), and the studies of Macmillan (1971, 1973) and Oxenham (1997), is that they measured *thresholds* (as opposed to psychometric functions) for increment and decrement detection as a function of duration. The small decreases in threshold with increasing duration were taken as evidence for a change detector. However, this evidence is not clear cut, because the change in threshold that would be expected from a temporal integrator or from a multiple-looks theory depends upon the shape of the psychometric function for increment or decrement detection; see Carlyon *et al.* (1990) and Donaldson *et al.* (1997). Consider as an example the effect predicted by a multiple-looks theory of multiplying the duration of an increment by a factor R . If the size of the increment remained fixed, the increase in duration should result in an increase in d' by a factor of \sqrt{R} . The data of Jeffress (1975) and of Green *et al.* (1979) suggest that d' for increment detection is proportional to $(\Delta I/I)^k$, where k is about 2.2 for brief increments. Therefore, to restore detectability of the longer increment to that of the shorter one, $\Delta I/I$ should be decreased by a factor $(\sqrt{R})^{1/k}$. For example, if $R=2$, then $\Delta I/I$ would have to be decreased by a factor of only 1.17 (equivalent to 0.7 dB). Thus, a large value of k could lead to performance

that looked like it was based on a change detector, when in fact it was not.

In our experiment, we measured psychometric functions for increment and decrement detection over a wide range of durations and for a wide range of pedestal frequencies, to gain further insight into the nature of temporal integration in increment and decrement detection and to provide a stronger test of the change-detection hypothesis.

I. METHOD

A. Procedure

Psychometric functions for detecting increments and decrements in level of sinusoidal pedestals were measured using a modified method of constant stimuli and a three-alternative forced-choice task. Data were gathered first for increments and later for decrements. The method was based on that described by Moore and Sek (1992). Five increment or decrement levels, expressed as ΔL , were employed, where ΔL was equal to the difference in level of the pedestal alone and the pedestal plus the increment or decrement; $\Delta L = 10 \log_{10} ((I + \Delta I)/I)$. The highest value of ΔL was chosen to be highly detectable. The other values were chosen to be progressively less detectable. All five values were used within a block of 55 trials. The first five trials were practice trials, and used the highest ΔL value. Subsequently, the five values were presented in a regular sequence going from highest to lowest values. Thus, the highest value was presented every five trials to help the subjects "remember" the detection cues (Taylor and Forbes, 1983). We have found that this procedure leads to better detectability for small values of ΔL than when stimuli are presented in blocks with a constant value of ΔL within each block. The five increment or decrement values were chosen on the basis of pilot data to give values of d' between about 0.56 and 3.25 (values of % correct between about 50 and 98).

Each trial consisted of three observation intervals, marked by lights, and with feedback given by lights on the response box. At the end of each block of trials the number of correct identifications for each value of ΔL was calculated. Data from ten blocks of trials were averaged, to determine the proportion of correct responses for each psychometric function for each subject. Thus, each point on each psychometric function was based on 100 judgments. The subjects were tested individually in a sound attenuating chamber.

B. Stimuli

The sinusoidal pedestal, with frequency either 250, 1000, or 4000 Hz, was generated by a Wavetek 182 function generator, and gated in three bursts each with 50-ms rise/fall time (Wilsonics BSIT, set to cosine squared, providing a raised-cosine envelope). The pedestal duration varied with increment or decrement duration (which was 5–200 ms), so as to maintain constant steady-state portions before and after the increment or decrement; the steady-state duration before the increment/decrement was 200 ms and the steady-state duration after the increment/decrement was 100 ms. The interstimulus interval was 200 ms. The level of the pedestal

was 70 dB SPL, specified in terms of eardrum sound pressure as measured by a probe microphone placed near the eardrum.

Another Wilsonics gate provided a burst of a sinusoid obtained from the same function generator. The burst was amplified by 10 dB, passed through a Wilsonics programmable attenuator (PATT) and added to one of the longer (pedestal) bursts via an active adder, either in phase to produce an increment in level or in antiphase to produce a decrement. The attenuator was used to control the magnitude of the increment or decrement in level, ΔL . The increment or decrement was presented in one of the three bursts of the pedestal, chosen at random on each trial.

The brief burst, which was gated without regard to phase of the sinusoid (although it was always either in phase or out of phase with the pedestal), had a raised-cosine rise/fall time of 5 ms. Psychometric functions were determined for steady-state increment/decrement durations of 5, 10, 20, 50, 100, and 200 ms. The 5-ms rise/fall time, while allowing for increments/decrements of short duration, also helped in reducing the detectability of spectral splatter. As another means to prevent the detection of splatter, the sinusoid was presented in continuous white noise (Grason-Stadler, 901B) low-pass filtered at 750, 2000, and 8000 Hz (Khron-Hite 3550, 48 dB/oct slope) for the pedestal frequencies of 250, 1000, and 4000 Hz, respectively. The spectrum level of the noise was 30.5, 34.6, and 40 dB below the level of the pedestal for the frequencies 250, 1000, and 4000, respectively. The overall intensity of each noise was roughly the same. The overall noise level at the output of the auditory filters was approximately 13.4 dB below the pedestal level for each center frequency. This is based on the assumption that the auditory filter equivalent rectangular bandwidths were 51.7, 133, and 457 Hz (Glasberg and Moore, 1990) at 250, 1000, and 4000 Hz, respectively. The subjects reported that no clicks or thuds were audible and that the increments and decrements were heard as short fluctuations in level of the pedestal.

The signal and noise were combined in an active adder and routed through a manual attenuator to one earpiece of a Sennheiser HD424 headset, selected for its relatively smooth frequency response. Stimulus timing and levels were controlled by an IBM PC-compatible computer that also recorded subjects' responses and provided feedback.

C. Subjects

The seven subjects all had absolute thresholds below 20 dB HL at all frequencies tested. Six of the subjects were university students, all less than 26 years old; the remaining subject was author RP, who was 75 years old. All subjects had considerable experience in comparable psychoacoustical listening tasks and were given several practice trials to familiarize them with the task. No significant practice effects were noted during the course of the experiment. Subject RP was tested in all conditions. Subjects CA and MD were tested in all conditions with increments. Subject WH was tested in all conditions with decrements. Subjects GS, SM,

TABLE I. Statistics of the intercepts with the abscissa of the best-fitting straight lines for d' vs ΔL , d' vs $\Delta I/I$, and d' vs $\Delta p/p$. Results are given for increment detection [part (a)] and for decrement detection [part (b)]. To assess whether the intercepts were significantly different from zero, t -tests were used. NS indicates not significant at the 0.05 level.

	ΔL	$\Delta I/I$	$\Delta p/p$
(a) Increment detection			
Number of positive intercepts (out of 12)	12	7	11
Mean intercept	1.85	-0.58	0.12
Standard deviation of intercepts	1.15	1.64	0.14
Significance of difference from zero	<0.001	NS	<0.02
(b) Decrement detection			
Number of positive intercepts (out of 12)	6	11	11
Mean intercept	0.71	0.30	0.12
Standard deviation of intercept	4.3	0.25	0.12
Significance of difference from zero	NS	<0.02	NS

and HL were tested with decrements for frequencies of 250, 1000, and 4000 Hz, respectively. Thus, three subjects were tested for each condition.

II. RESULTS

A. Determining the stimulus measure linearly related to d' for increments

The percent correct for each value of ΔL was converted to the detectability index d' (Macmillan and Creelman, 1991). We wished to find a quantity describing the magnitude of the increment such that d' was directly proportional to this quantity, for a given increment frequency and duration. The following measures of the increment magnitude were evaluated: ΔL , $\Delta I/I$, and $\Delta p/p$ (Green, 1993). We examined those cases where the values of ΔL covered a large range, namely for the two shortest increments at the two lower frequencies (4 conditions \times 3 subjects = 12 cases). Initially, straight lines were fitted to the psychometric functions with the constraint that the lines passed through the origin. The sum of the squared deviations of the data from the fitted lines had an average value of 0.22 for d' versus ΔL , 0.13 for $\Delta I/I$, and 0.12 for $\Delta p/p$. Hence the goodness of fit was similar for $\Delta I/I$ and $\Delta p/p$, while ΔL gave a somewhat poorer fit. In a further analysis, straight lines were fitted to the psychometric functions without the constraint that the lines passed through the origin. If the intercepts with the abscissa are positive, this implies that the psychometric functions have a slight upwards curvature, while if they are negative, the functions have a downward curvature. The results are summarized in part (a) of Table I. The only measure for which there is no evidence for curvature in the psychometric functions is $\Delta I/I$.

As a further way of determining the best fitting measure, the data were fitted with functions of the form $d' = m(X)^k$, where X is ΔL , $\Delta I/I$ or $\Delta p/p$. If d' is linearly related to a given X , then the best fitting value of k should be close to one. The results are summarized in part (a) of Table II. They indicate that d' for increments is approximately linearly related to $\Delta I/I$, whereas it is not linearly related to ΔL or to $\Delta p/p$; the values of k for the latter two quantities are significantly different from one.

TABLE II. Values of k derived from fitting the psychometric functions with the equation $d' = m(X)^k$, where X is ΔL , $\Delta I/I$, or $\Delta p/p$. Results are given for increment detection [part (a)] and for decrement detection [part (b)]. To assess whether the intercepts were significantly different from 1, t -tests were used. NS indicates not significant at the 0.05 level.

	ΔL	$\Delta I/I$	$\Delta p/p$
(a) Increment detection			
Mean value of k	1.78	1.05	1.34
Standard deviation of k	0.50	0.33	0.39
Significance of difference from 1	<0.001	NS	<0.01
(b) Decrement detection			
Mean value of k	1.18	2.54	1.69
Standard deviation of k	0.39	0.91	0.47
Significance of difference from 1	NS	<0.001	<0.001

B. Determining the stimulus measure linearly related to d' for decrements

An analysis similar to that described above was carried out for the data on decrement detection. For straight lines fitted to the psychometric functions with the constraint that the lines passed through the origin, and for the two shortest decrements at the two lower frequencies, the sum of the squared deviations of the data from the fitted lines had an average value of 0.10 for d' versus ΔL , 0.22 for $\Delta I/I$, and 0.16 for $\Delta p/p$. Hence the goodness of fit was best for ΔL . The results of fitting straight lines to the psychometric functions without the constraint that the lines passed through the origin are shown in part (b) of Table I. The intercept differs significantly from zero for $\Delta I/I$, but not for ΔL or $\Delta p/p$. The data were also fitted with equations of the form $d' = m(X)^k$. The results are summarized in part (b) of Table II. They indicate that d' for decrements is approximately linearly related to ΔL , whereas it is not linearly related to $\Delta I/I$ or to $\Delta p/p$.

C. Psychometric functions for increment detection

The psychometric functions for increment detection are plotted as d' versus $\Delta I/I$ in Figs. 1–3; $\Delta I/I$ was chosen as d' was linearly related to $\Delta I/I$. Each figure shows the individual results for the three subjects at one frequency. Note that the ordinate range differs across figures. Each symbol represents one increment duration. The best-fitting straight lines, forced through the origin, are also shown. We forced the lines to pass through the origin, since d' must be zero when $\Delta I/I$ is zero. The slopes of these lines are given in Table III. Note that the word “slopes” here refers to the slopes of the linear functions relating d' to $\Delta I/I$. It does not refer to the value of k . The results for subject RP did not differ clearly from those of the other subjects, even though he was much older than the other subjects. The bottom right panel in each figure shows lines with slopes corresponding to the average across subjects. The variability of the slopes across subjects tended to increase with increasing slope, but was roughly a constant proportion of the mean slope. Hence mean slopes were calculated as geometric means. The slopes increased with increasing frequency. For the frequency of 250 Hz, the mean slopes increased monotonically with increasing increment duration. All three subjects showed a

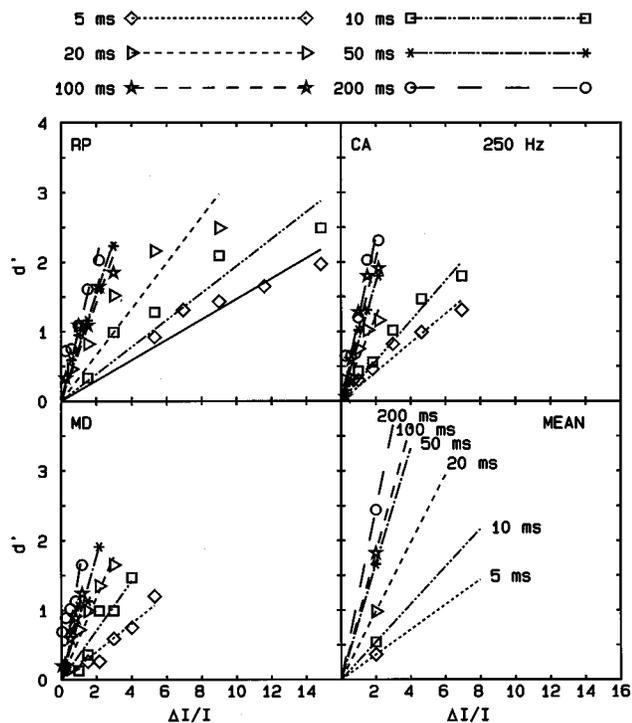


FIG. 1. Psychometric functions relating d' to $\Delta I/I$ for a frequency of 250 Hz. Each symbol represents one increment duration. The best-fitting straight lines, forced through the origin, are also shown. The bottom right panel shows lines with slopes corresponding to the geometric mean across subjects.

clear trend for slopes to increase with increasing duration. A similar trend was observed for the frequency of 1000 Hz, although the change in slope with increment duration was somewhat smaller. For the frequency of 4000 Hz, changes in slope with duration were even smaller, especially for subject

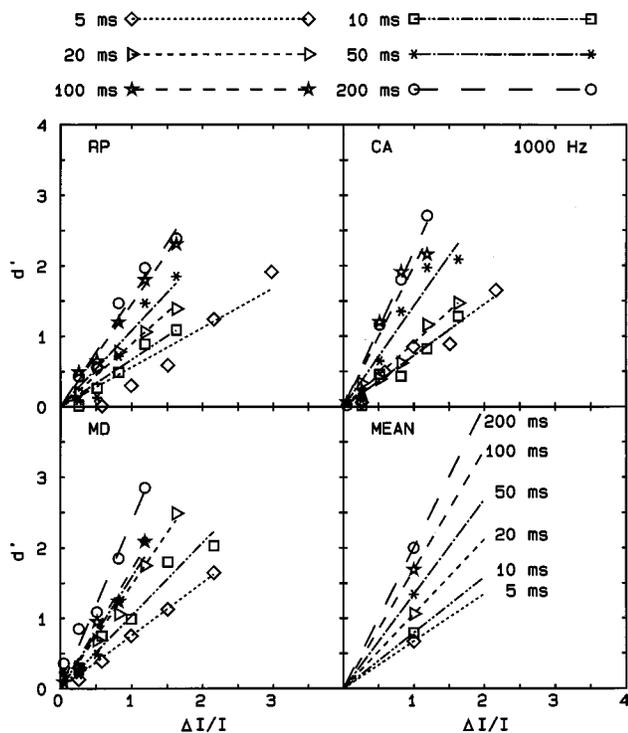


FIG. 2. As Fig. 1, but for a frequency of 1000 Hz.

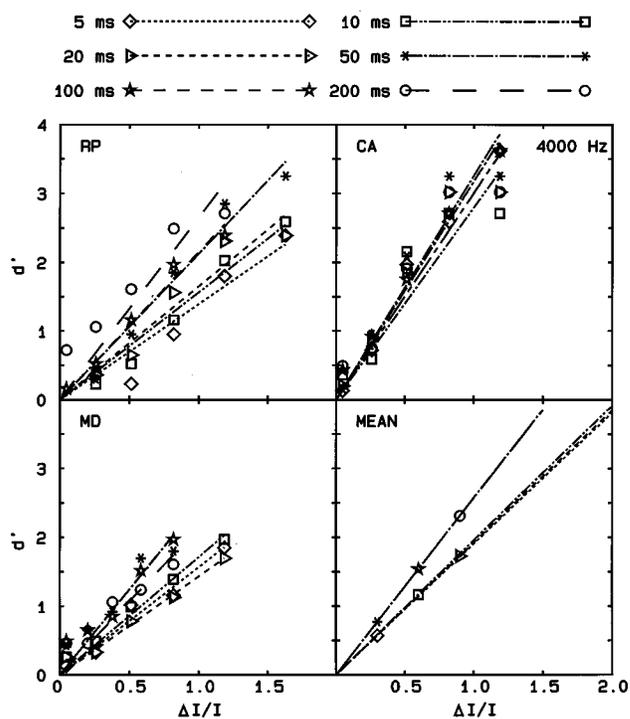


FIG. 3. As Fig. 1, but for a frequency of 4000 Hz.

CA. The mean slope at 4000 Hz increased with increasing increment duration up to 50 ms and then remained constant.

A within-subjects analysis of variance (ANOVA) was conducted on the logarithms of the slope values (as the variability of the slopes across subjects was roughly a constant proportion of the mean slope), with factors frequency and duration. Both main factors were highly significant: for frequency, $F(2,4)=52.8$, $p=0.001$; for duration, $F(5,10)=149.9$, $p<0.001$. The main effect of frequency reflects the fact that performance generally improved (the slopes became steeper) with increasing frequency. The interaction of frequency and duration was also highly significant: $F(10,20)=14.41$, $p<0.001$. This reflects the fact that the slopes changed less with duration as the frequency increased.

In summary, the slopes of the psychometric functions for increment detection increased with increasing frequency and with increasing duration. However, the effect of duration decreased with increasing frequency and was very small at 4000 Hz.

D. Psychometric functions for decrement detection

The psychometric functions for decrement detection are plotted as d' vs ΔL in Figs. 4–6; ΔL was chosen as d' was linearly related to ΔL . Note that the ordinate range differs across figures. The slopes of the best-fitting straight lines are given in Table IV. The results for subject RP did not differ clearly from those of the other subjects, even though he was much older than the other subjects. The bottom right panel in each figure shows lines with slopes corresponding to the geometric mean across subjects. Generally, the slopes of the functions increased (indicating better performance) with increasing frequency. For all three frequencies, performance

TABLE III. Slopes of the best-fitting straight lines to the data relating d' to $\Delta I/I$ for increment detection. The lines were constrained to pass through the origin. Slopes are given for each subject for each steady-state duration and frequency. Geometric mean slopes are also given.

Duration (ms)	Frequency (Hz)											
	RP			CA			MD			Mean		
	250	1000	4000	250	1000	4000	250	1000	4000	250	1000	4000
5	0.15	0.56	1.39	0.21	0.73	3.16	0.20	0.75	1.56	0.18	0.67	1.90
10	0.19	0.67	1.56	0.29	0.72	2.78	0.35	1.03	1.71	0.27	0.79	1.95
20	0.33	0.89	1.65	0.61	0.90	3.00	0.59	1.47	1.43	0.49	1.06	1.92
50	0.77	1.07	2.12	0.85	1.42	3.25	0.88	1.58	2.47	0.83	1.34	2.57
100	0.70	1.45	2.15	1.02	1.99	3.18	1.06	1.68	2.48	0.91	1.69	2.57
200	1.02	1.55	2.65	1.17	2.20	3.01	1.52	2.36	2.13	1.22	2.00	2.57

tended to improve progressively with increasing decrement duration up to 50 ms, and then to remain roughly constant.

As different subjects were used for each frequency, separate within-subjects ANOVAs were conducted for each frequency on the logarithms of the slope values, with factor duration. The effect of duration was highly significant for all three frequencies; $F(5,10) > 10.5$, $p < 0.001$. At 250 Hz, *post hoc* comparisons (based on the least-significant differences test) for each possible pair of durations showed that the slopes were significantly different ($p < 0.05$) for all pairs of decrement durations up to 50 ms. Slopes were not significantly different for durations of 50, 100, and 200 ms. Similar patterns of results were found at 1000 and 4000 Hz, except that at 1000 Hz slopes did not differ significantly for durations of 5 and 10 ms, and at 4000 Hz slopes did not differ significantly for durations of 5 and 10 ms and 20 and 50 ms.

Overall, these results indicate that performance at all three frequencies improves with increasing decrement duration up to about 50 ms, and then reaches an asymptote. Thus, temporal integration, whether mediated by multiple looks or by some other mechanism, does not extend over durations longer than 50 ms.

E. Changes in d' with increment duration

For a given value of $\Delta I/I$, the value of d' for increments is directly proportional to the slope of the psychometric function. Therefore, the changes in slope with increment duration can be used to assess how detectability changes with duration. If the improvement in detection of intensity changes with increasing duration reflects multiple looks, then d' should be linearly related to the square root of duration. To assess whether this was the case, we used the geometric mean slopes of the psychometric functions across subjects for each frequency and increment duration (see Table III). The mean slope for a given duration was divided by the square root of that duration (expressed as the equivalent rect-

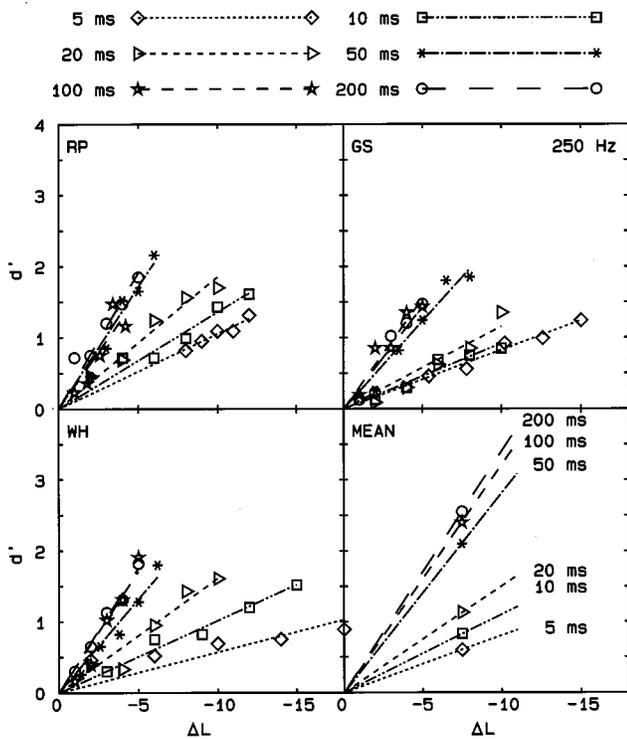


FIG. 4. Psychometric functions relating d' to ΔL for a frequency of 250 Hz. Each symbol represents one decrement duration. The best-fitting straight lines, forced through the origin, are also shown. The bottom right panel shows lines with slopes corresponding to the geometric mean across subjects.

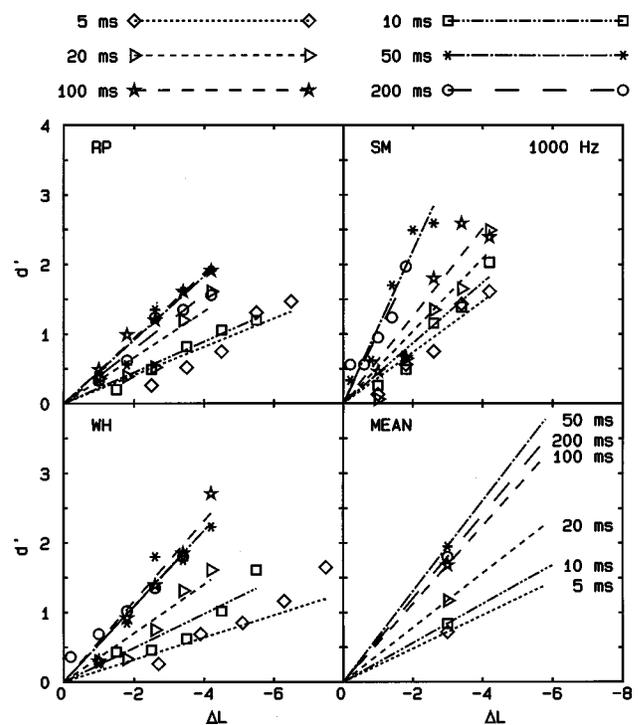


FIG. 5. As Fig. 4, but for a frequency of 1000 Hz.

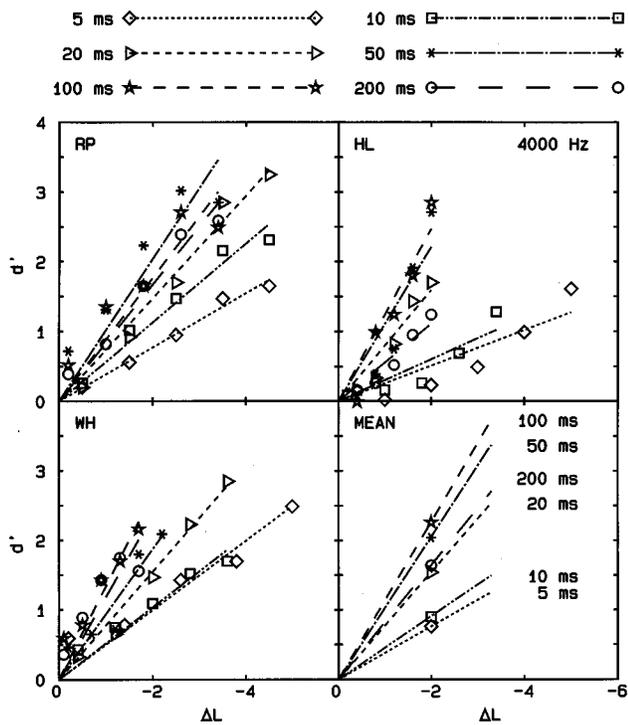


FIG. 6. As Fig. 4, but for a frequency of 4000 Hz.

angular duration, ERD). If the prediction from the multiple-looks hypothesis holds, then for a given frequency the values obtained in this way should be independent of duration. The results of these calculations are shown in the upper panel of Fig. 7. For the frequency of 250 Hz, the ratio slope/ $\sqrt{\text{ERD}}$ actually increased as the ERD increased up to 53.75 ms, and then decreased slightly for longer ERDs. The initial increase indicates that performance increased with increasing duration *more* than predicted by a multiple-looks hypothesis. For the frequency of 1000 Hz, the ratio was roughly constant for ERDs up to 23.75 ms, and then declined slightly. For the frequency of 4000 Hz, the ratio decreased monotonically with increasing ERD.

These results indicate that at 4000 Hz the increase in d' with increasing duration is less than predicted by the multiple-looks hypothesis. This implies either that the information from individual looks is not combined optimally, or that it is not weighted equally. For the 4000-Hz frequency, it appears that information from the onsets of the increments is

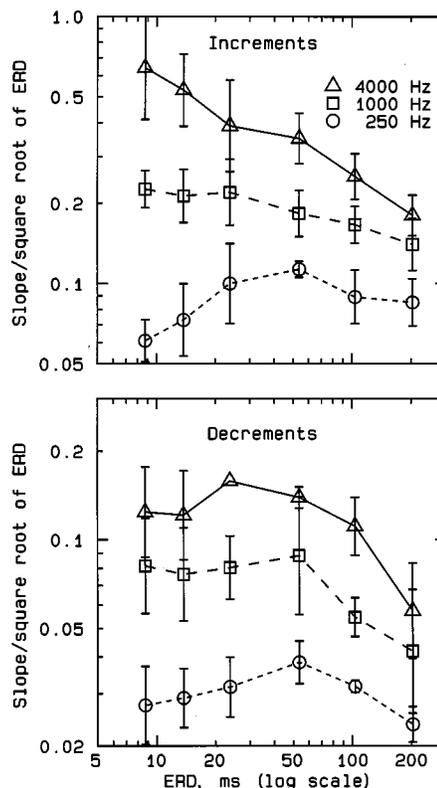


FIG. 7. The upper panel shows geometric means of the slopes of the psychometric function for increment detection (d' vs $\Delta I/I$) for a given equivalent rectangular duration (ERD), divided by the square root of the ERD and plotted as a function of the ERD. Each symbol represents one frequency. Error bars indicate \pm one standard deviation across subjects. A multiple-looks hypothesis leads to the prediction that the values obtained in this way should be independent of duration. The lower panel shows mean absolute values of the slopes of the psychometric functions (d' vs ΔL) for decrement detection, divided by the square root of the ERD.

more important than information from the ongoing parts of the increments.

Data in the literature can also be analyzed to assess their consistency with the multiple-looks hypothesis. If the duration of an increment is increased by a factor R , and d' is proportional to $(\Delta I/I)^k$, then to restore detectability of the longer increment to that of the shorter one, $\Delta I/I$ should be decreased by a factor $(\sqrt[k]{R})^{1/k}$. Green *et al.* (1979) used a signal frequency of 1000 Hz and increment durations of 10 and 100 ms, and found that, for a continuous pedestal, k

TABLE IV. Slopes of the best-fitting straight lines to the data relating d' to ΔL for decrement detection. The lines were constrained to pass through the origin. Slopes are given for each subject for each steady-state duration and frequency. Geometric mean slopes are also given.

Duration (ms)	Frequency (Hz)											
	RP			GS			WH			Mean		
	250	1000	4000	250	1000	4000	250	1000	4000	250	1000	4000
5	-0.11	-0.20	-0.39	-0.08	-0.37	-0.25	-0.06	-0.19	-0.50	-0.08	-0.24	-0.38
10	-0.14	-0.22	-0.57	-0.09	-0.43	-0.30	-0.10	-0.24	-0.52	-0.11	-0.28	-0.45
20	-0.19	-0.33	-0.74	-0.12	-0.52	-0.80	-0.16	-0.35	-0.77	-0.15	-0.39	-0.77
50	-0.34	-0.46	-1.02	-0.25	-1.09	-1.11	-0.26	-0.54	-0.94	-0.28	-0.65	-1.02
100	-0.32	-0.47	-0.88	-0.31	-0.63	-1.23	-0.34	-0.58	-1.35	-0.32	-0.56	-1.13
200	-0.38	-0.39	-0.84	-0.29	-1.01	-0.56	-0.35	-0.54	-1.18	-0.34	-0.60	-0.82

≈ 2.2 , so the value of $\Delta I/I$ at threshold should differ by a factor of $3.16^{1/2.2} = 1.68$. This corresponds to 2.3 dB. In fact, the data showed a difference in threshold of 4–7 dB between the two durations, greater than the predicted value. A similar analysis applied to the data of Jeffress (1975) for detection of increments in a continuous 500-Hz pedestal also indicates that performance improved more with increasing duration than predicted by the multiple-looks hypothesis (Note that, although Jeffress described his results in terms of $\Delta I/I$, this was actually used to describe the “signal voltage increment divided by the masker voltage,” his $\Delta I/I$ corresponds to $\Delta p/p$ in the present paper.) For the intensity discrimination of pulsed tones (pedestal duration equal to increment duration), Green *et al.* (1979) found that, for a 10-ms stimulus, d' was proportional to $\Delta I/I$ raised to the power 0.53 (value averaged across three subjects). Therefore, the value of $\Delta I/I$ at threshold should change by a factor of $3.16^{1/0.53} = 8.8$ when the duration is changed from 10 to 100 ms. This corresponds to 9.4 dB. In fact, the data showed a difference in threshold of only 2–3 dB, less than the predicted value. Thus, in this case, performance improved less with increasing duration than predicted by the multiple-looks hypothesis.

Our results at 1000 Hz appear to be intermediate between the results obtained by Green *et al.* (1979) for a continuous pedestal and for a pedestal gated with the signal. Thus, it may be just a coincidence that performance in our experiment at 1000 Hz was *roughly* consistent with the multiple-looks hypothesis. Green *et al.* (1979) discussed the possibility that differences in the psychometric functions for continuous and gated pedestals, as quantified by the value of k , might be related to the amount of signal uncertainty; with a gated pedestal the signal starting time and duration are precisely defined, whereas for a continuous pedestal these quantities are less well defined. This could lead to lower values of k for a gated than for a continuous pedestal. We will return to this point later.

In summary, the results indicate that, with increasing duration, performance at 250 Hz increased more than predicted by a multiple-looks hypothesis, while performance at 4000 Hz increased less than predicted by such a hypothesis. The change in performance with duration at 1000 Hz was roughly consistent with a multiple-looks hypothesis.

F. Changes in d' with decrement duration

An analysis similar to that described in Sec. II E was carried out for decrement detection. The results are shown in the lower panel of Fig. 7. For ERDs up to 54 ms, the ratio $\text{slope}/\sqrt{\text{ERD}}$ shows a slight increase with increasing ERD, especially at 250 Hz. In other words, performance improves at a rate slightly greater than predicted by a multiple-looks hypothesis. For ERDs greater than 54 ms, the ratio $\text{slope}/\sqrt{\text{ERD}}$ decreases, as expected given that the slopes did not increase systematically for ERDs beyond 54 ms. The results for durations beyond 54 ms indicate either that the information from individual looks is not combined optimally, or that it is not weighted equally.

TABLE V. The change in threshold ($\Delta I/I$) for a tenfold increase in increment duration (from 10 to 100 ms) for our study and for several studies in the literature.

Study	Frequency (Hz)	Pedestal gating	Change (dB)
Present	250	Before signal	-5.3
Moore and Peters (1997)		Before signal	-6.4
Henning (1970)		With signal	-5.3
Florentine (1986)		With signal	-3.6
Moore and Peters (1997)		Continuous	-6.5
Present	1000	Before signal	-3.3
Moore and Peters (1997)		Before signal	-6.0
de Boer (1986)		Before signal	-1.7
Henning (1970)		With signal	-2.8
Green <i>et al.</i> (1979)		With signal	-3.0
Florentine (1986)		With signal	-5.3
Harris (1963)		Continuous	-4.1
Leshowitz and Raab (1967)		Continuous	-7.0
Green <i>et al.</i> (1979)		Continuous	-6.0
Moore and Peters (1997)		Continuous	-3.2
Present	4000	Before signal	-1.2
Moore and Peters (1997)		Before signal	-1.8
Oxenham (1997)		Before signal	-2.5
Henning (1970)		With signal	-1.1
Carlyon and Moore (1984)		With signal	-2.5
Dai <i>et al.</i> (1999)		With signal	-2.0
Moore and Peters (1997)		Continuous	-5.3
Dai <i>et al.</i> (1999)		Continuous	-4.3

G. Changes in threshold with increment duration—comparison with earlier results

Our data can be used to estimate changes in threshold as a function of increment duration. This provides a useful basis for comparison with earlier results. For this purpose, we define the threshold as the value of $\Delta I/I$ giving $d' = 1$. Previous studies have used a great variety of conditions, differing, for example, in signal rise/fall times, type of background noise, and level of the pedestal. To compare earlier results with our own, we have selected stimulus conditions where we can be reasonably sure that spectral splatter was not detectable, either because the rise/fall times used were sufficiently long, or because a background noise was used. The pedestal level was chosen to be as close as possible to that used in our study (70 dB SPL). The studies considered used pedestals gated with the signal (Henning, 1970; Green *et al.*, 1979; Carlyon and Moore, 1984; Florentine, 1986; Dai *et al.*, 1999), pedestals gated before the signal (de Boer, 1986; Moore and Peters, 1997; Oxenham, 1997), and continuous pedestals (Harris, 1963; Leshowitz and Raab, 1967; Green *et al.*, 1979; Moore and Peters, 1997; Dai *et al.*, 1999). We estimated the change in threshold, expressed as $10 \log_{10}(\Delta I/I)$, for a tenfold increase in duration, from 10 to 100 ms. In some cases this involved interpolation or (slight) extrapolation of the measured thresholds. Where data were presented at several durations, a smooth curve was fitted to the data, and the thresholds at 10 and 100 ms were derived from the curve. The results are shown in Table V.

For the signal frequency of 250 Hz, the changes are mostly around -5 to -6 dB (with the exception of the study by Florentine, 1986), and do not seem to depend systemati-

cally on whether the pedestal is gated with the signal, gated before the signal, or is on continuously. For the signal frequency of 1000 Hz, the variability across studies is somewhat larger. For a pedestal gated with the signal, the mean change over studies is -3.7 dB. For a pedestal gated before the signal, the mean change is also -3.7 dB. For a continuous pedestal, the mean change is somewhat larger at -5.1 dB. For the signal frequency of 4000 Hz, the changes are small for a pedestal gated with the signal (mean -2.0 dB) or before the signal (mean -1.8 dB), but the change is larger for a continuous pedestal (mean -4.8 dB). In summary, when the pedestal is gated with the signal or before the signal, the change in threshold for a tenfold increase in duration changes progressively with frequency, from about -5 dB at 250 Hz to -3.7 dB at 1000 Hz and to -1.8 dB at 4000 Hz. When the pedestal is continuous, the change in threshold with duration is roughly constant across frequency at about -5 to -6 dB.

III. ANALYSIS USING A TEMPORAL WINDOW MODEL

Previous studies of increment and decrement detection have often been analyzed using a model based on the concept of a sliding temporal integrator or “window” (Moore *et al.*, 1988; Plack and Moore, 1991; Peters *et al.*, 1995; Oxenham and Moore, 1994; Moore *et al.*, 1996, 1993; Oxenham, 1997). We used a similar model to analyze the present results.

A. Stages in the model

The model has the following stages:

(1) An array of bandpass filters (simulating the auditory filters). These are gammatone filters (Patterson *et al.*, 1995) with bandwidths as suggested by Glasberg and Moore (1990). As our stimuli were presented in background noise, we considered only the output of the auditory filter centered at the signal frequency; the outputs of the auditory filters at other center frequencies would have had a more adverse signal-to-noise ratio.

(2) A nonlinearity. Initially the output was full-wave rectified. Then it was raised to the power 0.7, i.e., the nonlinearity was compressive, consistent with other studies (Oxenham and Moore, 1994, 1995; Moore *et al.*, 1996; Moore and Oxenham, 1998).

(3) A sliding temporal integrator with each side modeled as an exponential function, described by

$$W(-t) = \exp(-t/T_b), \quad t < 0, \quad (1)$$

and

$$W(t) = \exp(-t/T_a), \quad t > 0, \quad (2)$$

where t is time measured relative to the time at the peak of the function, and T_b and T_a are parameters determining the sharpness of the function for times before the peak and times after the peak, respectively. The equivalent rectangular duration (ERD) of this function is defined as $T_b + T_a$. The data did not provide any information about the asymmetry of the temporal integrator, and we assumed that the asymmetry was similar to that measured in earlier studies (Oxenham and

Moore, 1994); the value of T_b was assumed to be 0.63 times the value of T_a .

(4) A decision device. When the stimulus contains a brief increment or decrement, the output of the temporal integrator shows a corresponding brief dip or peak. Several previous studies of increment and decrement detection (Moore *et al.*, 1988; Plack and Moore, 1991; Peters *et al.*, 1995; Oxenham and Moore, 1994; Moore *et al.*, 1996, 1993) have assumed that threshold is reached when the magnitude of the dip or peak, relative to the response to the steady-state portion of the stimulus, reaches a certain value. A similar detection criterion was suggested by Plomp (1964) and by Buus and Florentine (1985). Oxenham (1997) suggested that threshold might correspond to a fixed positive rate of change at the output of the temporal integrator. In the present analysis, we used a more general approach, based on the concept of a template (Dau *et al.*, 1996). We assumed that an internal representation of the signal, a template, was built up over a series of trials. This template is based on the output of the temporal integrator as a function of time in response to easily detectable signals. In a given trial, the interval for which the internal representation of the stimulus most closely matched the template was chosen as the signal interval. Dau *et al.* proposed that similarity should be measured by the correlation of the internal representation of each stimulus with the template. However, a drawback of using the correlation alone is that it does not take into account the extent to which the internal representation of a given stimulus matches the template in *magnitude*; small fluctuations in the internal representation caused by the background noise might give a shape similar to that of the template, and therefore a high correlation, but a listener would probably not identify such a small fluctuation as a signal. Therefore, we used a measure of similarity that took into account the size of features in the internal representation. The measure was the product of two quantities:

- (1) the correlation between the template and the internal representation of a given stimulus and
- (2) the slope of the linear regression line relating the sample values of the internal representation of the stimulus to the sample values of the template; the template and internal representation were sampled at 1-ms intervals. On average, the slope increases monotonically with increasing size of the signal.

Studies using measures of increment and decrement detection to estimate the shape of the temporal integrator have mainly used very short increment and decrement durations, up to about 20 ms, to avoid longer-term temporal integration effects. However, our data included durations much longer than this, as we were interested in these longer-term processes. Hence we based the ERD of the integrator on previous estimates, rather than estimating it from our own data. The value chosen was 7 ms for the center frequencies of 1000 and 4000 Hz, which is typical of the value found when a compressive nonlinearity is included in the model (Oxenham and Moore, 1994; Moore *et al.*, 1996). The value for the center frequency of 250 Hz was 9 ms, as previous studies have suggested that the ERD of the integrator increases

slightly for very low center frequencies (Plack and Moore, 1990).

We modeled the effect of the background noise by using a Gaussian white noise at the input to the simulated auditory filter. The noise level was chosen so that the signal-to-noise ratio at the output of the filter was, on average, 13.4 dB. A different sample of the noise was used for every stimulus presentation.

B. The formation of the template

We simulated the three-interval forced-choice task actually used in our experiment. The template was constructed, for a given block of trials, by averaging the output of the sliding temporal integrator as a function of time for five presentations of the increment or decrement at the highest magnitude used; recall that five such stimuli were presented at the start of each block in the actual experiment. It was necessary to choose an appropriate duration and range of temporal positions for the template. One possibility is that the template duration is equal to the whole duration of the pedestal; essentially, this assumes that the start and end of the template are perfectly defined by the start and end of the pedestal, and that there is no uncertainty as to when the signal might occur. Several researchers have assumed that listeners are able to select a time interval that is exactly matched to the duration and time of occurrence of a signal (Green, 1960; Dai and Wright, 1995). This seems unlikely for our experiment, as previous experiments using increments presented at predictable times or uncertain times in a continuous masker indicate that subjects have limited accuracy in selecting the appropriate time to listen (Leis-Rossio and Small, 1986; Chang and Viemeister, 1991). Under the conditions of our experiment, it seems likely that there was some temporal uncertainty, but not as much as for a continuous masker; we return to this point later. We assumed that the duration of the template was only slightly longer than the duration of the increment or decrement; the template was chosen to extend 8 ms on either side of the time when the increment or decrement occurred, allowing for the delay produced by the simulated auditory filter. Thus, the template duration was equal to the total duration of the increment or decrement (including the rise/fall times) plus 16 ms. To accommodate the effects of temporal uncertainty, for a given stimulus, the two quantities described earlier (the correlation of the template with the internal representation of the stimulus and the slope of the regression line of the sample values) were determined for time delays over a range ± 15 ms (in 1-ms steps) around the time corresponding to synchrony of the signal and template. The highest value of the product of the two quantities was determined for each of the three intervals in a trial, and the interval giving the highest value was chosen as the signal interval.

C. Incorporating the effects of internal noise

We found that we needed to assume internal noise as well as external noise to predict performance accurately; the external noise alone led to predicted performance better than actually observed. To model the effect of internal noise, the

internal representation of the stimulus on each trial was corrupted by addition of a Gaussian random variable, with zero mean, to each sample point (the template was represented by points at 1-ms intervals). This approach is similar to that used by Dau *et al.* (1996). To determine appropriate values for the standard deviation of the Gaussian variable, σ , we took the following approach. The value of σ was expressed relative to the mean output of the temporal window during the steady-state portions of the stimuli and was assumed to be a constant proportion of that mean output at a given center frequency; in other words, we assumed that Weber's law held, an assumption made by many previous researchers (Tanner, 1961; Pfafflin and Mathews, 1962; Green and Swets, 1974). For each frequency separately, stimuli with increments or decrements corresponding to the mean threshold value ($d' = 1$) measured at each duration were used as input to the model. For each of several values of σ , 1000 forced-choice trials were run. The percent correct values obtained were converted to d' values. The initial value of σ was chosen so that the mean value of d' across durations was close to 1. After this, the value of σ was "fine tuned" by generating predicted psychometric functions and by adjusting σ in small steps until the slopes of the predicted functions matched the obtained slopes as closely as possible. Each point on each psychometric function was based on 1000 forced-choice trials. For increments, the values of σ obtained in this way were 1.86, 1.25, and 1.00 at 250, 1000 and, 4000 Hz, respectively. For decrements, the values were 1.17, 0.75, and 0.60. It would be unwise to attach much significance to the fact that the values of σ were smaller for decrements than for increments at each frequency. This may simply reflect individual differences, as only subject RP ran using both increments and decrements. The finding that the values of σ decreased with increasing frequency is discussed later. Note that the values of σ did not vary with increment or decrement duration. Once the value of σ was chosen for a given frequency and direction of change (increment or decrement), that same value of σ was used to predict *all* of the psychometric functions for that frequency and direction of change.

D. Predicted psychometric functions

An example of a predicted psychometric function for increments is shown in the left column of Fig. 8. The function is plotted as a function of ΔL , $\Delta I/I$, and $\Delta p/p$ (from bottom to top). The solid lines are best-fitting linear functions, constrained to pass through the origin. For ΔL , the psychometric function shows a distinct upward curvature. The sum of the squared deviations of the points from the fitted line (hereafter denoted *ssq*) is 18.7. For $\Delta I/I$, the curvature is less distinct; the *ssq* is 10.5. For $\Delta p/p$, the function again shows a distinct upward curvature; the *ssq* is 13.9. Based on a one-tailed test *F* test (as we expected the fit to be better for $\Delta I/I$ than for the other measures), the *ssq* for $\Delta I/I$ is significantly less than for $\Delta p/p$ ($p < 0.025$) or for ΔL ($p < 0.0001$). This pattern is consistent with that observed in the experimental results; see Sec. II A. The example shown is typical of the results obtained for all frequencies and increment durations.

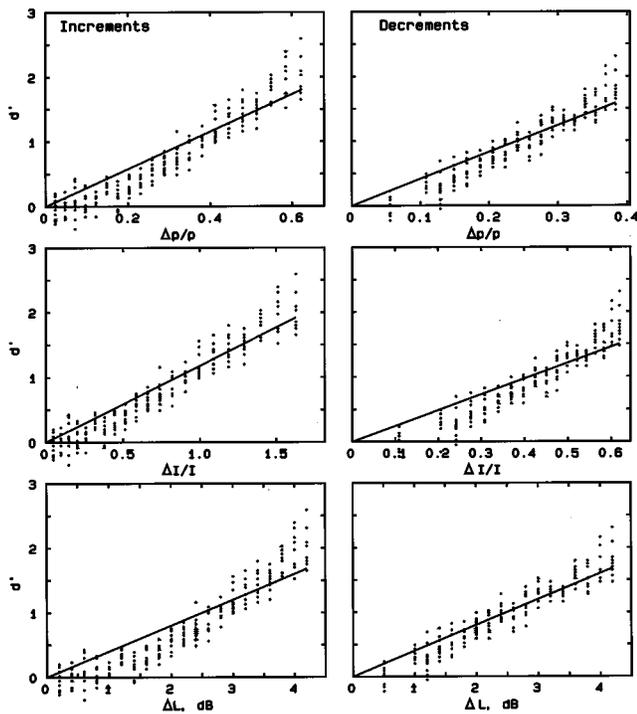


FIG. 8. Predictions of the model for an increment (left column) or decrement (right column) duration of 20 ms and a center frequency of 1000 Hz. The predicted value of d' is plotted as a function of the magnitude of the increment, expressed as ΔL , $\Delta I/I$, or $\Delta p/p$. Each point is based on 1000 forced-choice trials.

The right column of Fig. 8 shows an example of a predicted psychometric function for decrements, plotted as a function of ΔL , $\Delta I/I$, and $\Delta p/p$. The function is almost linear when plotted against ΔL ; the ssq is 7.0. The function shows a distinct upward curvature for $\Delta I/I$; the ssq is 11.2. The function for $\Delta p/p$ also shows a distinct upward curvature; the ssq is 8.8. The ssq for ΔL is significantly less than for $\Delta I/I$ ($p < 0.0004$) or for $\Delta p/p$ ($p < 0.05$). Again, this pattern of results is consistent with the experimental data. The example shown is typical of the results obtained for all frequencies and increment durations.

E. Predicted slopes of the psychometric functions

Straight lines were fitted to the predicted psychometric functions, using d' vs $\Delta I/I$ for increments and d' vs ΔL for decrements. Just as was done when fitting the data, the lines were constrained to pass through the origin. The range of predicted d' values used in the fitting procedure was chosen for each frequency and each increment and decrement duration to correspond to the range in the experimental data. The slopes of the resulting lines are plotted as solid lines in Fig. 9. For comparison, the experimentally obtained slopes are shown as symbols. Error bars indicate \pm one standard deviation across subjects.

Consider first the results for increment detection (top panel). At 250 and 1000 Hz the fit is reasonably good, and the form of the results is captured well. At 4000 Hz the model predicted a somewhat larger change in slope with duration than was actually observed, and the predicted slope for the shortest duration was smaller than observed. However, it

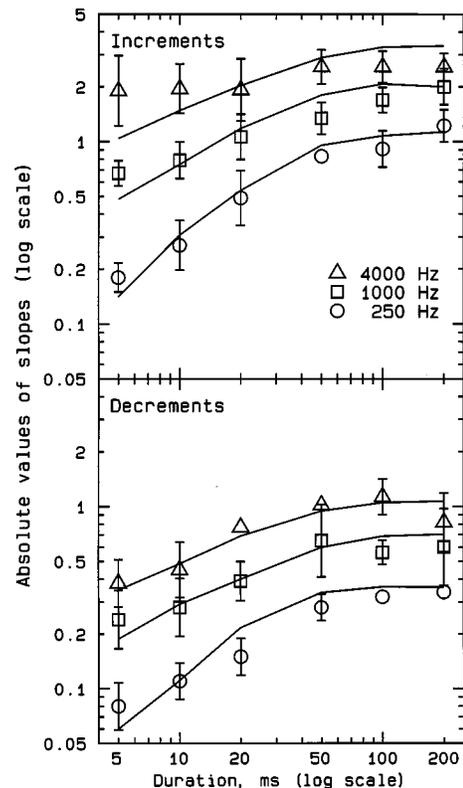


FIG. 9. The lines show the predicted slopes of the psychometric functions, plotted as a function of increment (top) or decrement (bottom) duration, with center frequency as parameter. The symbols show the experimentally obtained slopes. Slopes for increments were calculated from data/predictions relating d' to $\Delta I/I$. Slopes for decrements were calculated from data/predictions relating d' to ΔL . Error bars indicate \pm one standard deviation across subjects.

should be noted that individual variability was marked for the shortest duration; two subjects did show a distinct decrease in slope for the shortest duration, while one did not. Other data, summarized in Table V, suggest somewhat greater changes in threshold (and therefore in the slope of the psychometric function) with duration at 4000 Hz. The model correctly predicted that the ratio of slopes for the longest and shortest increment durations decreases with increasing frequency.

Consider now the results for decrement detection (bottom panel). The predicted slopes fit the general pattern of the obtained slopes very well; most (14 out of 18) of the predicted slopes lie within \pm one standard deviation of the obtained slopes. The model correctly predicts that the ratio of slopes for the longest and shortest increment durations decreases with increasing frequency.

In summary, the model captures the main features of the experimental results for both increment and decrement detection. It correctly predicts that the psychometric functions for increment detection are almost linear when d' is plotted against $\Delta I/I$, while those for decrement detection are almost linear when d' is plotted against ΔL . It also predicts the way that the slopes of the psychometric functions change with frequency and with increment/decrement duration.

F. Accounting for the effects of gating of the pedestal on the shape of the psychometric function

In the model, the effects of temporal uncertainty are accounted for by allowing the template to fall in a range of times, T_r , around the actual time of occurrence of the signal; in the modeling presented so far, T_r was ± 15 ms. To assess whether changes in temporal uncertainty could account for changes in the value of k with gating conditions, we used the model to generate psychometric functions for increment detection for three different degrees of temporal uncertainty. In one case, simulating the minimal temporal uncertainty associated with gating the pedestal and increment together, T_r was set to 0. In the second case, simulating the high temporal uncertainty associated with a continuous pedestal, T_r was set to ± 150 ms. In the third case, simulating the moderate temporal uncertainty of our experiment, T_r was set to ± 15 ms. Curves of the form $d' = m(\Delta I/I)^k$ were fitted to the model predictions. The best fitting value of k increased systematically with increasing T_r . For $T_r = 0$, the value of k was less than 1, while for $T_r = \pm 150$ ms, the value of k was markedly greater than 1. For $T_r = \pm 15$ ms, k was close to, but slightly greater than, 1. For example, for a frequency of 1000 Hz, and an increment duration of 5 ms, the values of k were 0.77 for $T_r = 0$, 1.19 for $T_r = \pm 15$ ms, and 1.75 for $T_r = \pm 150$ ms. For a frequency of 1000 Hz, and an increment duration of 200 ms, the values of k were 0.88 for $T_r = 0$, 1.26 for $T_r = \pm 15$ ms, and 1.46 for $T_r = \pm 150$ ms. Consistent with the data of Dai *et al.* (1999), the value of k tends to increase slightly with increasing duration for conditions of minimal temporal uncertainty ($T_r = 0$, pedestal gated with signal), but to decrease with increasing duration for conditions of high temporal uncertainty ($T_r = \pm 150$ ms, continuous pedestal). Overall, the changes in the values of k predicted by the model are in line with those observed experimentally.

The model fails in one respect in that it predicts lower thresholds when the pedestal is gated with the signal than when the pedestal is gated before the signal or is continuous; the predicted thresholds are lowest when the temporal uncertainty is smallest. This failure probably occurs because, when the pedestal is gated with the signal, the nature of the detection task changes (Harris, 1963). The extent to which the shape of the internal representation matches the shape of the template becomes largely irrelevant. Instead, the magnitude of the internal response is critical, and this magnitude must be remembered and compared across the stimuli comprising a forced-choice trial (Harris, 1963). This memory and comparison process, sometimes referred to as "trace mode," may involve an additional source of internal noise (Durlach and Braida, 1969). In contrast, when the increment is imposed on a longer-duration pedestal, the extent to which the shape of the internal response matches the shape of the template is critical. The similarity between the two may be coded in a way which is resistant to the effects of memory decay. The effects of "memory noise" in the trace mode could be incorporated in the model as an additional internal noise that corrupted the overall magnitude of the internal representation of each stimulus. However, we have not attempted to incorporate such a noise in the model.

G. Discussion

The value of σ necessary to fit the data decreased with increasing frequency, especially between 250 and 1000 Hz. A possible explanation for this is that the value of σ is partly influenced by the random fluctuations in the output of the sliding temporal integrator, which in turn result from random fluctuations in the background noise used in the experiment. As described earlier, the effects of random fluctuations in the background noise are greater at low frequencies than at high; this happens because the bandwidth of the auditory filter increases with increasing center frequency. Greater bandwidths lead to more rapid fluctuations in the noise, and these are more effectively smoothed by the sliding temporal integrator. The model clearly revealed such effects. When σ was set to zero, simulating the hypothetical case where performance was limited only by the external noise, performance was predicted to be markedly better at high frequencies than at low. However, in this case, predicted performance was better than obtained performance at all center frequencies.

If the internal noise represented by σ was linked to the random fluctuations at the output of the sliding temporal integrator, this could explain why σ decreased with increasing frequency. A possible reason for the linkage is connected with the phenomenon of modulation detection interference (MDI); the ability to detect or discriminate amplitude modulation at a given center frequency is often impaired by the presence of amplitude modulation at a remote center frequency (Yost *et al.*, 1989; Yost and Sheft, 1989). The detection of increments or decrements is probably closely related to amplitude modulation detection. The background noise in our experiments would have produced amplitude fluctuations at the outputs of auditory filters centered away from the signal frequency, and these fluctuations may have impaired performance. Consistent with this idea, Dau and Verhey (1999) have shown that the detection of amplitude modulation of a sinusoidal carrier can be impaired by the presence of a narrow-band noise centered at a frequency remote from that of the target carrier. In our experiment, the deleterious effect of fluctuations at the outputs of auditory filters centered away from the signal frequency may have been greater at low center frequencies, because the fluctuations there would be slower and would be less smoothed by the sliding temporal integrator.

If the value of σ is partly determined by the random fluctuations at the output of the sliding temporal integrator (for both on-frequency and off-frequency channels), then the variation of performance across frequency should be reduced or absent if increment/decrement detection is measured in the absence of background noise. However, in this case, the possibility of the detection of spectral splatter needs to be taken into account. Macmillan (1973) reported thresholds for increment and decrement detection in a continuous pedestal, without any background noise, for pedestal frequencies of 250 and 4000 Hz. Rise/fall times of 1 and 10 ms were used. It is unlikely that spectral splatter was detectable for the 10-ms rise/fall times. Performance was slightly better at 4000 than at 250 Hz for both rise/fall times, but the difference was much smaller than found in our experiment. The small effect of frequency is consistent with our suggestion

that the large effects of frequency in our experiment were partly caused by the background noise.

IV. CONCLUSIONS

The following conclusions can be drawn from this study:

- (1) The detectability, d' , of increments in level of a gated sinusoidal pedestal presented in a moderate level of background noise was approximately linearly related to $\Delta I/I$ (see Tables I and II). The detectability of decrements was approximately linearly related to ΔL (see Tables I and II).
- (2) The slopes of the psychometric functions relating d' to $\Delta I/I$ (for increments) or d' to ΔL (for decrements) increased with increasing frequency, i.e., the detectability of increments and decrements increased with increasing frequency (see Figs. 1–6).
- (3) The slopes of the psychometric functions for increment detection increased markedly with increasing increment duration at 250 Hz, increased somewhat less at 1000 Hz, and increased very little at 4000 Hz (see Figs. 1–3).
- (4) The slopes of the psychometric functions for decrement detection increased with decrement duration up to about 50 ms, and then remained roughly constant, for all three frequencies tested (see Figs. 4–6).
- (5) For a center frequency of 250 Hz, the slopes of the psychometric functions for increment detection increased with duration more rapidly than predicted by a “multiple-looks” hypothesis, i.e., more rapidly than the square root of duration. For center frequencies of 1000 and 4000 Hz, the slopes decreased less rapidly than predicted by a multiple-looks hypothesis, for durations greater than about 20 ms (see Fig. 7).
- (6) The slopes of the psychometric functions for decrement detection increased with decrement duration at a rate slightly greater than the square root of duration, for durations up to 50 ms. For greater durations, the increase in slope was less than proportional to the square root of duration (see Fig. 7).
- (7) The results were analyzed using a model incorporating a simulated auditory filter, a compressive nonlinearity, a sliding temporal integrator, and a decision device based on a template. The model took into account the effects of both the external noise and an assumed internal noise. The model was able to account for the main features of the results for both increment and decrement detection, including the form of the psychometric functions (Fig. 8), and the way that the slopes of the functions changed with frequency and with increment/decrement duration (Fig. 9). The model was also able to account for changes in the shape of the psychometric function with gating condition (pedestal gated with the signal or on continuously) found in previous studies.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council (UK). We thank Deborah Vickers, Emily Buss, Will Hula, Christin Adkins, Meghan Durham, Heidi Liefer,

Sharon Moore, and Virginia Smith for their assistance with various aspects of this study. We also thank two reviewers for their comments, which resulted in substantial improvements to the paper.

- Bacon, S. P., and Moore, B. C. J. (1986). “Temporal effects in simultaneous pure-tone masking: Effects of signal frequency, masker/signal frequency ratio, and masker level,” *Hearing Res.* **23**, 257–266.
- Bacon, S. P., and Viemeister, N. F. (1985). “The temporal course of simultaneous tone-on-tone masking,” *J. Acoust. Soc. Am.* **78**, 1231–1235.
- Buus, S., and Florentine, M. (1985). “Gap detection in normal and impaired listeners: the effect of level and frequency,” in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, New York).
- 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.
- Carlyon, R. P., and Moore, B. C. J. (1984). “Intensity discrimination: a severe departure from Weber’s Law,” *J. Acoust. Soc. Am.* **76**, 1369–1376.
- Carlyon, R. P., and Moore, B. C. J. (1986). “Continuous versus gated pedestals and the ‘severe departure’ from Weber’s Law,” *J. Acoust. Soc. Am.* **79**, 453–460.
- Carlyon, R. P., Buus, S., and Florentine, M. (1990). “Temporal integration of trains of tone pulses by normal and by cochlearly impaired listeners,” *J. Acoust. Soc. Am.* **87**, 260–268.
- Chang, P., and Viemeister, N. F. (1991). “Temporal windows for signals presented at uncertain times,” *J. Acoust. Soc. Am.* **90**, 2248.
- Dai, H., and Wright, B. A. (1995). “Detecting signals of unexpected or uncertain durations,” *J. Acoust. Soc. Am.* **98**, 798–806.
- Dai, H., Neff, D., and Jesteadt, W. (1999). “Psychometric function for detecting level increments added to continuous or gated pedestals,” *J. Acoust. Soc. Am.* (submitted).
- Dau, T., and Verhey, J. L. (1999). “Modelling across-frequency processing of amplitude modulation,” in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, V. Hohmann, and B. Kollmeier (World Scientific, Singapore).
- Dau, T., Püschel, D., and Kohlrausch, A. (1996). “A quantitative model of the ‘effective’ signal processing in the auditory system. I. Model structure,” *J. Acoust. Soc. Am.* **99**, 3615–3622.
- de Boer, E. (1986). “On thresholds of short-duration intensity increments and decrements,” in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson (Plenum, New York).
- Donaldson, G. S., Viemeister, N. F., and Nelson, D. A. (1997). “Psychometric functions and temporal integration in electric hearing,” *J. Acoust. Soc. Am.* **101**, 3706–3721.
- Durlach, N. I., and Braida, L. D. (1969). “Intensity Perception. I. Preliminary theory of intensity resolution,” *J. Acoust. Soc. Am.* **46**, 372–383.
- Eddins, D. A., and Green, D. M. (1995). “Temporal integration and temporal resolution,” in *Hearing*, edited by B. C. J. Moore (Academic, San Diego).
- Egan, J. P., Lindner, W. A., and McFadden, D. (1969). “Masking-level differences and the form of the psychometric function,” *Percept. Psychophys.* **6**, 209–215.
- Florentine, M. (1986). “Level discrimination of tones as a function of duration,” *J. Acoust. Soc. Am.* **79**, 792–798.
- Forrest, T. G., and Green, D. M. (1987). “Detection of partially filled gaps in noise and the temporal modulation transfer function,” *J. Acoust. Soc. Am.* **82**, 1933–1943.
- Glasberg, B. R., and Moore, B. C. J. (1990). “Derivation of auditory filter shapes from notched-noise data,” *Hear. Res.* **47**, 103–138.
- Grantham, D. W., and Yost, W. A. (1982). “Measures of intensity discrimination,” *J. Acoust. Soc. Am.* **72**, 406–410.
- Green, D. M. (1960). “Auditory detection of a noise signal,” *J. Acoust. Soc. Am.* **32**, 121–131.
- Green, D. M. (1988). *Profile Analysis* (Oxford U. P., Oxford).
- Green, D. M. (1993). “Auditory intensity discrimination,” in *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer-Verlag, New York).
- Green, D. M., and Swets, J. A. (1974). *Signal Detection Theory and Psychophysics* (Krieger, New York).

- Green, D. M., Nachmias, J., Kearney, J. K., and Jeffress, L. A. (1979). "Intensity discrimination with gated and continuous sinusoids," *J. Acoust. Soc. Am.* **66**, 1051–1056.
- Hafter, E. R., Bonnel, A. M., and Gallun, E. J. (1996). "Detection of change without regard to its valence," *J. Acoust. Soc. Am.* **99**, 2541.
- Harris, J. D. (1963). "Loudness discrimination," *J. Speech. Hear. Disord. Monographs, Supplement* **11**, 1–63.
- Henning, G. B. (1970). "Comparison of the effects of signal duration on frequency and amplitude discrimination," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden).
- Jeffress, L. A. (1975). "Masking of tone by tone as a function of duration," *J. Acoust. Soc. Am.* **58**, 399–403.
- Leis-Rossio, B., and Small, A. M. (1986). "Temporal specificity," *J. Acoust. Soc. Am.* **80**, S92.
- Leshowitz, B., and Raab, D. H. (1967). "Effects of stimulus duration on the detection of sinusoids added to continuous pedestals," *J. Acoust. Soc. Am.* **41**, 489–496.
- Macmillan, N. A. (1971). "Detection and recognition of increments and decrements in auditory intensity," *Percept. Psychophys.* **10**, 233–238.
- Macmillan, N. A. (1973). "Detection and recognition of intensity changes in tone and noise: the detection-recognition disparity," *Percept. Psychophys.* **13**, 65–75.
- Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide* (Cambridge U. P., Cambridge, England).
- Moore, B. C. J., and Oxenham, A. J. (1998). "Psychoacoustic consequences of compression in the peripheral auditory system," *Psychol. Rev.* **105**, 108–124.
- Moore, B. C. J., and Peters, R. W. (1997). "Detection of increments and decrements in sinusoids as a function of frequency, increment and decrement duration and pedestal duration," *J. Acoust. Soc. Am.* **102**, 2954–2965.
- Moore, B. C. J., and Sek, A. (1992). "Detection of combined frequency and amplitude modulation," *J. Acoust. Soc. Am.* **92**, 3119–3131.
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1993). "Effects of frequency on the detection of decrements and increments in sinusoids," *J. Acoust. Soc. Am.* **94**, 3190–3198.
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1996). "Detection of decrements and increments in sinusoids at high overall levels," *J. Acoust. Soc. Am.* **99**, 3669–3677.
- Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (1988). "The shape of the ear's temporal window," *J. Acoust. Soc. Am.* **83**, 1102–1116.
- Oxenham, A. J. (1997). "Increment and decrement detection in sinusoids as a measure of temporal resolution," *J. Acoust. Soc. Am.* **102**, 1779–1790.
- Oxenham, A. J., and Moore, B. C. J. (1994). "Modeling the additivity of nonsimultaneous masking," *Hearing Res.* **80**, 105–118.
- Oxenham, A. J., and Moore, B. C. J. (1995). "Additivity of masking in normally hearing and hearing-impaired subjects," *J. Acoust. Soc. Am.* **98**, 1921–1935.
- Patterson, R. D., Allerhand, M. H., and Giguère, C. (1995). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.* **98**, 1890–1894.
- Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (1995). "Effects of level and frequency on the detection of decrements and increments in sinusoids," *J. Acoust. Soc. Am.* **97**, 3791–3799.
- Pfafflin, S. M., and Mathews, M. V. (1962). "Energy detection model for monaural auditory detection," *J. Acoust. Soc. Am.* **34**, 1842–1852.
- Plack, C. J., and Carlyon, R. P. (1995). "Loudness perception and intensity coding," in *Hearing*, edited by B. C. J. Moore (Academic, Orlando, FL).
- Plack, C. J., and Moore, B. C. J. (1990). "Temporal window shape as a function of frequency and level," *J. Acoust. Soc. Am.* **87**, 2178–2187.
- Plack, C. J., and Moore, B. C. J. (1991). "Decrement detection in normal and impaired ears," *J. Acoust. Soc. Am.* **90**, 3069–3076.
- Plomp, R. (1964). "The rate of decay of auditory sensation," *J. Acoust. Soc. Am.* **36**, 277–282.
- Rabinowitz, W. M., Lim, J. S., Braidia, L. D., and Durlach, N. I. (1976). "Intensity perception. VI. Summary of recent data on deviations from Weber's law for 1000-Hz tone pulses," *J. Acoust. Soc. Am.* **59**, 1506–1509.
- Tanner, W. P. (1961). "Application of the theory of signal detectability to amplitude discrimination," *J. Acoust. Soc. Am.* **33**, 1233–1244.
- Taylor, M. M., and Forbes, S. M. (1983). "PEST reduces bias in forced-choice psychophysics," *J. Acoust. Soc. Am.* **74**, 1367–1374.
- Viemeister, N. F. (1988). "Psychophysical aspects of auditory intensity coding," in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. A. Cowan (Wiley, New York).
- Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal integration and multiple looks," *J. Acoust. Soc. Am.* **90**, 858–865.
- Yost, W. A., and Sheft, S. (1989). "Across-critical-band processing of amplitude-modulated tones," *J. Acoust. Soc. Am.* **85**, 848–857.
- Yost, W. A., Sheft, S., and Opie, J. (1989). "Modulation interference in detection and discrimination of amplitude modulation," *J. Acoust. Soc. Am.* **86**, 2138–2147.

Context dependence of fundamental-frequency discrimination: Lateralized temporal fringes

Hedwig Gockel^{a)}

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom
and Department of Experimental Psychology, University of Cambridge, Downing Street,
Cambridge CB2 3EB, United Kingdom

Robert P. Carlyon

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom

Christophe Micheyl

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom
and Laboratoire "Neurosciences & Systèmes Sensoriels" UPRESA CNRS 5020, Pavillon U-Hôpital E.
Herriot, 69437 Lyon cedex 03, France

(Received 4 March 1999; revised 9 July 1999; accepted 2 September 1999)

In a two-interval, two-alternative, forced-choice (2I-2AFC) adaptive procedure, listeners discriminated between the fundamental frequencies (F_0 s) of two 100-ms harmonic target complexes. This ability can be impaired substantially by the presence of another complex (the "fringe") immediately before and after each target complex. It has been shown that for the impairment to occur (i) target and fringes have to be in the same frequency region; (ii) if all harmonics of target and fringes are unresolved then they may differ in F_0 ; otherwise, they have to be similar [C. Micheyl and R. P. Carlyon, *J. Acoust. Soc. Am.* **104**, 3006–3018 (1998)]. These findings have been discussed in terms of information about the fringe's F_0 being included in the estimate of the F_0 of the target, and in terms of auditory streaming. The present study investigated the role of perceived location and ipsilateral versus contralateral presentation of the fringes on F_0 discrimination of the target. Experiment 1 used interaural level differences (ILDs), and experiment 2 used interaural time differences (ITDs) to create a range of lateralized perceptions of the 200-ms harmonic fringes. Difference limens for the F_0 of the monaural target complex were measured in the presence and absence of the fringes. The nominal F_0 was 88 or 250 Hz and could be the same or different for target and fringes. Stimuli were bandpass filtered between 125–625, 1375–1875, or 3900–5400 Hz. In both experiments, the effect of the fringes was reduced when their subjective location differed from that of the target. This reduction depended on the resolvability of both the fringes and the target. The effect of the fringes was reduced most (but still present), when fringes were presented purely contralaterally to the target. The results are consistent with the idea that the fringes produce interference when the listeners have difficulty segregating the target from the fringes, and that a difference in perceived location enhances segregation of the sequentially presented stimuli. © 1999 Acoustical Society of America. [S0001-4966(99)05712-4]

PACS numbers: 43.66.Fe, 43.66.Hg, 43.66.Rq, 43.66.Mk [DWG]

INTRODUCTION

In everyday life we are often confronted with a mixture of competing sounds. When listeners try to hear a specific target sound, e.g., a voice, they exploit the fact that the frequencies of the components emitted by a single periodic source usually correspond to integer multiples of a fundamental frequency (F_0). Even when there is no physical energy at the F_0 , such harmonically related components elicit a pitch sensation corresponding to F_0 (Schouten, 1940; Moore, 1997). For the purpose of structuring our complex auditory environment, i.e., determining the sound sources present, this process of grouping harmonically related components into a single percept has been shown to be crucial

(Brokx and Nooteboom, 1982; Scheffers, 1983; Moore *et al.*, 1986; Assmann and Summerfield, 1990; Bregman, 1990; Hartmann *et al.*, 1990; Hartmann, 1996).

The present study is part of a series investigating the mechanisms of F_0 encoding and how they are affected by the presence of other sounds (Carlyon *et al.*, 1992; Carlyon, 1994a, b; Carlyon and Shackleton, 1994; Shackleton and Carlyon, 1994; Plack and Carlyon, 1995; Carlyon, 1996a, b; Micheyl and Carlyon, 1998). Specifically, Carlyon (1996b) showed that F_0 discrimination can be impaired substantially by the presence of another harmonic complex (the "fringe") immediately before and after the target complex. He attributed this sequential effect to listeners "overintegrating" information, so that the estimate of the target- F_0 was influenced by F_0 information from the fringe, and rejected other explanations based on peripheral adaptation and on mechanisms such as backward recognition masking or "blanking"

^{a)}Electronic mail: hedwig.gockel@mrc-cbu.cam.ac.uk

(Divenyi and Hirsh, 1975; Massaro, 1975; Kelly and Watson, 1986). More recent experiments (Micheyl and Carlyon, 1998) showed that for the impairment to occur (i) the target and fringes have to be in the same frequency region; (ii) if all harmonics of the target and fringes are unresolved by the peripheral auditory system, then they may differ in F_0 ; otherwise they have to be similar. Micheyl and Carlyon (1998) discussed their results in terms of auditory streaming. Due to the measurement procedure used, the triplet of fringe–target–fringe was presented repeatedly. The authors argued that this resulted in subjects perceiving the fringes and the targets in separate streams, *if* they differed in F_0 or spectral content. For these conditions, subjects reported being able to focus on the target, and the fringes did not impair performance. Thus, the authors suggested that overintegration of information about the fringes' F_0 when estimating that of the target, might depend on perceiving the stimuli in one stream.

While there is evidence in the literature that pure or complex tones presented sequentially can lead to streaming effects (Bregman and Campbell, 1971; van Noorden, 1975; Bregman, 1990), the stimuli in those studies differed not only in F_0 , but also in spectral content. Thus, the streaming effects observed could rely on spectral differences between the tones. Micheyl and Carlyon (1998) included a condition (experiment 4) where target and fringes differed in spectral content but not in F_0 , which was fixed at 250 Hz. This was achieved by using a method originally devised by Moore and Glasberg (1990), which involved using components with harmonic numbers 1, 4, 5, 8, 9, 12, 13, etc. for the target, and components with harmonic numbers 2, 3, 6, 7, 10, 11, etc. for the fringes, and vice versa. Both stimuli were then band-pass filtered between 1375 and 1875 Hz (mid region). The spectral differences between targets and fringes, as determined through excitation-pattern simulations, were considerably larger than when target and fringes differed in F_0 (being 88 and 250 Hz) but consisted of intact, overlapping harmonics series. For the former stimuli they found a strong impairment in F_0 discrimination performance, while for the latter stimuli no impairment was observed. Thus, the impairment depended less on local spectral similarity of target and fringes than it depended on similarity of F_0 , *per se*. This showed that if interference depended on perceiving the target and fringes in one stream, then F_0 , *per se*, was an effective cue in streaming too. A role for F_0 in streaming is also supported by the results of Vliegen and Oxenham (1999), and Vliegen *et al.* (1999).

The purpose of the present experiments was to investigate further the possible influence of streaming on the effects of fringes on F_0 discrimination. In Micheyl and Carlyon's (1998) study all stimuli were presented monaurally. The present study investigated the effect of ear of entry and perceived location of the fringes relative to the monaurally presented target on F_0 discrimination of the target. Specifically, it was asked whether contralateral presentation and perceived contralateral location of the fringes relative to the target would reduce their deleterious effect. This would provide some information concerning the possible "site" of the af-

ected mechanism and its susceptibility to perceptual segregation.

Pitch perception for complex tones (at least for tones with resolved harmonics) generally is thought to be a central process which is largely unaffected by differences in either ear of presentation or perceived location of components. For example, Darwin and Ciocca (1992) found pitch shifts—produced by mistuning one harmonic—to be only slightly larger when the mistuned harmonic was presented to the same ear as the rest of the complex as opposed to the opposite ear (see also Houtsma and Goldstein, 1972; Beerends and Houtsma, 1989). Additionally, Darwin *et al.* (1995) showed that not only the pitch shifts but also the reduction in the pitch shift by a tonal-sequence precursor remained when the mistuned component and the precursor were presented contralaterally to the remaining components. Similarly, for detection of mistuning of a single, low-numbered harmonic, Lee and Green (1994) found thresholds to be unaffected by the side of presentation of the mistuned harmonic relative to the remainder of the complex. Thus, if the observed sensitivity of F_0 discrimination to the context in which the complex target sounds are presented were based purely on this central pitch-perception process (by means of overintegration of information), one would expect no substantial effect of contralateral vs ipsilateral presentation, and perceived location of the fringes.

However, the situation may not be that clear-cut. While, for steady-state sounds, contralateral presentation or a different perceptual location of components made no difference, the findings of Gockel and Carlyon (1998) showed that it played a role for complex tones with dynamic changes in F_0 . Gockel and Carlyon asked their listeners to detect mistuning imposed on one component of a complex tone. All components were frequency modulated by a 5-Hz sinusoid. Mistuning was introduced by inverting the phase of the frequency modulation for a single target component. Performance was impaired when the nontarget components were presented contralaterally to the target component. However, performance was not impaired when the nontarget components were presented dichotically with an ILD of 20 dB, so that they were perceived contralaterally to the target. Thus, relative ear of entry but not perceived location was important in a task relying crucially on *concurrent* sound segregation.

Sequential streaming, on the other hand, has been found to be affected by both relative ear of entry (Deutsch, 1979; Hartmann and Johnson, 1991; Hukin and Darwin, 1995), and perceived location (Kubovy and Howard, 1976; Hartmann and Johnson, 1991; Hukin and Darwin, 1995; Darwin and Hukin, 1999). Thus, if F_0 discrimination performance were affected by sequential streaming of targets and fringes, one might expect better performance when target and fringes were perceived at different locations than when target and fringes were presented monaurally to the same side.

Experiment 1 used interaural level differences (ILDs), and experiment 2 used interaural time differences (ITDs), to create a range of different lateralizations of the fringes. A crucial distinction is drawn between those effects driven by the perceived location of the fringes, and those determined by their "relative ear of entry" (ipsilateral versus contralat-

eral presentation of the fringes relative to the monaurally presented target). When the objective was to investigate the effect of perceived location of the fringes, there were no level differences between the target and the fringes in the ear ipsilateral to the target. Since level differences themselves can lead to sequential segregation (Hartmann and Johnson, 1991; Rogers and Bregman, 1998), this is an important point which has not been taken into account sometimes (e.g., demonstration number 38 on Bregman and Ahad's CD (1995) intends to illustrate streaming by spatial location, but does not prevent streaming due to level differences between the tones presented to a specific side).

I. GENERAL METHOD

In this section, the general procedure and stimuli will be described. Variations in stimulus presentation in the different conditions will be given in the sections describing the individual experiments.

The difference limen for F_0 (DLF_0) for a 100-ms harmonic target complex was measured in the presence and absence of 200-ms harmonic temporal fringes. In a two-interval, two-alternative, forced-choice (2I-2AFC) task, subjects were required to indicate the interval with the higher target F_0 . Feedback was provided following each response. The difference in F_0 (ΔF_0) between the two targets was adjusted using a two-down, one-up geometric adaptive tracking rule (Levitt, 1971), corresponding to 71%-correct responses. Initially, the ΔF_0 was divided or multiplied by a factor of 2. Following four reversals, the factor was reduced to 1.41, and 12 further reversals were obtained. The threshold estimate was defined as the geometric mean of the ΔF_0 values at the last 12 reversals.

The F_0 s of the two successive targets were geometrically centered on a nominal F_0 of either 88 or 250 Hz. The fringes (when present) were presented immediately before and after the target (zero silent interval), but did not overlap with it. The F_0 of the fringes was either 88 or 250 Hz; all four combinations of target and fringe F_0 were tested. Both the target and the fringes consisted of a large number of harmonics which were summed up in sine phase. All stimuli were bandpass filtered into the same frequency region, which in three different conditions could be the low, mid, or high region; these regions had passbands (3-dB down points) of: 125–625, 1375–1875, and 3900–5400 Hz, respectively. According to Shackleton and Carlyon's (1994) definition of resolvability, harmonics of an 88-Hz F_0 are resolved only in the low region, whereas harmonics of a 250-Hz F_0 are resolved in the low and mid region, but unresolved in the high region. The number of harmonics presented varied with F_0 and region, but always "filled" the passband and skirts down to a level of 0 dB SPL. Each target and each fringe was gated with 20-ms raised-cosine ramps. The interstimulus interval (ISI) between the end of the first and the beginning of the second presentation interval in each 2AFC trial (between targets when no fringes were presented, and between fringe–target–fringe triplets when fringes were presented) was 500 ms. All stimuli were generated digitally, played out by a 16-bit digital-to-analog converter (CED 1401 plus) at a sampling rate of 40 kHz, and passed through an antialiasing

filter (Kemo VBF25.01) with a cutoff frequency of 17.2 kHz (slope of 100 dB/oct). They were then bandpass filtered using a pair of cascaded filters (one high pass, one low pass, Kemo VBF25.03, each with slopes of 48 dB/oct). The stimuli were presented in a background of pink noise (low-pass filtered at 20 kHz, with a spectrum level of 10 dB at 1 kHz), using Sennheiser HD 414 headphones. Subjects were seated individually in an IAC double-walled sound-attenuating booth.

In both experiments at least four normal hearing subjects, one of whom was the first author, participated in all conditions. The total duration of a single session was about 2 h, including rest times. After at least 10 h of practice, six threshold estimates were obtained for each condition and subject. The geometric mean of these six estimates was defined as the final threshold for each condition and subject. Data collection for the three frequency regions was done in separate blocks of runs. The order of the frequency regions was balanced over subjects. Within a given frequency region the various lateralization conditions were run in separate sub-blocks. In each session, two thresholds in the no-fringe condition, and four thresholds for each combination of target and fringe– F_0 (in separate sub-blocks) were collected in one specific lateralization condition ($2+4\times 4=18$). For each of the four target–fringe– F_0 combinations, the first run was regarded as a practice run and discarded from the data analysis. In this way, three thresholds were collected first for each F_0 combination in each lateralization condition. Then, three more thresholds were collected in each condition in the same way, but with the order of the lateralization conditions reversed. The order of running through the individual lateralization conditions was counterbalanced over subjects, as was the order of running through the target–fringe– F_0 combinations. This design was chosen to enable subjects to get used to the new stimulus and lateralization, and to reduce the contrasting influence of the condition in the run before.

II. EXPERIMENT 1: LATERALIZATION DUE TO ILD

A. Rationale and method

There is some evidence that lateralization due to ILD affects sequential streaming (Deutsch, 1979; Bregman, 1990; Hartmann and Johnson, 1991; Rogers and Bregman, 1998). In the present F_0 discrimination task, fringes and target were presented sequentially. The triplet fringe/target/fringe was presented twice in each trial (separated by an ISI of 500 ms), and a run consisted of at least 24 trials. This relatively long sequence of stimuli might have encouraged subjects to perceptually segregate fringes and target into separate streams, given that they differ on a dimension which is relevant for auditory streaming (Bregman, 1978; Anstis and Saida, 1985). Thus, if F_0 discrimination performance were affected by sequential streaming of targets and fringes, one might expect a difference in perceived location of the fringes relative to the target to reduce their deleterious effect on performance. Experiment 1 investigated the role of ear of entry and perceived location of the fringes—due to ILDs—relative to the target.

The target was always presented monaurally at a level of

TABLE I. DLF0s for complex tones presented in isolation, expressed as a percentage of the F_0 . Values specify the geometric mean thresholds over subjects. The associated geometrical standard errors, expressed as factor, are indicated in parentheses. (a) Results in experiment 1. (b) Results in experiment 2.

F_0 (Hz)	Low	Mid	High
(a)			
88	0.43 (1.29)	1.43 (1.09)	2.23 (1.14)
250	0.34 (1.44)	0.37 (1.38)	1.49 (1.13)
(b)			
88	0.36 (1.23)	1.11 (1.07)	1.56 (1.21)
250	0.28 (1.21)	0.22 (1.39)	1.31 (1.03)

45 dB SPL to the left ear of the subjects (all levels are specified as the level per component before any filtering was applied). The variable of main interest, the lateralization of the fringes (if present) relative to the target was manipulated using the following conditions: (i) contralateral at a level of 45 dB SPL (“*c45*”); (ii) contralateral at a level of 65 dB SPL (“*c65*”); (iii) dichotic at a level of 45 dB SPL ipsilateral, and 65 dB SPL contralateral (“*i45c65*”); (iv) diotic at a level of 45 dB SPL (“*i45c45*”); and (v) ipsilateral at a level of 45 dB SPL (“*i45*”). Comparison of thresholds in the *c45* condition with those in the no-fringe condition would reveal whether contralateral presentation of the fringes had any effect on F_0 discrimination of the target. A difference between thresholds in the *c45* and the *i45* condition (lower thresholds in condition *c45* would be expected) would show either an effect of relative ear of entry, or of perceived location, or both. If an observed difference between thresholds in the *c45* and the *i45* condition were at least partly due to the difference in perceived location, then one would also expect thresholds in condition *i45c65* to be lower than in condition *i45*, at least if there were no negative effect of increased level of the fringes on the contralateral side. Condition *c65* was introduced to check on this possible level effect. If perceived location of the fringes is the crucial factor, then thresholds in condition *i45c45* should be intermediate between those in condition *i45* and those in condition *c45*.

Six subjects were tested using the mid-frequency region, four of whom were also tested using the low- and high-frequency regions. One of the six subjects in the mid region (the third author) produced results very different from the general pattern produced by all the other subjects. This deviation will be described in the Results section. However, this subject’s data were excluded from further analysis, and from the group data shown below.

B. Results

1. No fringe

Table I(a) shows the geometric mean DLF0s and the corresponding geometrical standard errors for the target tones presented in isolation. In the absence of the fringes, the overall sizes of the DLF0s were a function of (un)resolvability in all three frequency regions. That is, the observed DLF0s were small (values between 0.3% to 0.4%) when the individual components in the target complex were resolved (low region: F_0 equals 88 or 250 Hz; mid region: F_0 equals

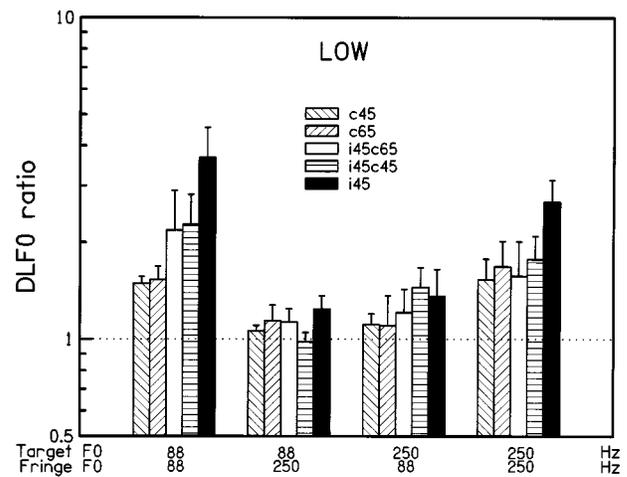


FIG. 1. Results in the low-frequency region, in experiment 1. The bars show the DLF0s for the 88- and 250-Hz targets in the presence of 88- or 250-Hz fringes, divided by the DLF0 for each target when presented alone. The four groups of bars correspond to the four combinations of target–fringe– F_0 . The individual bars correspond to conditions (from left to right): *c45*, *c65*, *i45c65*, *i45c45*, *i45*. Indicated are the geometric mean ratios, averaged over the ratios obtained for the individual subjects, and the associated geometrical standard errors. The ratio within one subject was calculated by dividing the geometric mean threshold derived from six runs in the presence of a fringe by the geometric mean threshold derived from six runs when the target was presented alone.

250 Hz), and large (values between 1.4% to 2.2%) when they were unresolved (Hoekstra, 1979; Moore *et al.*, 1984; Shackleton and Carlyon, 1994; Micheyl and Carlyon, 1998). The impairment caused by the presence of the fringes depended on condition and frequency region.

2. Low-frequency region

Figure 1 shows the results in the presence of the fringes for the low-frequency region. The bars show the DLF0s for the 88- and 250-Hz targets in the presence of 88- or 250-Hz fringes, divided by the DLF0 for each target when presented alone. There are four groups of bars corresponding to the four combinations of target–fringe– F_0 . Each group consists of five bars corresponding to the five different conditions of lateralization of the fringes. The individual bars show the conditions (from left to right): *c45*, *c65*, *i45c65*, *i45c45*, and *i45*.

In the low-frequency region, the fringes raised the DLF0s significantly (the DLF0 ratio was significantly higher than 1), but only when target and fringes had the same nominal F_0 . This was supported by the outcome of a three-way analysis of variance (ANOVA¹) (target– $F_0 \times$ fringe– $F_0 \times$ lateralization) which showed a significant two-way interaction between the F_0 of the target and that of the fringe [$F(1,3) = 15.8, p < 0.05$]. Thus, Micheyl and Carlyon’s (1998) finding that ipsilaterally presented fringes have no effect on DLF0s if target and fringes differ in nominal F_0 could be replicated and generalized to different lateralizations. Therefore, two separate two-way ANOVAs were calculated. One was based exclusively on ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes was the same [$F(1,3) = 19.09, p < 0.05$] and the other

TABLE II. Planned contrasts between log-transformed ratios of DLF0s measured in experiment 1 in the presence and in the absence of temporal fringes, for ipsilateral presentation of the fringes versus other lateralization conditions. The significance levels specified were corrected for multiple comparisons. (a) Results in low-frequency region. (b) Results in mid-frequency region. (c) Results in high-frequency region for different nominal F_0 of target and fringes. (d) Results in high-frequency region for identical nominal F_0 of target and fringes.

(a) Low Target and fringes at identical F_0			(b) Mid Target and fringes at identical F_0		
Contrast	$F(1,12)$	p	Contrast	$F(1,16)$	p
$i45-c45$	36.78	<0.001	$i45-c45$	66.84	<0.001
$i45-c65$	31.2	<0.01	$i45-c65$	60.43	<0.001
$i45-i45c65$	19.06	<0.01	$i45-i45c65$	26.56	<0.001
$i45-i45c45$	13.74	<0.05	$i45-i45c45$	12.62	<0.05
(c) High Target and fringes at different F_0			(d) High Target and fringes at identical F_0		
Contrast	$F(1,12)$	p	Contrast	$F(1,12)$	p
$i45-c45$	43.01	<0.001	$i45-c45$	32.65	<0.001
$i45-c65$	32.13	<0.01	$i45-c65$	28.27	<0.001
$i45-i45c65$	36.1	<0.01	$i45-i45c65$	2.65	>0.50
$i45-i45c45$	37.49	<0.01	$i45-i45c45$	0.00	>0.50

one was based exclusively on ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes differed. In the latter case, there was no significant effect of lateralization.

When target and fringes had the same nominal F_0 , thresholds were highest for ipsilateral presentation (filled black bars) of the fringes. In conditions $i45c45$ and $i45c65$, where the fringes were lateralized away from the target, the deleterious effect of the fringes was reduced when compared with condition $i45$ (ratios were smaller). Note that in these conditions, identical sounds were delivered to the left (target) ear. This effectiveness of perceived location observed in the present task stands in marked contrast to its ineffectiveness if listeners have to detect a mistuning in complex tones with dynamic changes in F_0 (Gockel and Carlyon, 1998). Finally, DLF0s were smallest in conditions $c45$ and $c65$ (strictly true for condition $F_0=88/88$; smallest or equal to smallest in condition $F_0=250/250$) where the fringes were presented purely contralaterally to the target. These results were confirmed by the outcome of the two-way ANOVA calculated exclusively for the ratios of DLF0s obtained in conditions with identical nominal target and fringe F_0 , which showed a highly significant main effect of lateralization [$F(4,12) = 11.49, p < 0.01, HH = 0.87$].² Planned contrasts showed that for all lateralizations of the fringes away from the target performance improved significantly when compared with ipsilateral presentation [see Table II(a) for statistical details]. However, even for purely contralateral presentation (conditions $c45$ and $c65$) the fringes still impaired performance [$F(1,3) = 15.91, p < 0.05$]. Note also that the 20-dB increase in level of the purely contralateral fringes did not result in a significant increase in thresholds.

To summarize the most important findings for the low-frequency region, where harmonics were resolved for both F_0 s: (i) F_0 discrimination was only impaired by the presence of temporal fringes, when target and fringes had identical nominal F_0 ; (ii) The amount of impairment was reduced when the fringes were perceived at a location different

from that of the target; (iii) Purely contralateral fringes produced a significant impairment which was, however, smaller than that in all other conditions, and was independent of their presentation levels.

3. Mid-frequency region

In the mid-frequency region (Fig. 2), results in the presence of the fringes were in all respects very similar to those observed in the low-frequency region.³ The data of the third author, who was tested using only the mid region, differed from the above-described general pattern in the following way: (i) in condition $c45$ his performance was impaired more than for the other subjects; (ii) a distinct level effect was observed for the purely contralateral fringes, i.e., the more intense fringes led to more impairment; (iii) thresholds were highest in condition $i45c65$, and second highest in $i45c45$. A possible explanation for these data will be discussed later, after presenting his data from the second (ITD) experiment.

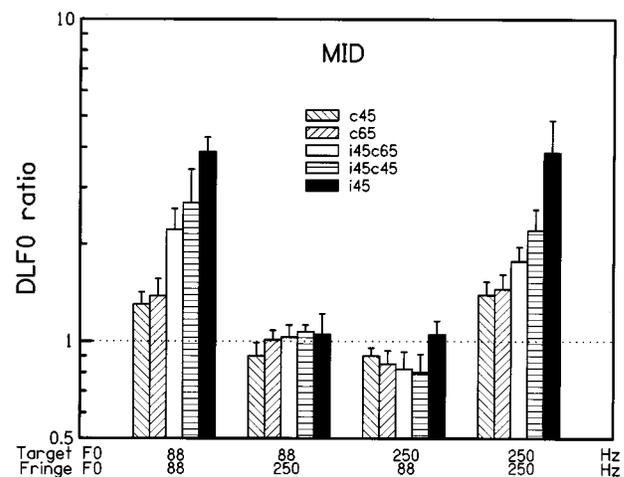


FIG. 2. As Fig. 1, but for the mid region.

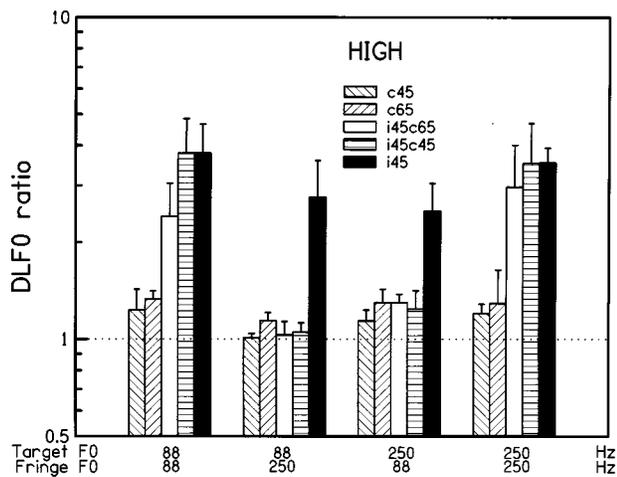


FIG. 3. As Fig. 1, but for the high region.

4. High-frequency region

In the high-frequency region, the presence of the fringes impaired F_0 discrimination of the target even when target and fringes differed in F_0 (Fig. 3). This contrasts markedly with the results in the low- and mid-frequency regions, and replicates Michey and Carlyon's (1998) finding with an ipsilateral interferer. These results were supported by the outcome of two separate two-way ANOVAs, one calculated exclusively for ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes was the same [$F(1,3) = 37.4$, $p < 0.01$], and the other calculated exclusively for ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes differed [$F(1,3) = 26.5$, $p < 0.05$]. In both cases, there were significant main effects of lateralization [$F(4,12) = 15.7$, $p < 0.001$, and $F(4,12) = 15.0$, $p < 0.001$, $HH = 0.75$, for identical and different F_0 s of target and fringes, respectively].

When target and fringes differed in nominal F_0 , the pattern of results was similar to that observed in the low- and mid-frequency regions for identical F_0 s, in that performance improved for all conditions in which the subjective location of the fringes was moved away from that of the target [see Table II(c) for statistical details].

Unexpectedly, for identical nominal F_0 s of target and fringes, a shift in subjective location of the fringes away from that of the target (conditions $i45c45$ and $i45c65$) did not reduce the deleterious effect of the fringes. Performance improved only for purely contralateral presentation [see Table II(d) for planned comparisons]. As in the low- and mid-frequency regions, the purely contralaterally presented fringes still impaired performance, and did so equally for both presentation levels.

The most important findings for the high-frequency region were: (i) F_0 discrimination was impaired by the presence of temporal fringes, even when target and fringes differed in nominal F_0 ; (ii) Only when target and fringes differed in nominal F_0 was the amount of impairment reduced when the fringes were perceived at a location different from that of the target; (iii) Presenting the fringes only contralaterally improved performance relative to ipsilateral presentation for all combinations of target and fringe F_0 s.

However, performance was still significantly impaired.

C. Discussion

In the low- and mid-frequency regions, where harmonics were resolved for at least one F_0 , ipsilaterally presented fringes only raised thresholds in an F_0 discrimination task when they had the same nominal F_0 as the target. In contrast, they raised thresholds in the high region even when their F_0 differed markedly from that of the target. This replicates the finding of Michey and Carlyon (1998).

In all frequency regions, the amount of the impairment could be reduced by presenting the fringes only contralaterally to the target. A shift in perceived location of the fringes away from that of the target (conditions $i45c65$ and $i45c45$) improved performance too, but not as much as purely contralateral presentation. Especially in the high-frequency region, a shift in perceived location only improved performance in combination with a marked difference in the nominal F_0 of target and fringes.

Note that the observed improvements in conditions $i45c65$ and $i45c45$ over condition $i45$ must truly rely on binaural mechanisms. Since in all three conditions exactly the same stimuli were presented to the left ear of the subjects, the superiority of the binaural conditions over the monaural one cannot be attributed to advantages in monaural processing.

The fact that a difference in "ear of entry" and perceived location of target and fringes improved performance seems to indicate that grouping processes might play a role in the present task. There is evidence for the effectiveness of "relative ear of entry" as a cue in sequential (Deutsch, 1979; Bregman, 1990; Hartmann and Johnson, 1991; Shackleton *et al.*, 1994; Darwin *et al.*, 1995) and concurrent sound segregation (Summerfield and Assmann, 1991; Culling and Summerfield, 1995; Hukin and Darwin, 1995; Gockel and Carlyon, 1998). A possible explanation for how streaming might contribute to the observed overall pattern of results is that fringes only impair performance when they are not perceptually segregated from the target. Differences in F_0 , relative ear of entry, and perceived location provided cues for sequential streaming which differed with respect to their effectiveness.

In the low and mid regions, a big difference in F_0 was sufficient to induce segregation and eliminate the interference of the fringes. In the high region, F_0 differences may be much less effective at triggering segregation, due to the absence of concomitant spectral differences (Vliegen *et al.*, 1999). The second-most effective cue for segregation was relative ear of entry. Presenting the fringes only contralaterally always improved performance, but not as much as F_0 differences if concomitant spectral cues were available.

The least effective way of reducing the effect of the fringes was introducing a lateralization due to ILDs. In the high-frequency region, it showed an effect only when the targets and fringes had different nominal F_0 s. This explanation is compatible with other evidence in the literature showing that lateralization through binaural mechanisms, even if insufficient on its own to produce segregation, nevertheless might enhance segregation when combined with other cues.

For example, Shackleton *et al.* (1994) reported that interaural difference cues (ILDs and ITDs) alone did not facilitate recognition of two concurrently presented synthetic vowels. However, in combination with a voice pitch difference of one semitone they markedly increased performance above the improvement found for the pitch difference alone. Similarly, Darwin and Hukin (1998) found that a difference in ITD increased the segregation of a mistuned harmonic from the rest of a vowel.

There could be two reasons for the finding that in the present task relative ear of entry was more effective than perceived location due to ILDs. First, it could be a true effect of relative ear of entry. Second, it could be attributed to an even more lateralized percept in the purely contralateral condition. At present, we cannot decide between these explanations, since listeners noticed a difference in perceived location of the fringes in the two conditions (when asked to pay attention only to the fringes, as done after the experiment proper).

The next experiment investigated the influence of various ITDs of the fringes. Lateralization due to ITDs depends on binaural processing, without any complications due to possible level effects.

III. EXPERIMENT 2: LATERALIZATION DUE TO ITD

A. Rationale and method

Experiment 2 investigated the role of perceived location of temporal fringes—manipulated using ITDs—on F_0 discrimination of two sequentially presented target complexes. If ITDs reduce the deleterious effect of the fringes, this would provide additional support for an effect of perceived location *per se*, beyond a possible effect of relative ear of entry.

As in experiment 1, the target was always presented monaurally at a level of 45 dB SPL per component to the left ear of the subjects. The fringes (if present) were presented at a level of 45 dB SPL to both ears of the subjects. Lateralization of the fringes relative to that of the target was manipulated using the following conditions: (i) leading to the right with an ITD of 1 ms, denoted PERceived CONtra (“percon”); (ii) diotic presentation, denoted PERceived MID (“permid”); and (iii) leading to the left with an ITD of 1 ms, denoted PERceived IPSI (“peripsi”). The whole waveform of the sound at one side was delayed relative to the opposite side, thus providing onset- and ongoing phase-difference cues. An ITD of 1 ms was chosen because informal listening tests with a range of ITDs showed that, for this ITD, the perceived shift in location seemed to be strongest; the shift seemed to level off with further increases in ITD⁴ (see also Mossop and Culling, 1998). Higher thresholds in condition peripsi than in condition percon would unambiguously show an effect of perceived location. Thresholds in condition permid were expected to be intermediate between those in conditions peripsi and percon.

Four subjects were tested in all frequency regions, all of whom took part in experiment 1. In addition, the third author was tested in the mid region in the two extreme conditions (peripsi, percon), to check a possible explanation for his de-

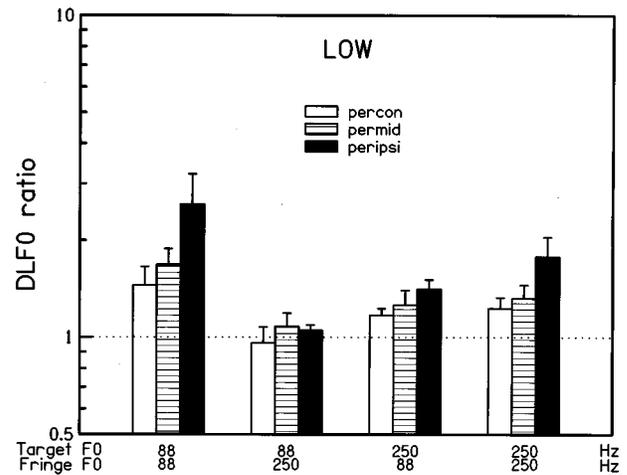


FIG. 4. As Fig. 1, but for experiment 2. The individual bars correspond to conditions (from left to right): leading to the right with an ITD of 1 ms (percon), diotic (permid), leading to the left with an ITD of 1 ms (peripsi).

viant behavior in the first experiment (see Sec. II B 3). However, the data for this subject were excluded from the group data shown below.

B. Results

1. No fringe

As in experiment 1, the overall sizes of the DLF0s in the absence of the fringes depended on (un)resolvability of the individual components in the target complex. They are, however, somewhat smaller than the corresponding values observed in experiment 1 for all three frequency regions, presumably due to practice effects [see Table I(b)].

2. Low-frequency region

Figure 4 shows the results in the presence of the fringes for the low-frequency region. They are plotted in the same way as the results of experiment 1. The individual bars correspond to the conditions (from left to right): percon, permid, and peripsi. The DLF0s increased significantly in the presence of the fringes, but only when target and fringes had the same nominal F_0 . This was supported by the outcome of two separate two-way ANOVAs. One was calculated for ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes were the same [$F(1,3) = 18.14, p < 0.05$] and the other one was calculated for ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes differed. This replicated the finding from experiment 1 that in the low-frequency region, where all components are resolved, fringes with an F_0 different from the nominal target F_0 have no significant effect on DLF0s.

When target and fringes had the same nominal F_0 , thresholds were highest in condition peripsi. Relative to the no-fringe conditions, thresholds increased by a factor of about 2.4 and 2.0 for the 88 and 250-Hz F_0 s, respectively. In conditions permid and percon, where the fringes were lateralized away from the target, their deleterious effect on F_0 discrimination was reduced when compared with condition peripsi. On average, DLF0 ratios were smallest in con-

TABLE III. Planned contrasts between log-transformed ratios of DLF0s measured in experiment 2 in the presence and in the absence of temporal fringes, for condition peripsi versus other lateralization conditions. The significance levels specified were corrected for multiple comparisons. (a) Results in low-frequency region. (b) Results in mid-frequency region. (c) Results in high-frequency region for identical nominal F_0 of target and fringes. (d) Results in high-frequency region for all combinations of target and fringe F_0 .

(a) Low Target and fringes at identical F_0			(b) Mid Target and fringes at identical F_0		
Contrast	$F(1,6)$	p	Contrast	$F(1,6)$	p
peripsi-percon	39.44	<0.05	peripsi-percon	28.35	<0.01
peripsi-permid	23.42	<0.05	peripsi-permid	16.39	<0.05
(c) High Target and fringes at identical F_0			(d) High All combinations of target and fringe F_0		
Contrast	$F(1,6)$	p	Contrast	$F(1,12)$	p
peripsi-percon	48.53	<0.01	peripsi-percon	28.25	<0.01
peripsi-permid	12.39	<0.05	peripsi-permid	8.17	>0.05

dition percon. These results were confirmed by the outcome of the two-way ANOVA calculated for the ratios of DLF0s obtained in conditions with identical nominal target and fringe F_0 . This showed a significant main effect of lateralization [$F(2,6) = 21.65, p < 0.05, HH = 0.61$]. Planned contrasts showed that for both lateralizations of the fringes away from that of the target, performance in F_0 discrimination improved significantly over that in condition peripsi [see Table III(a) for statistical details].

3. Mid-frequency region

In the mid-frequency region (see Fig. 5), results with the fringes were very similar to those observed in the low-frequency region.⁵ The data of the third author, who participated only in the mid region in conditions peripsi and percon, showed the same general pattern as for the other subjects. This supports an explanation of his deviant pattern of results in the binaural fringe conditions in experiment 1 in terms of a negative effect due to an increase in perceived loudness of the fringes.⁶

4. High-frequency region

In the high-frequency region (see Fig. 6), the fringes impaired F_0 discrimination of the target even when target

and fringes differed in F_0 . This was supported by the outcome of two two-way ANOVAs, one calculated for ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes corresponded to each other [$F(1,3) = 22.66, p < 0.05$], and the other, calculated for ratios of DLF0s obtained in conditions where the nominal F_0 of target and fringes differed [$F(1,3) = 12.24, p < 0.05$]. There was a significant main effect of lateralization [$F(2,6) = 24.3, p < 0.01, HH = 0.82$] if target and fringes had identical nominal F_0 , but *no* significant main effect was observed if they differed in F_0 . This was opposite to the pattern observed in experiment 1. The finding of statistically equal performance in all conditions when the target and fringes differed in F_0 was probably due to the fact that the DLF0 ratios were already quite small in condition peripsi for unequal F_0 s.

When target and fringes had the same nominal F_0 , both lateralizations of the fringes away from that of the target led to significant improvement in performance over that in condition peripsi. Table III(c) and (d) show the results of the planned contrasts.

C. Discussion

In the low- and mid-frequency regions, where harmonics were resolved for at least one F_0 , the pattern of results from

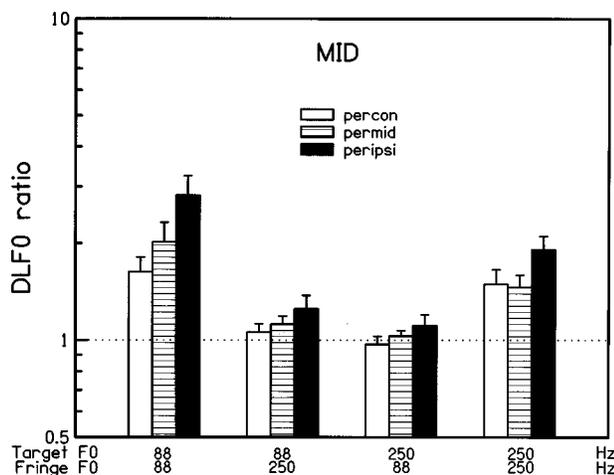


FIG. 5. As Fig. 4, but for the mid region.

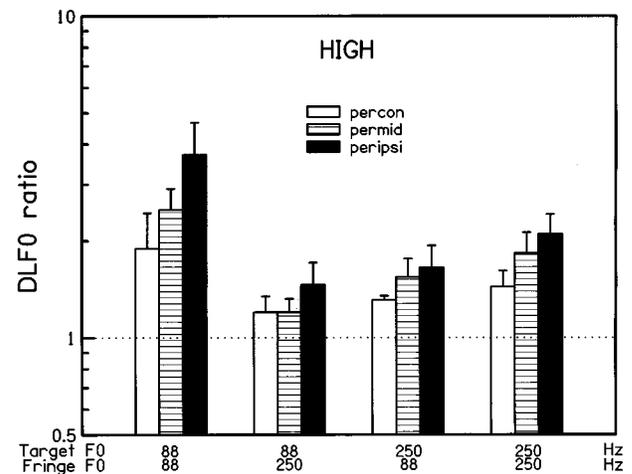


FIG. 6. As Fig. 4, but for the high region.

experiment 2, which used ITDs to vary the perceived location of the fringes, was very similar to that for experiment 1, which used various ILDs. First, the fringes only raised thresholds when they had the same nominal $F0$ as the target. Second, and more importantly, the deleterious effect of the fringes could be reduced by moving the perceived location of the fringes away from that of the target. That is, fringes leading by an ITD of 1 ms at the side contralateral to the monaurally presented target produced less impairment than fringes leading at the side ipsilateral to the target. An improvement was also observed for diotically presented fringes when compared with fringes leading at the ipsilateral side. Note, however, that the negative effect of the fringes in condition *peripsi* appears weaker than in condition *i45* in experiment 1. This could be due to increased practice of the subjects, and/or to the fact that the ipsilateral lateralized image due to ITD was more fuzzy, i.e., more distributed in space, than the image for monaural presentation. However, since thresholds in condition *permid* in experiment 2 were also lower than thresholds for the identical condition *i45c45* in experiment 1, it is quite likely that there was a general improvement due to practice.

In the high-frequency region, fringes raised thresholds even when their $F0$ differed markedly from that of the target. In contrast to experiment 1, subjective location of the fringes away from the target, due to ITDs, significantly reduced $DLF0$ ratios even when the fringes and the target had identical nominal $F0$ s. However, we would not interpret this as evidence that ITDs were more effective than ILDs in reducing the deleterious effect of the fringes in the present task. It is much more likely that the origin of this difference lay in increased practice of subjects which might have led to a more efficient use of the available cues, i.e., the small pitch differences and the subjective difference in location of target and fringes. This interpretation gains support from the fact that $DLF0$ ratios in condition *permid* in experiment 2 were smaller than in the identical condition in experiment 1. Additional support comes from the finding that, in contrast to experiment 1, there was no significant main effect of lateralization in the high-frequency region when target and fringes differed in nominal $F0$. This was probably due to the extreme decrease in $DLF0$ ratios in condition *peripsi* in experiment 2 compared to the $DLF0$ ratios observed in experiment 1 in condition *i45*. This decrease markedly reduced the range of possible improvement due to a shift of perceived location of the fringes away from that of the target.

Apart from the general improvement due to practice effects, the overall pattern of results in experiment 2 was very similar to that observed in experiment 1. Overall, the data provide strong evidence for the role of a truly binaural mechanism partly determining the effect of temporal fringes in $F0$ discrimination. This influence may well be mediated by perceptual segregation. The present findings and interpretation are compatible with evidence in the literature showing that perceived location due to ITDs is an effective cue for sequential perceptual segregation (Hartmann and Johnson, 1991; Darwin and Hukin, 1999). For concurrent sound segregation, it is less effective (see, e.g., Culling and Summerfield, 1995; Hukin and Darwin, 1995), but it can become a

significant cue to segregation in conjunction with other cues (Shackleton *et al.*, 1994; Darwin and Hukin, 1998).

IV. SUMMARY AND CONCLUSIONS

The present study measured $DLF0$ s for complex tones presented with lateralized fringes, which were present before and after but not simultaneously with the target. The fringes were complex tones, filtered into the same frequency region as the target, and had either identical or different nominal $F0$ s as the target. The fringes were lateralized either by ILDs (experiment 1) or by ITDs (experiment 2). In the low- and mid-frequency regions, where harmonics for at least one $F0$ were resolved, the fringes impaired performance only if they had the same nominal $F0$ as the target. In the high-frequency region, where all harmonics were unresolved, they increased thresholds independently of $F0$. The results of the present experiments showed that:

- (1) The deleterious effect of the ipsilateral fringes was reduced most, but was still present, when the fringes were presented purely contralaterally to the monaural target.
- (2) The effect of the ipsilateral fringes was significantly reduced when they were perceived at a location different from that of the target, due either to ILDs or ITDs.
- (3) Besides resolvability, at least one other factor, namely perceived location *per se*, determines the effect of the fringes on $F0$ discrimination of the target.
- (4) The findings are compatible with the assumption that fringes only impair performance when they are not perceptually segregated from the target. Segregation is not "all or none," and depends on a combination of available cues, which may differ in effectiveness.
- (5) In the high region where all components were unresolved, $F0$ differences may be much less effective at triggering segregation than in the low and mid regions, where components for at least one $F0$ were resolved.

The results are consistent with the idea that temporal fringes produce interference in $F0$ discrimination when the listeners have difficulty segregating the target from the fringes. A difference between the perceived location of the target and that of the fringes enhances segregation of the sequentially presented stimuli.

ACKNOWLEDGMENTS

The authors wish to thank Brian Moore, Chris Darwin, an anonymous reviewer, and Wesley Grantham for helpful comments on a previous version of this manuscript. This work was supported by the Wellcome Trust.

¹Repeated-measurements ANOVAs were run on the logarithms of the ratios of the geometric mean $DLF0$ s in the presence of the fringes and the geometric mean $DLF0$ s without fringes.

²Throughout the paper, if appropriate, the Huynh-Feldt correction was applied to the degrees of freedom (see, e.g., Howell, 1997).

³Results of ANOVAs on log-transformed ratios of $DLF0$ s measured in the mid-frequency region in experiment 1 were as follows: (a) Three-way ANOVA (target- $F0 \times$ fringe- $F0 \times$ lateralization); significant interaction between target- $F0$ and fringe- $F0$ [$F(1,4) = 33.01, p < 0.01$]. (b) Two-way ANOVA (nominal target- $F0$ differs from fringe- $F0$); no effect of pres-

ence of fringes. (c) Two-way ANOVA (nominal target- $F0$ equals fringe- $F0$): (i) significant effect of presence of fringes [$F(1,4)=93.5, p<0.01$]; (ii) significant main effect of lateralization [$F(4,16)=22.42, p<0.001$]. (d) Two-way ANOVA (nominal target- $F0$ equals fringe- $F0$, purely contralaterally presented fringes): (i) significant effect of presence of fringes [$F(1,4)=89.39, p<0.01$]; (ii) no significant level effect. The results of the planned contrasts are given in Table II(b).

⁴More formal experiments were run with three subjects out of the total of five participating in experiment 2, to investigate the perceived location of the fringes presented in isolation as a function of ITD. In a 2I-2AFC paradigm, subjects had to indicate which of the two successive sounds they perceived as more to the right. The ITD was varied randomly for each and every stimulus presentation in the range from 0 to 1 ms in 0.2-ms steps. The ITDs of the two sounds always had the same sign, and thus they were both perceived in the same hemifield. Each hemifield was tested in separate blocks of trials. A "hit" was defined as judging the stimulus with the greater ITD as more to the right (when the right ear led) or more to the left (when the left ear led). In this way, a whole set of d' scores was obtained for each reference ITD (defining reference ITD as the smallest absolute value of the two ITDs used in a specific trial); i.e., five d' scores resulted for the reference ITD of 0 ms in combination with comparison ITDs of 0.2, 0.4, 0.6, 0.8, and 1 ms; four d' scores resulted for the reference ITD of 0.2 ms in combination with comparison ITDs of 0.4, 0.6, 0.8, and 1 ms, etc. Average performance across subjects always improved monotonically with increasing comparison ITD, indicating that perceived location changed monotonically with increasing ITD, including the case where the reference and comparison ITDs were 0.8 and 1 ms, respectively. Performance was better than chance in all frequency regions and for all $F0$ s. However, performance levels were dependent on both $F0$ and frequency region. In the low region, performance was generally very good, although somewhat better for the 88-Hz $F0$ than for the 250-Hz $F0$. In the mid region, performance was poorest, especially for the 250-Hz $F0$. In the high region, performance was nearly as good as in the low region, especially for the lower $F0$. These results were expected, since in the low-frequency region, where all components were resolved, components lay well within the frequency range in which ITD discrimination performance for pure tones is known to be good (see, e.g., Blauert, 1983; Moore, 1997). In the mid region, resolved components of the 250-Hz $F0$ fringe lay only partly below the frequency limit of 1500 Hz for which ITD discrimination performance for pure tones was reported to be around chance (Moore, 1997). In the high region, where all components were unresolved, discrimination of ITDs presumably was based on the interaural delay of the envelopes of the signals (Henning, 1974).

⁵Results of ANOVAs on log-transformed ratios of $DLF0$ s measured in the mid-frequency region in experiment 2 were as follows: (a) Three-way ANOVA (target- $F0 \times$ fringe- $F0 \times$ lateralization); significant interaction between target- $F0$ and fringe- $F0$ [$F(1,3)=136.1, p<0.01$]. (b) Two-way ANOVA (nominal target- $F0$ differs from fringe- $F0$); no significant effect of presence of fringes. (c) Two-way ANOVA (nominal target- $F0$ equals fringe- $F0$): (i) significant effect of presence of fringes [$F(1,3)=436.09, p<0.001$]; (ii) significant main effect of lateralization [$F(2,6)=15.46, p<0.01$].

⁶In experiment 1, performance in condition $c45$ relative to the no-fringe condition was more impaired for the third author than for the other subjects. Additionally, his data showed a distinct level effect for the purely contralateral fringes, i.e., the more intense fringes led to more impairment. His unusually high susceptibility for contralateral fringes, and especially for relatively loud ones, presumably added to the deleterious effect of the ipsilateral fringes in the binaural conditions. This in turn would have prevented any advantage due to the shift of perceived location from emerging in the data. That would explain why his thresholds were highest in condition $i45c65$ and second highest in $i45c45$. In experiment 2, all conditions involved binaural presentation of the fringes at the same sound-pressure level. If the above explanation was correct, one would expect no differences between the pattern of the third author's data and that of the other subjects. This was confirmed.

Anstis, S., and Saida, S. (1985). "Adaptation to auditory streaming of frequency-modulated tones," *J. Exp. Psychol. Hum. Percept. Perform.* **11**, 257-271.

Assmann, P., and Summerfield, Q. (1990). "Modelling the perception of

concurrent vowels: Vowels with different fundamental frequencies," *J. Acoust. Soc. Am.* **88**, 680-697.

Beerends, J. G., and Houtsma, A. J. M. (1989). "Pitch identification of simultaneous diotic and dichotic two-tone complexes," *J. Acoust. Soc. Am.* **85**, 813-819.

Blauert, J. (1983). *Spatial Hearing* (MIT Press, Cambridge, MA).

Bregman, A. S. (1978). "Auditory streaming is cumulative," *J. Exp. Psychol. Hum. Percept. Perform.* **4**, 380-387.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA).

Bregman, A. S., and Ahad, P. A. (1995) (Compact Disc). "Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound" (Department of Psychology, Auditory Perception Laboratory, McGill University).

Bregman, A. S., and Campbell, J. (1971). "Primary auditory stream segregation and the perception of order in rapid sequences of tones," *J. Exp. Psychol.* **89**, 244-249.

Brokx, J. P. L., and Nootboom, S. G. (1982). "Intonation and the perceptual separation of simultaneous voices," *J. Phonetics* **10**, 23-36.

Carlyon, R. P. (1994a). "Detecting mistuning in the presence of synchronous and asynchronous interfering sounds," *J. Acoust. Soc. Am.* **95**, 2622-2630.

Carlyon, R. P. (1994b). "Detecting pitch-pulse asynchronies and differences in fundamental frequency," *J. Acoust. Soc. Am.* **95**, 968-979.

Carlyon, R. P. (1999a). "Encoding the fundamental frequency of a complex tone in the presence of a spectrally overlapping masker," *J. Acoust. Soc. Am.* **99**, 517-524.

Carlyon, R. P. (1996b). "Masker asynchrony impairs the fundamental-frequency discrimination of unresolved harmonics," *J. Acoust. Soc. Am.* **99**, 525-533.

Carlyon, R. P., Demany, L., and Semal, C. (1992). "Detection of across-frequency differences in fundamental frequency," *J. Acoust. Soc. Am.* **91**, 279-292.

Carlyon, R. P., and Shackleton, T. M. (1994). "Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms?," *J. Acoust. Soc. Am.* **95**, 3541-3554.

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.

Darwin, C. J., and Ciocca, V. (1992). "Grouping in pitch perception: Effects of onset asynchrony and ear of presentation of a mistuned component," *J. Acoust. Soc. Am.* **91**, 3381-3390.

Darwin, C. J., Hukin, R. W., and Al-Khatib, B. Y. (1995). "Grouping in pitch perception: Evidence for sequential constraints," *J. Acoust. Soc. Am.* **98**, 880-885.

Darwin, C. J., and Hukin, R. W. (1998). "Perceptual segregation of a harmonic from a vowel by interaural time difference in conjunction with mistuning and onset asynchrony," *J. Acoust. Soc. Am.* **103**, 1080-1084.

Darwin, C. J., and Hukin, R. W. (1999). "Auditory objects of attention: The role of interaural time-differences," *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 617-629.

Deutsch, D. (1979). "Binaural integration of melodic patterns," *Percept. Psychophys.* **25**, 399-405.

Divenyi, P. L., and Hirsh, I. J. (1975). "The effect of blanking on the identification of temporal order in three-tone sequences," *Percept. Psychophys.* **17**, 246-252.

Gockel, H., and Carlyon, R. P. (1998). "Effects of ear of entry and perceived location of synchronous and asynchronous components on mistuning detection," *J. Acoust. Soc. Am.* **104**, 3534-3545.

Hartmann, W. M. (1996). "Pitch, periodicity and auditory organization," *J. Acoust. Soc. Am.* **100**, 3491-3502.

Hartmann, W. M., and Johnson, D. (1991). "Stream segregation and peripheral channeling," *Music Percept.* **9**, 155-184.

Hartmann, W. M., McAdams, S., and Smith, B. K. (1990). "Hearing a mistuned harmonic in an otherwise periodic complex tone," *J. Acoust. Soc. Am.* **88**, 1712-1724.

Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex waveforms," *J. Acoust. Soc. Am.* **55**, 84-90.

Hoekstra, A. (1979). "Frequency discrimination and frequency analysis in hearing," Ph.D. thesis, Institute of Audiology, University Hospital, Groningen, The Netherlands.

Houtsma, A. J. M., and Goldstein, J. L. (1972). "The central origin of the pitch of pure tones: Evidence from musical interval recognition," *J. Acoust. Soc. Am.* **51**, 520-529.

- Howell, D. C. (1997). *Statistical Methods for Psychology* (Duxbury, Belmont, CA), 4th ed.
- Hukin, R. W., and Darwin, C. J. (1995). "Effects of contralateral presentation and of interaural time differences in segregating a harmonic from a vowel," *J. Acoust. Soc. Am.* **98**, 1380–1387.
- Kelly, W. J., and Watson, C. S. (1986). "Stimulus-based limitations on the discrimination between different temporal orders of tones," *J. Acoust. Soc. Am.* **79**, 1934–1938.
- Kubovy, M., and Howard, F. P. (1976). "Persistence of a pitch-segregating echoic memory," *J. Exp. Psychol. Hum. Percept. Perform.* **2**, 531–537.
- Lee, J., and Green, D. M. (1994). "Detection of a mistuned component in a harmonic complex," *J. Acoust. Soc. Am.* **96**, 716–725.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Massaro, D. W. (1975). "Backward recognition masking," *J. Acoust. Soc. Am.* **58**, 1059–1065.
- Micheyl, C., and Carlyon, R. P. (1998). "Effects of temporal fringes on fundamental-frequency discrimination," *J. Acoust. Soc. Am.* **104**, 3006–3018.
- Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing* (Academic, London), 4th ed.
- Moore, B. C. J., and Glasberg, B. R. (1990). "Frequency discrimination of complex tones with overlapping and nonoverlapping harmonics," *J. Acoust. Soc. Am.* **87**, 2163–2177.
- Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1986). "Thresholds for hearing mistuned partials as separate tones in harmonic complexes," *J. Acoust. Soc. Am.* **80**, 479–483.
- Moore, B. C. J., Glasberg, B. R., and Shailer, M. J. (1984). "Frequency and intensity difference limens for harmonics within complex tones," *J. Acoust. Soc. Am.* **75**, 550–561.
- Mossop, J. E., and Culling, J. F. (1998). "Lateralization of large interaural delays," *J. Acoust. Soc. Am.* **104**, 1574–1579.
- Plack, C. J., and Carlyon, R. P. (1995). "Differences in frequency modulation detection and fundamental frequency discrimination between complex tones consisting of resolved and unresolved harmonics," *J. Acoust. Soc. Am.* **98**, 1355–1364.
- Rogers, W. L., and Bregman, A. S. (1998). "Cumulation of the tendency to segregate auditory streams: Resetting by changes in location and loudness," *Percept. Psychophys.* **60**, 1216–1227.
- Scheffers, M. T. M. (1983). "Sifting vowels: auditory pitch analysis and sound segregation," Doctoral dissertation, University of Groningen, The Netherlands.
- Schouten, J. F. (1940). "The residue and the mechanism of hearing," *Proceedings of the Koninklijke Akademie van Wetenschap*, Vol. 43, pp. 991–999.
- Shackleton, T. M., and Carlyon, R. P. (1994). "The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination," *J. Acoust. Soc. Am.* **95**, 3529–3540.
- Shackleton, T. M., Meddis, R., and Hewitt, M. J. (1994). "The role of binaural and fundamental frequency difference cues in the identification of concurrently presented vowels," *Q. J. Exp. Psychol.* **47A**, 545–563.
- Summerfield, Q., and Assmann, P. F. (1991). "Perception of concurrent vowels: Effects of harmonic misalignment and pitch-period asynchrony," *J. Acoust. Soc. Am.* **89**, 1364–1377.
- van Noorden, L.P.A.S. (1975). "Temporal coherence in the perception of tone sequences," Doctoral dissertation, Technische Hogeschool Eindhoven, Eindhoven, The Netherlands.
- Vliegen, J., Moore, B. C. J., and Oxenham, A. J. (1999). "The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task," *J. Acoust. Soc. Am.* **106**, 938–945.
- Vliegen, J., and Oxenham, A. J. (1999). "Sequential stream segregation in the absence of spectral cues," *J. Acoust. Soc. Am.* **105**, 339–346.

The influence of musical training on the perception of sequentially presented mistuned harmonics

Edward M. Burns^{a)}

University of Washington, 1417 NE 42nd Street, Seattle, Washington 98105-6246

Adrianus J. M. Houtsma

IPO, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands

(Received 5 May 1999; revised 4 January 1999; accepted 2 August 1999)

The question of whether musical scales have developed from a processing advantage for frequency ratios based on small integers, i.e., ratios derived from relationships among harmonically related tones, is widely debated in musicology and music perception. In the extreme position, this processing advantage for these so-called “natural intervals” is assumed to be inherent, and to apply to sequentially presented tones. If this is the case, evidence for this processing advantage should show up in psychoacoustic experiments using listeners from the general population. This paper reports on replications and extensions of two studies from the literature. One [Lee and Green, *J. Acoust. Soc. Am.* **96**, 716–725 (1994)] suggests that listeners from the general population can in fact determine whether sequentially presented tones are harmonically related. The other study [Houtgast, *J. Acoust. Soc. Am.* **60**, 405–409 (1976)] is interpreted in different terms, but could be confounded by such an ability. The results of the replications and extensions, using listeners of known relative pitch proficiency, are consistent with the idea that only trained musicians can reliably determine whether sequentially presented tones are harmonically related. © 1999 *Acoustical Society of America*. [S0001-4966(99)06111-1]

PACS numbers: 43.66.Hg [JWH]

INTRODUCTION

Most trained musicians have the ability, termed relative pitch, to identify the musical interval (frequency ratio) between two sequentially presented tones. Those with the best relative pitch are able to identify intervals with an accuracy of about one-third semitone, or 2% (Burns and Campbell, 1994). Since many of the frequency ratios which correspond to the standardized intervals (such as octaves, fifths, and thirds) comprising the musical scales of most cultures are equal to, or nearly equal to, small integer ratios, they also correspond to the frequency ratios (or octave transpositions thereof) among harmonics of a complex tone. In fact, there is considerable debate over whether musical scales have developed from an innate processing advantage for these so-called “natural” intervals (e.g., Burns, 1999). If such an advantage exists, the ability of relative pitch merely implies that musicians have learned labels for these harmonic relationships for which all people have innate predisposition. In this case, this predisposition for harmonic relationships should show up in psychoacoustic experiments which do not require the specific labeling ability acquired through musical training. In this report we focus on two studies in the literature, one of which suggests that the ability to tell whether a (sequentially presented) tone is harmonically related to a reference tone is a universal ability, and the other of which has been interpreted in an entirely different manner, but whose interpretation could also be confounded by this ability.

The latter study was reported more than 20 years ago by Houtgast (1976). He found that most subjects could success-

fully discriminate between the pitches formed by the missing fundamental frequency of a six-tone harmonic complex and the “fundamental frequency” of a single sinusoid not contained in the six-tone complex, provided that there was enough background noise to keep the tone signals barely audible. The observation that under those conditions subjects could consistently detect a small up or down step in fundamental frequency led to the conclusion that people can hear the missing fundamental pitch from even one single harmonic, provided the S/N ratio is sufficiently low. However, listeners with relative pitch could presumably perform this task for the single harmonics simply by noting if the mistuned harmonics comprised sharp or flat standard musical intervals relative to the fundamental of the complex. If the ability to tell if sequentially presented tones are mistuned from harmonicity is universal, presumably even musically naive listeners could perform the task without hearing a fundamental pitch.

The other experiment was reported by Lee and Green (1994). Measuring subjects’ sensitivity to a mistuned harmonic of a complex tone, they removed one partial from a harmonic tone complex and presented the remainder monaurally to a listener, followed by the isolated partial. This partial was either in tune (unaltered), or slightly changed in frequency (mistuned). In a forced-choice paradigm subjects had to identify the stimulus sequence with the mistuned partial. A rather constant mistuning threshold of 1% was found for harmonics 1–7, and an increasing threshold for harmonics 8 and beyond. Although Lee and Green mention the musical training of their subjects, they do not suggest that this training was crucial to the outcome of the experiment, and the implication is that these mistuning thresholds are gener-

^{a)}Electronic mail: pwa@u.washington.edu

alizable to the population of normal-hearing listeners.

In this report, data are presented from two replications and extensions of Houtgast's experiment and a replication of the Lee and Green experiment. In all cases, subjects with known relative pitch abilities are used in order to determine the extent to which the results are dependent on this ability.

I. METHOD AND EXPERIMENTS

Since both the Houtgast (1976) and Lee and Green (1994) experiments are published in this journal, and the replications are essentially exact, only synopses of the protocols of these experiments are provided and important differences between our protocols and those of the original experiments are noted.

A. First Houtgast replication

The basic paradigm of Houtgast's experiment was a two-alternative forced choice (2AFC) up/down pitch discrimination task between a fixed complex-tone reference with a fundamental frequency of 200 Hz and a comparison signal with a 2% higher or lower fundamental and a variable harmonic composition. Harmonic composition (three, two, or one harmonics), harmonic order, and presence or absence of background noise were independent variables.

Two aspects of Houtgast's procedure appear to be crucial for eliciting a low (missing fundamental) pitch sensation from a single harmonic. The tones must be presented at a low sensation level (SL), accomplished by presenting the tones in the presence of a masking noise. The listener must also be cued to listen in the region of the fundamental, accomplished by a special trial sequence. On every seventh trial the comparison tone was a three-(adjacent) harmonic complex, on the next three trials the comparison tones were two-harmonic complexes, and on the following three trials they were three single harmonics. The cueing is a logical necessity since a pure tone does not, by itself, have a uniquely defined missing fundamental.

The stimuli and procedures employed for the replication were the same as those used by Houtgast with the following exceptions: Houtgast used a six-harmonic standard complex (with no harmonics in common with the three harmonics treated in a block), sine-phase addition of harmonics, $\pm 2\%$ change in fundamental frequency for the comparison tone, and two dummy trials for practice at the beginning of each (70 trial) stimulus block; we used a three-tone standard complex, random-phase addition of harmonics, $\pm 3\%$ change of fundamental frequency, and no dummy trials. We specifically instructed listeners to judge whether the comparison tone was higher or lower than the standard, whereas Houtgast's listeners were "suggested to concentrate on pitch of second tone relative to first" but instructions emphasized minimization of incorrect responses by "any criterion or strategy."

In addition to the replication, the listeners performed a "simple mistuning" experiment, where the standard tone in a 2AFC task was a sinusoid with a frequency of 200 Hz and the comparison tone was also a sinusoid, the m th harmonic of 200 Hz, mistuned by either plus or minus 3%, with m taken randomly between 2 and 12 on each trial. Since there

is no cueing, the only way listeners can perform above chance level in this task is if they are able to determine whether the harmonics are mistuned high or low relative to the 200-Hz standard. If performance in this task is the same (as a function of harmonic number) as that in the Houtgast single-harmonic condition, it would suggest that the listeners used the same relative pitch cues in both experiments.

Six subjects participated in the experiments. They were divided into two groups. Group 1 consisted of three listeners who had participated in experiments requiring good relative pitch, and had demonstrated their ability to perform at essentially 100% correct in a melodic musical interval identification task. Group 2 consisted of three listeners who reported having only minimal musical training, did not play an instrument, and did not possess relative pitch.

The results of both the original Houtgast experiment and the replication are presented in Fig. 1. Houtgast's results, shown in the right column, are based on three subjects, 50 trials per harmonic condition per subject, and a fundamental frequency difference of 2%. The replication results, shown in the left column, are based on 20 trials per harmonic condition per subject, and a fundamental frequency difference of 3%. The essential features of Houtgast's results were that listeners performed better for the with-noise (low SL) condition, and in particular performed well above chance for the single harmonics up to 10, but performed essentially at chance for the single harmonics in the absence of noise. These results, along with the subjective impressions of the listeners, suggest that the low SL condition enhances the propensity to hear synthetically, i.e., to hear a "low" pitch corresponding to the fundamental rather than pitches corresponding to the individual harmonics, to the extent that the subject hear a low pitch even for a single harmonic.

The results of the replication do not fully support this interpretation. For the single harmonic case, the performance of the musicians, group 1, for both the with-noise and without-noise conditions, was essentially identical to that of Houtgast's subjects in the with-noise condition, whereas the performance of the nonmusicians, group 2, was similar to the performance of Houtgast's subjects in the without-noise condition, basically at chance level for harmonics higher than 4.

The results of group 1 in the simple-mistuning experiment, where the listeners had to determine if a mistuned harmonic of 200 Hz was mistuned high or low relative to a pure tone of 200 Hz, are shown in Fig. 2. Again, performance patterns are similar to performance in the Houtgast replication, and are basically independent of the presence of noise, which suggests that the same discrimination strategy was used in all cases. These results verify that listeners who possess good relative pitch can determine if a pure tone is mistuned from harmonicity relative to a reference fundamental in a procedure where they could not be hearing any missing fundamental pitch cue. None of the listeners in group 2 could perform above chance in this task (results not shown), but we did not attempt any extensive training.

In light of the results from the simple-mistuning experiment, the simplest interpretation of the results from the Houtgast replication is that neither subject group heard a pitch corresponding to the fundamental, in either noise con-

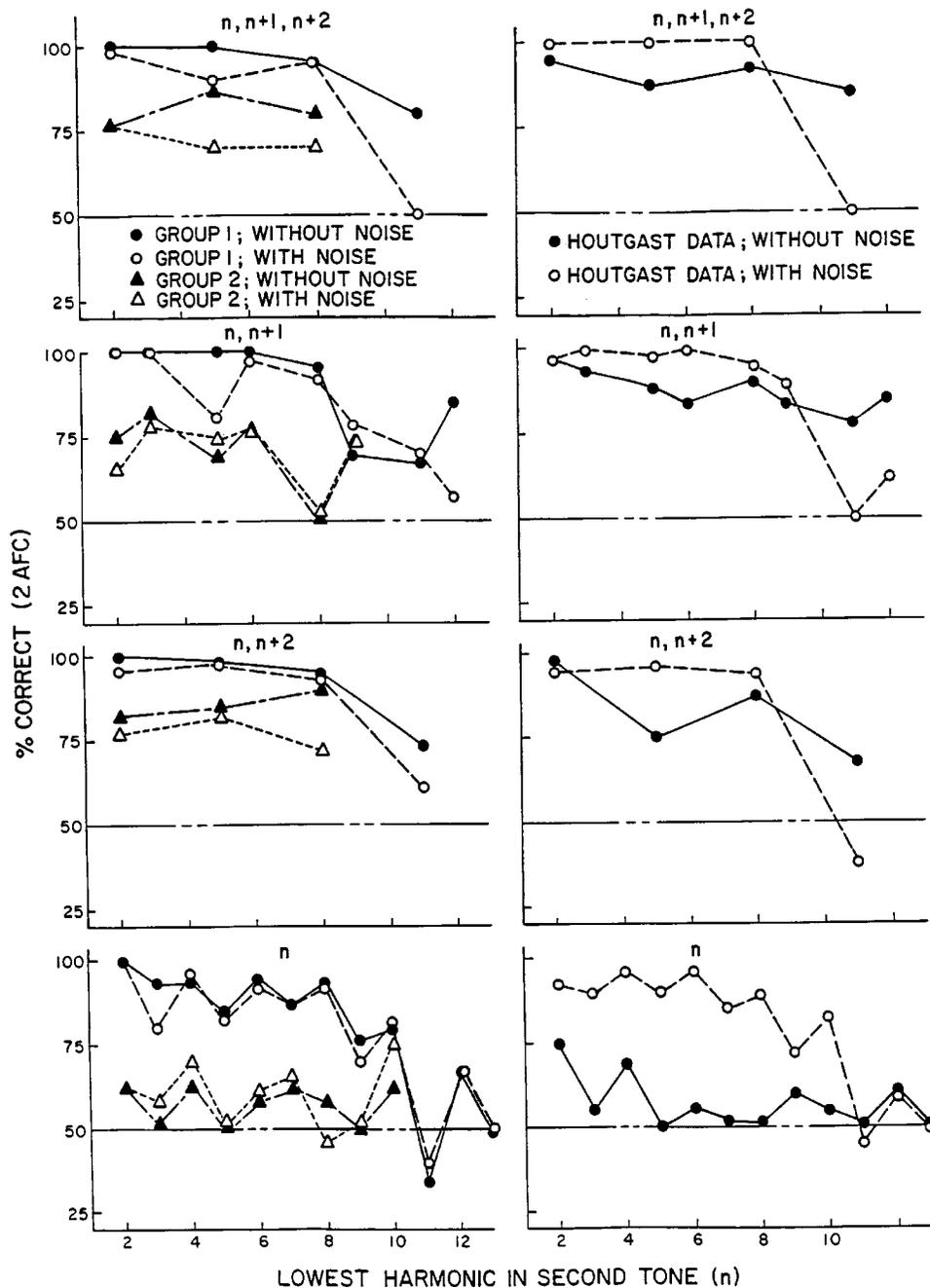


FIG. 1. The right column shows the results of the original Houtgast (1976) study, and the left column the first Houtgast replication. The %-correct scores represent average scores for three subjects in a 2AFC fixed-frequency-increment/decrement task where the comparison tone was composed of either one, two, or three harmonics, as indicated, and for with- and without noise conditions as indicated. The value along the abscissa is the lowest harmonic of the comparison tone. Group 1 consisted of three musicians with good relative pitch, group 2 consisted of three non-musicians.

dition, but that the musicians could perform the task based on relative pitch judgments. However, despite the fact that their performance was virtually identical in the with- and without-noise conditions in the Houtgast replication, two of the three musicians (the two authors) claimed that they did in fact hear a low pitch in the with-noise condition and that they used a different strategy in this condition than they did in the without-noise condition.

B. Second Houtgast replication

Because of the somewhat discrepant findings of Houtgast (1976) and our replication, it was decided to repeat the replication. Whereas in the previous replication there had been some slight differences in the stimuli between the replication and the original experiment, in this case both the stimuli and experimental protocol of the original Houtgast

experiment were repeated exactly with the exception that the difference between fundamental frequency of the comparison tones was 6 Hz (3%). The simple mistuning experiment was also repeated, the only difference between this version and the previous version being that, in a single block of trials, comparison-tone harmonics were limited to one of three adjacent harmonics, as in the Houtgast protocol.

The six subjects were divided into three groups of two subjects each, based on their performance on a relative pitch screening test in which they had to identify ten presentations each of the 12 equal-tempered melodic intervals from minor 2nd through octave, presented randomly, and with the first tone of each interval randomized over a range of 261.6 Hz, ± 200 cents. The two excellent relative pitch (ERP) listeners both scored perfectly, the two mediocre relative pitch (MRP) listeners scored 76% and 68% correct, and the two no rela-

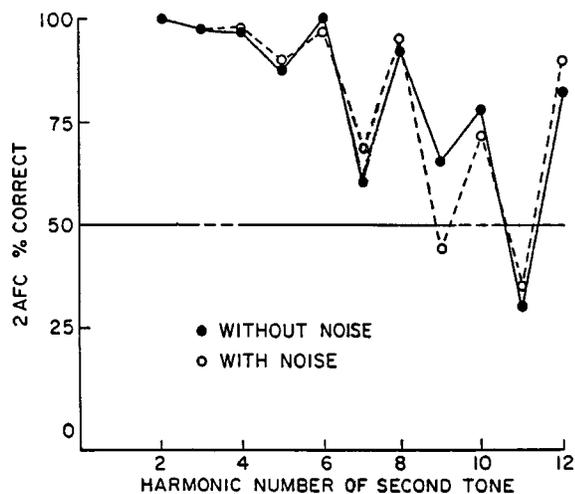


FIG. 2. Scores for group 2 subjects in the "simple mistuning" task where they determined if a randomly selected harmonic of 200 Hz, mistuned by $\pm 3\%$ and presented sequentially to a 200-Hz pure tone, was mistuned "high" or "low."

tive pitch (NRP) listeners could not identify any intervals consistently.

The results for both the Houtgast procedure and the simple-mistuning procedure are shown in Fig. 3, with the ERP, MRP, and NRP listeners shown in the top, middle, and bottom rows, respectively. Performance in the simple mistuning procedure correlates with the degree of relative pitch ability and is essentially independent of noise background: ERP listeners show nearly perfect performance for harmonics 2–10, with slight exceptions at the 7th and 9th harmonics; MRP listeners also generally show well above chance performance for most harmonics through 10; the NRP listeners, with the exception of harmonics 2–4 for NRP1, show chance level, or marginally above chance level, performance for all harmonics. Since in this version of the simple mistuning procedure only three harmonics were tested in a block of trials, only six comparison tone frequencies were presented per block. Nonpossessors of absolute pitch adopting an absolute identification strategy could thus be expected to achieve above chance performance (Pollack, 1952), which may account for the marginally above chance performance of the NRP listeners. Listeners using this strategy would not be expected to achieve near perfect performance, and performance would not be a function of frequency (harmonic number). Consequently, this strategy cannot explain the performance of the ERP and MRP listeners.

Performance in the Houtgast protocol is much less consistent, and ranges from essentially perfect performance through harmonic 10 for both with- and without-noise conditions for ERP1, to essentially chance performance for both noise conditions for NRP2. Only MRP2 shows performance similar to the typical Houtgast subject: near perfect, with the exception of some odd harmonics, in the with-noise case, and close to chance for harmonics greater than 4 in the without-noise case. Excepting MRP2, the most parsimonious explanation would seemingly be that the listeners were using, or trying to use, relative pitch cues in both noise conditions of the Houtgast protocol. However, as was the case with the first replication, several of the subjects insisted they

heard a low pitch for all trials in the with-noise condition and used this to perform the task, but did not hear it for some of the trials in the without-noise condition. If performance is compared for both noise conditions of both the Houtgast and simple-mistuning protocols, whenever there is a large difference in performance between noise conditions it is always in the Houtgast procedure, with performance being worse in the without-noise condition. There is no obvious reason why a high sensation level would degrade relative pitch cues for the Houtgast protocol but not for the simple mistuning protocol, nor is there any reason why a low sensation level would improve relative pitch cues for the Houtgast protocol but not for the simple mistuning protocol.

C. Lee and Green replication

The fact that some subjects could not do the simple mistuning experiments performed in conjunction with the Houtgast replications suggests that the ability to tell if sequentially presented tones are harmonically related is not universal; performance in this task was directly correlated with relative pitch competence. However, it is possible that the poor performance of the nonmusicians in the simple mistuning experiment is related to the experimental protocol and that a different protocol might be more successful at tapping such an ability. In fact, the protocol used by Lee and Green (1994) differed substantially from the simple mistuning task. The latter is essentially a single- (listening) interval task, the two sequentially presented tones comprise a single stimulus which the listeners must judge as being mistuned high or low relative to some internal standard. Lee and Green used a two- (listening) interval task wherein the subjects had to choose which of a pair of sequentially presented tones contained the mistuned harmonic. They also used an adaptive procedure where the degree of mistuning varied as a function of performance and only tested a single harmonic in a block of trials. This protocol may be more effective in tapping a general ability to tell if sequentially presented tones are harmonically related. It was therefore decided to replicate the Lee and Green experiment with the same subjects used in the second Houtgast replication.

The replication used the exact procedures of Lee and Green, with minor exceptions in the adaptive procedure: their criterion for ending a trial block was a fixed number of trials (60), threshold estimates from a trial block were based on the last even number of reversals, excluding the first three, and thresholds were based on eight 60-trial blocks. In our study, trial blocks were terminated after a fixed number of reversals (usually ten), threshold estimates from a trial block based on (usually) the last six reversals, and thresholds were based on at least eight blocks averaging about 50 trials each.

The mistuning thresholds for four of the six subjects are shown in Fig. 4. Also shown are the mean thresholds of the three subjects of Lee and Green. It is clear that only the two ERP subjects have thresholds which are in agreement with those of the Lee and Green subjects over the entire range of harmonic numbers, and even in this case ERP2 has discrepant results for harmonics 7 and 9 (the largest threshold

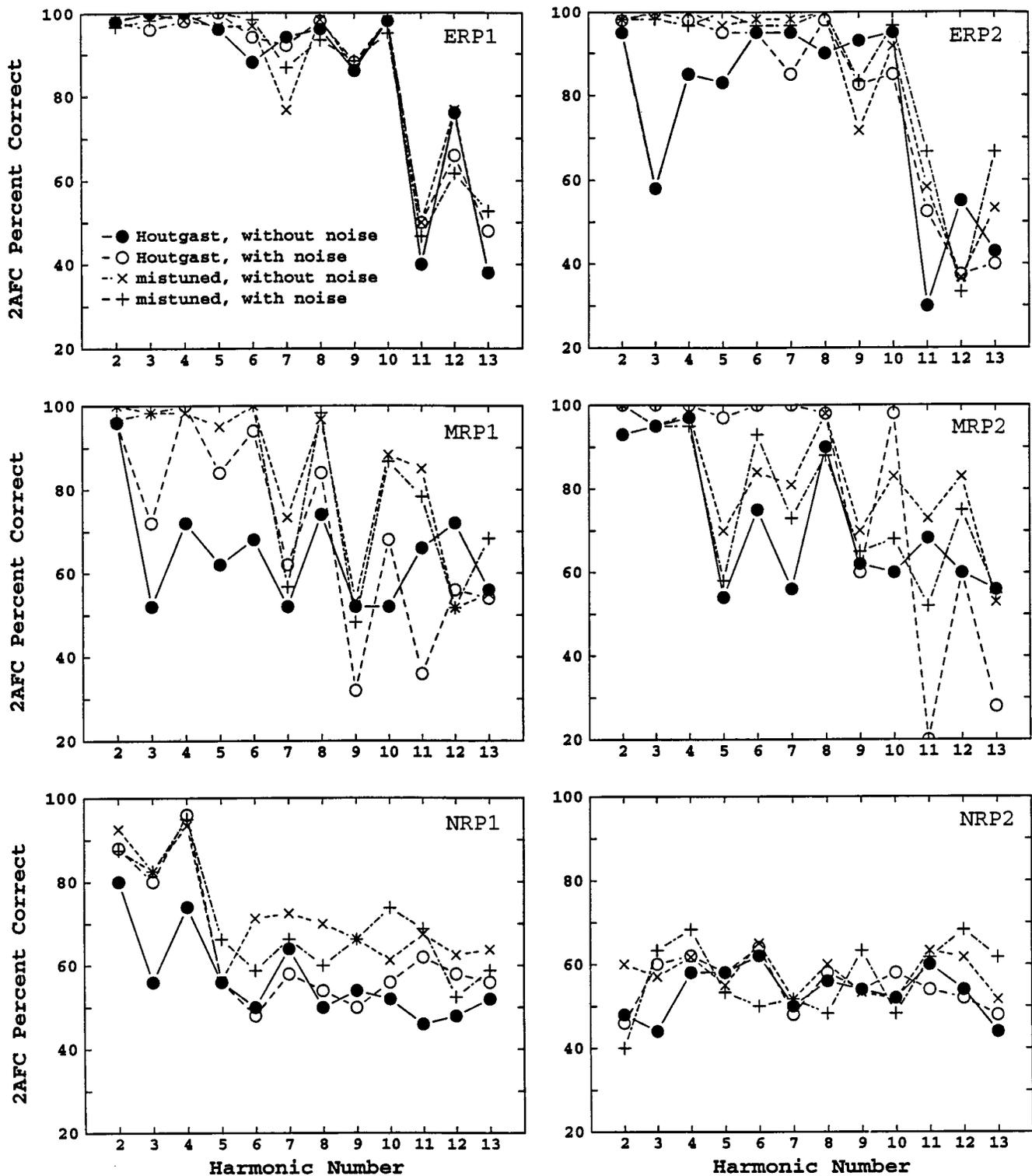


FIG. 3. Scores for the six subjects in the second Houtgast replication and in the second "simple mistuning" experiment. The three rows from top to bottom contain scores for subjects with excellent (ERP), mediocre (MRP), and no (NRP) relative pitch, respectively.

shown by any of the three Lee and Green subjects was 58 Hz for harmonic 9; harmonic 10 was not tested). Subjects MRP1 and MRP2 showed comparable results for harmonics 6 and below, but could not perform the task consistently for harmonics 7 and above. For subject NRP1, consistent threshold estimates could only be obtained for harmonics 2–4, and only for harmonic 2 was the threshold estimate in agreement with Lee and Green's result. Subject NRP1 could not per-

form the task consistently for any harmonic number, despite almost 2000 total trials. She also could not perform a non-adaptive version (i.e., a fixed-mistuning over a block of trials) of the task, for any amount of mistuning. This is consistent with her results in the simple mistuning and Houtgast protocol above, where she performed at near chance level for all conditions. Thus the performance of these subjects is completely consistent with their performance in the simple

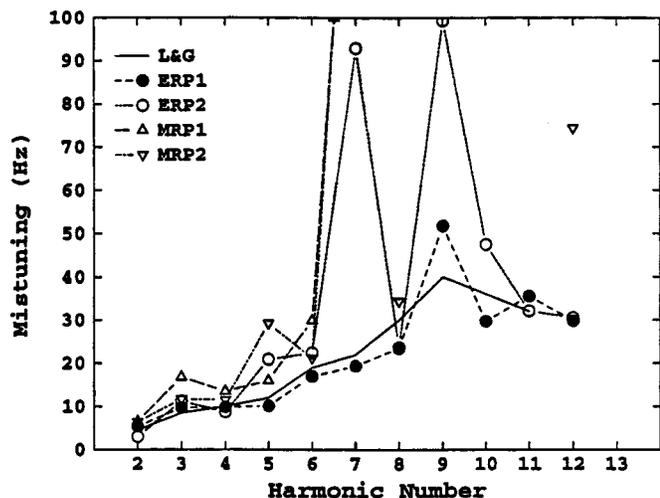


FIG. 4. Thresholds for harmonic mistuning as a function of harmonic number, determined using the 2I-2AFC "successive-monotic" procedure of Lee and Green (1994). The lines with symbols are the individual results for six subjects, as indicated. The average thresholds for three subjects from Lee and Green are shown by the solid line without symbols; no thresholds were obtained by Lee and Green for harmonic no. 10.

mistuning task and is correlated with relative pitch proficiency.

II. DISCUSSION

The results of both simple mistuning experiments and the Lee and Green replication show conclusively that some listeners can tell if sequentially presented pure tones are harmonically related. The results of the simple mistuning experiment, in particular, show that listeners with good relative pitch proficiency can use relative pitch cues even for a broad range of interval widths when multiple octave transpositions are necessary, and for a task where overt interval labeling is not required. However, there was no evidence suggesting that this is a universal ability. In all three experiments performance was strongly correlated with relative pitch proficiency. This finding is consistent with the vast majority of the literature (e.g., Allen, 1967; Kallman, 1982; Terhardt *et al.*, 1986; Thurlow and Erchul, 1977) which suggests that any affinity for harmonic relationships for sequential tones is based on exposure to music and a strong affinity is found only in subjects with extensive musical training. The finding is not consistent with the concept of "natural intervals" (e.g., Schellenberg and Trehub, 1996), i.e., a "prewired" propensity for harmonically based melodic intervals.

It seems likely that the thresholds for mistuning of harmonics determined by Lee and Green were also based on musical interval judgments by their subjects and are not representative of the abilities of the general population. Only our best relative pitch subjects gave thresholds comparable to those obtained by their subjects, and subjects without relative pitch could not even perform the task. While Lee and Green did not specifically test the relative pitch proficiency of their subjects, they do mention that all three subjects had musical training and played an instrument.

Our results cast some doubt on the generality of the phenomenon, reported by Houtgast, of a subharmonic pitch

from a pure tone presented at a low sensation level. Some listeners can achieve results identical to those shown by Houtgast's listeners in the low SL condition by using relative pitch cues. Nevertheless, the fact that performance is usually poorest in the Houtgast without-noise condition suggests that the phenomenon reported by Houtgast does exist. This conclusion is strongly supported by the introspective comments of the subjects. In particular, the two authors were two of the subjects in group 1 in the original replication, and our subjective impressions mirror those of many of Houtgast's subjects: i.e., in the with-noise condition the low pitch cue was so compelling that it was not possible to tell which trials contained complex comparison tones, and which contained comparison tones comprising a single harmonic. Another subject, ERP2, was especially insistent that his poorer performance for some lower harmonics in the Houtgast without-noise condition was due to having to switch strategies among trials because he heard a low pitch for the comparison tone on some trials and a high pitch on others.

This conclusion raises the question of why, for some listeners, performance using entirely different strategies would be so similar, including the "finestructure." That is, performance tends to be poorer for odd harmonics and falls to chance at about harmonic 11. For the relative-pitch strategy, the decrement at odd harmonics stems from the fact that, in general, the ratios formed by the odd harmonics correspond to less salient intervals (e.g., major seconds) than those formed by even harmonics (e.g., octaves) and some odd harmonics (e.g., the 7th and 11th) do not form ratios corresponding closely to any standard musical intervals. The falloff at high harmonics (wide frequency ratios) presumably comes from the distortion of the subjective musical pitch scale at high frequencies [i.e., the stretching of the scale and musical paracusis (Ward, 1954)], which distorts the (multiple) octave transpositions. For the subharmonic pitch strategy, the falloff in performance at high harmonics reflects the "existence region" for the fundamental pitch of complex tones, an upper limit (for a fundamental frequency of 200 Hz) of about the 10th harmonic. This upper limit is independent of frequency resolution (e.g., Houtsma and Goldstein, 1972) and presumably reflects a degradation of the tonotopic or temporal cues upon which the fundamental pitch is based.

The reason for the unequal performance for odd and even harmonics in the subharmonic pitch strategy is not clear. Although all current pitch models are, in general, compatible with the concept of subharmonic pitches from a pure tone, none specifically predicts this asymmetry. Houtgast proposed an explanation for the asymmetry which also explains the above chance but incorrect (reversed) responses for some high odd harmonics, as seen, e.g., in the results of MRP1 and MRP2. Specifically, he suggested that for high harmonics, where the "fundamental" would be a high-order subharmonic and the fundamental pitch cue deteriorates, the listener can resort to listening to the subharmonic in the region an octave above the fundamental. This strategy would give the correct direction of pitch change for even harmonics of the fundamental, but would lead to confusion, or in some cases response reversal, for odd harmonics. It is noteworthy in this regard that for both Houtgast's listeners, and for the

group 1 listeners in our first replication, the performance asymmetry for harmonics below the 11th seen in the averaged data primarily stems from the results of one of the three listeners. The explanation is somewhat unsatisfying in that it is a quasi-relative-pitch explanation; it requires an affinity for octave relationships.

Another question is why so few of our listeners showed results like that of the typical Houtgast subject? Houtgast did not specify the criteria for choosing the three subjects who were extensively tested, and whose data we show in Fig. 1, but the basic response pattern of these subjects is typical of the majority of subjects in his initial experiment in which he tested 50 subjects on harmonics 5–7. Two-thirds of these listeners were above 70% correct for single harmonics in the with-noise condition, whereas only two or three subjects were above 70% correct in the without-noise condition. There are at least two factors involved. One factor was obviously that musical training allowed some of our subjects to use relative pitch cues in the without-noise condition. Houtgast also does not specify the musical training of his subjects so it is not possible to estimate what proportion might have been able to use relative pitch cues. Another factor would be the propensity for subjects to be either synthetic or analytic listeners for complex tones with only a few harmonics (Smooenberg, 1970). The performance for two- and three-harmonic comparison tones in the without-noise condition for our listeners without relative pitch was poorer than that of Houtgast's extensively tested subjects. The data indicate that for whatever reasons, we happened to select strongly analytic listeners for whom even a low SL presentation did not engender a low pitch.

III. CONCLUSIONS

The ability to tell whether consecutively presented tones are harmonically related is not universal and depends prima-

rily upon specific musical training. This ability, when present, can confound the results of pitch-related psychoacoustic experiments and should be taken into account when interpreting such results.

ACKNOWLEDGMENTS

Experiment A was performed at MIT and was supported by a NIH National Research Service Award to the first author. The remaining experiments were performed at the University of Washington and were supported by the Virginia Merrill Bloedel Hearing Research Center. Publication was supported by IPO.

- Allen, D. (1967). "Octave discriminability of musical and non-musical subjects," *Psychonomic Sci.* **7**, 421–422.
- Burns, E. M. (1999). "Intervals, scales, and tuning," in *Psychology of Music*, 2nd ed., edited by D. Deutsch (Academic, New York), pp. 215–264.
- Burns, E. M., and Campbell, S. L. (1994). "Resolution of frequency and frequency ratios by possessors of absolute and relative pitch: Examples of categorical perception?" *J. Acoust. Soc. Am.* **96**, 2704–2719.
- Houtgast, T. (1976). "Subharmonic pitches of a pure tone at low S/N ratio," *J. Acoust. Soc. Am.* **60**, 405–409.
- Houtsma, A. J. M., and Goldstein, J. (1972). "The central origin of the pitch of complex tones: Evidence from musical interval recognition," *J. Acoust. Soc. Am.* **51**, 520–529.
- Kallman, H. J. (1982). "Octave equivalence as measured by similarity ratings," *Percept. Psychophys.* **32**, 37–36.
- Lee, J. Y., and Green, D. M. (1994). "Detection of a mistuned component in a harmonic complex," *J. Acoust. Soc. Am.* **96**, 716–725.
- Pollack, I. (1952). "The information in elementary auditory displays," *J. Acoust. Soc. Am.* **24**, 745–749.
- Schellenberg, E., and Trehub, S. (1996). "Natural musical intervals: Evidence from infant listeners," *Psychological Sci.* **7**, 272–277.
- Smooenberg, G. F. (1970). "Pitch perception of two-frequency stimuli," *J. Acoust. Soc. Am.* **48**, 924–942.
- Terhardt, E., Stoll, G., Schermbach, R., and Parncutt, R. (1986). "Tonhöhereindeutigkeit, tonverwandschaft, und identifikation von sukzessiv-intervallen," *Acustica* **61**, 57–66.
- Thurlow, W. R., and Erchul, W. P. (1977). "Judged similarity in pitch of octave multiples," *Percept. Psychophys.* **22**, 177–182.
- Ward, W. D. (1954). "Subjective musical pitch," *J. Acoust. Soc. Am.* **26**, 369–380.

The effect of temporal placement on gap detectability

Karen B. Snell

Audiology Department, Rochester Institute of Technology, Rochester, New York 14623

Hue-Lu Hu

141 Glen Ellyn Way, Rochester, New York 14618

(Received 16 February 1999; revised 16 June 1999; accepted 2 September 1999)

The detectability of a masked sinusoid increases as its onset approaches the temporal center of a masker. This study was designed to determine whether a similar change in detectability would occur for a silent gap as it was parametrically displaced from the onset of a noise burst. Gap thresholds were obtained for 13 subjects who completed five replications of each condition in 3 to 13 days. Six subjects were inexperienced listeners who ranged in age from 18 to 25 years; seven subjects were highly experienced and ranged in age from 20 to 78 years. The gaps were placed in 150-ms, 6-kHz, low-passed noise bursts presented at an overall level of 75 dB SPL; the bursts were digitally shaped at onset and offset with 10-ms cosine-squared rise-fall envelopes. The gated noise bursts were presented in a continuous, unfiltered, white noise floor attenuated to an overall level of 45 dB SPL. Gap onsets were parametrically delayed from the onset of the noise burst (defined as the first nonzero point on the waveform envelope) by 10, 11, 13, 15, 20, 40, 60, 110, 120, and 130 ms. Results of ANOVAs indicated that the mean gap thresholds were longer when the gaps were proximal to signal onset or offset and shorter when the gaps approached the temporal center of the noise burst. Also, the thresholds of the younger, highly experienced subjects were significantly shorter than those of the younger, inexperienced subjects, especially at placements close to signal onset or offset. The effect of replication (short-term practice) was not significant nor was the interaction between gap placement and replication. *Post hoc* comparisons indicated that the effect of gap placement resulted from significant decreases in gap detectability when the gap was placed close to stimulus onset and offset. © 1999 Acoustical Society of America.

[S0001-4966(99)06312-2]

PACS numbers: 43.66.Mk, 43.66.Ba [SPB]

INTRODUCTION

Studies have demonstrated that masked signal detectability increases as the signal onset is delayed relative to masker onset and decreases as the signal approaches masker offset (Elliott, 1965; Zwicker, 1965; McFadden, 1989; Bacon, 1990). Zwicker (1965) referred to the phenomenon at signal onset as “overshoot” and reported that it had a time constant of about 15 ms. These temporal effects were attributed to various processes including adaptation and transient masking. The effects of temporal placement on gap detection are less well known. Penner (1977) reported that the gap thresholds increased with an increase in the duration of the leading noise burst from 2 to 200 ms. In contrast, Forrest and Green (1987) varied the location of a gap in a 100-ms noise burst from 10 ms after noise burst onset to 10 ms before noise burst offset, and reported that gap detection varied little with temporal position except that gaps were slightly more detectable when located about 30 ms after burst onset. Phillips *et al.* (1998) varied the temporal position of a gap by varying the length of the leading noise burst from 5 to 300 ms while holding the duration of the trailing noise burst at 300 ms. They reported that gap thresholds decreased significantly with increases in leading marker duration, though more so when the leading and trailing bursts differed spectrally.

In designing this experiment, our primary motivation was to explore whether gap detection was affected by the

temporal location of the gap. To this end, we obtained gap thresholds close to the onset of the signal but also at nine other placements throughout the stimulus waveform. A second motivation was to examine short-term practice effects. Neff *et al.* (1982) reported that individual means and patterns present during the first three replications persist through 150 hours of training. Using a different method, we compared the gap thresholds obtained during the second through fifth replications by six inexperienced and seven highly experienced subjects.

I. METHOD

A. Subjects

Six of the subjects (JR, KD, YA, LG, DC, ZC) were college students without prior experience in psychoacoustic experiments who were recruited by posted advertisements. All were between the ages of 18 and 25 years. Seven experienced listeners served as subjects and varied in age from 18 to 78 years (HH, 18 y; AS, 20 y; JB, 25 y; KS, 45 y; FM, 48 y; JL, 65 y; EB, 78 y). Two of the experienced subjects were the authors (KS and HH). HH had participated for eight hours as a subject in a previous gap detection study and listened to gaps for several months during monitoring of other subjects. KS had extensive listening experience. The remaining five experienced subjects had previously participated in four or five gap detection experiments for a total of

25 to 50 hours of listening experience each. In addition, subject FM had assisted with data collection in a variety of studies over the previous nine months and subject JB had listened to gap stimuli for several months while assisting in an animal experiment. Twelve of the 13 subjects had audiometric thresholds at octave frequencies between 0.25 and 8 kHz of 25 dB HL or less (ANSI S3.6-1989). Subject EB's thresholds fell within normal audiometric limits except for a 35-dB HL threshold at 8 kHz. All subjects were paid for their participation.

B. Stimuli

The gaps were carried by digitally generated, gated noise bursts that were low-passed at 6 kHz (TDT AP1), transduced by a 16-bit D/A converter (TDT DA1), and attenuated to an overall level of 75 dB SPL (TDT PA4). The noise bursts were mixed (TDT SM3) with a continuous white noise (TDT WG1) floor that had been attenuated (TDT PA4) to an overall level of 45 dB SPL ($N_0=4$ dB) with a bandwidth limited by the frequency response of the earphone, partially filling the gaps. The output of the mixer was led to a headphone buffer (TDT HB4) and then to a single earphone (Beyer DT48) mounted in a circumaural cushion. The rise portion of the leading noise burst and the fall portion of the trailing noise burst were shaped by 10-ms cosine-squared envelopes. Gap onsets and offsets were shaped by 1-ms cosine-squared envelopes. The standard noise bursts were also shaped by 1-ms cosine-squared envelopes at comparable placements to minimize the possibility that detection could be based on spectral and intensity artifacts unique to the signal.

Thus the standard and signal stimuli consisted of leading noise bursts, gaps, and trailing noise bursts. Overall stimulus duration (defined as the interval between first and last non-zero points on the stimulus waveform) was maintained at 150 ms throughout the experiment. In each of the ten experimental conditions the duration of the leading noise burst (defined as the interval between first and last nonzero points on the leading noise burst) was held constant throughout a run at 10, 11, 13, 15, 20, 40, 60, 110, 120, or 130 ms. Note that at a delay of 10 ms, the onset of the gap falls at the first point of full amplitude on the waveform envelope.

C. Procedure

Subjects sat in a double-walled sound booth (Acoustic Systems RE243). At the beginning of each run, the subjects listened repeatedly to samples of the standard noise burst paired with a signal noise burst containing an easily discriminable gap until familiar with the gap placement for that condition. Generally, subjects chose to sample no more than four or five noise burst pairs before beginning a run; many preferred to listen to only one or two. Gap thresholds were determined in a three-interval, forced-choice (3IFC) procedure using 75 trial runs. During each run, the placement of the gap onset was held constant and the duration of the gap (and thus the gap offset and the duration of the trailing marker) adaptively varied using a three-down, one-up rule to estimate the 79.4% point on the psychometric function (Lev-

itt, 1971). The step size was logarithmically varied to minimize the standard error of the threshold estimate. Feedback was provided after each response and the interstimulus intervals were 450 ms. In a 1-hour session, one threshold in each of the ten conditions was obtained in random order. All subjects completed five replications in 5 to 13 days. Four subjects completed two sessions in one day. The last four replications were used to estimate gap thresholds and examine the effects of gap placement and short-term practice.

II. RESULTS

A. Inexperienced subjects

The mean gap detection thresholds of the inexperienced subjects are shown in Fig. 1. In each panel, the placement of the gap onset relative to the noise burst onset is shown on the abscissa in ms. Gap detection thresholds in ms are plotted on the ordinate. Replication is the parameter with gray symbols representing the earlier and black symbols representing the later replications. The first six panels represent the individual data for the six inexperienced subjects. The bottom panel contains the mean gap thresholds for the six inexperienced subjects with error bars indicating \pm one standard error. The shortest mean gap thresholds for the inexperienced subjects were at gap placements of 40 and 60 ms (both 2.2 ms), whereas the longest mean gap thresholds were at delays of 10 (6.8 ms) and 11 ms (6.5 ms). The mean gap detection thresholds for the four replications averaged across all delays were 4.2, 4.0, 4.0, and 3.6 ms, respectively. A two-way repeated measures ANOVA (SYSTAT, 1997) was performed on the data for the six inexperienced subjects with gap placement and replication as the two main factors. The results were significant for the main effect of gap placement [$F(9,45)=10.862, p<0.003$]. Neither replication [$F(3,15)=1.335, p=0.306$] nor the interaction between gap location and replication [$F(27,135)=1.450, p=0.234$] were significant. *Post hoc* contrasts (SYSTAT, 1997) between the mean gap thresholds for adjacent conditions revealed significant differences ($p<0.05$) between the mean gap thresholds at delays of 11 and 13 ms, 15 and 20 ms, and 20 and 40 ms, with poorer acuity in each case for the location closer to stimulus onset. Similarly, significant differences were found between the mean gap thresholds at delays of 60 and 110 ms and between the mean gap thresholds at delays of 120 and 130 ms, with poorer acuity in both cases for the location closer to stimulus offset. In summary, the effect of gap placement on gap thresholds of the inexperienced listeners was limited mainly to changes in detectability within 40 ms of burst onset and offset.

B. Results with experienced listeners

The gap thresholds for the seven highly experienced subjects are shown in Fig. 2. As in Fig. 1, the placement of the gap onset relative to the noise burst onset is shown on the abscissa and the gap detection thresholds are plotted on the ordinate. Replication is the parameter. The bottom panel on the right contains the mean thresholds for each replication for all seven experienced subjects (KS, HH, FM, JB, AS, JL, and EB), with error bars indicating \pm one standard error of

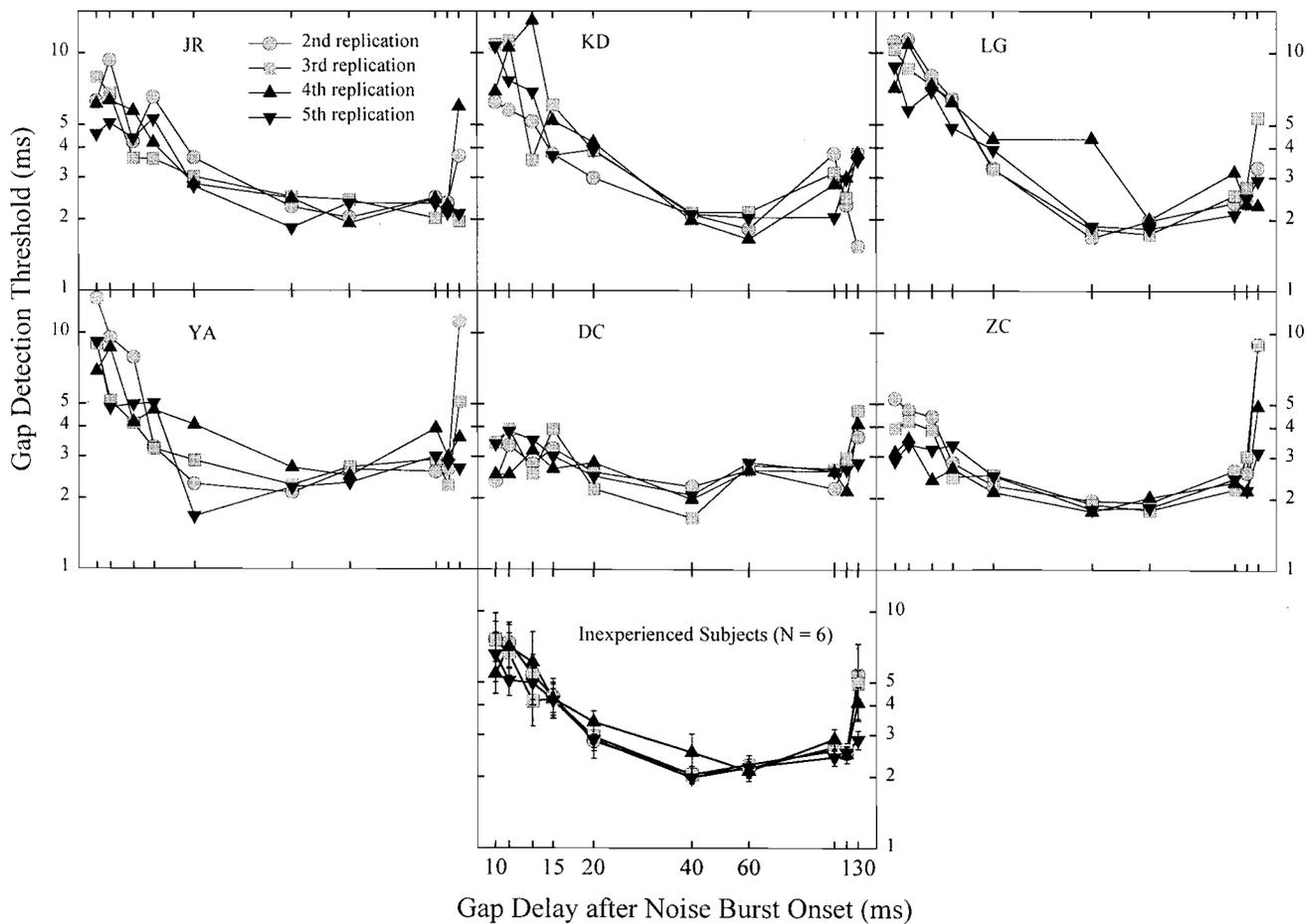


FIG. 1. Gap detection thresholds (ms) as a function of the temporal placement of the gap. The gap was parametrically displaced (delayed) from the onset of the stimulus noise burst. Replication is the parameter. The top six panels contain individual data for the inexperienced subjects. The bottom panel contains the mean gap thresholds for the inexperienced subjects. The lighter symbols represent thresholds obtained during the second and third replications; the darker symbols represent those obtained during the last (the fourth and fifth) replications.

the mean. The smallest mean gap threshold was at a delay of 40 ms (2.1 ms). The largest mean gap thresholds were at delays of 10 (4.2 ms) and 11 ms (3.9 ms). The mean gap detection thresholds for the four replications averaged across all delays were 3.2, 2.9, 3.0, and 3.0 ms, respectively.

A two-way repeated measures ANOVA (SYSTAT, 1997) was performed on the gap thresholds of the five younger highly experienced subjects with gap placement and replication as the two main factors. The data of the older subjects were excluded in part because of the recent study by Schneider and Hamstra (1999) which suggested the possibility of a significant interaction between age and proximity of gap placement to stimulus onset. The results were significant for the main effect of gap placement [$F(9,36)=4.673$, $p < 0.004$]. Neither replication [$F(3,12)=0.646$, $p = 0.600$] nor the interaction between gap location and replication [$F(27,108)=1.025$, $p = 0.442$] were significant. *Post hoc* contrasts between the mean gap thresholds for adjacent conditions revealed a significant ($p < 0.05$) decrease in mean gap thresholds as the delay in gap placement increased from 20 to 40 ms.

C. The effect of long-term experience

The mean gap thresholds for the three subject groups averaged across replication are shown in Fig. 3. As before,

the placement of the gap relative to the burst onset is shown on the abscissa in ms. Gap detection thresholds in ms are plotted on the ordinate. The light squares represent the younger, inexperienced subjects, the light circles represent the younger, highly experienced subjects, and the dark circles represent the older, experienced subjects. Note that the mean gap thresholds of the younger experienced subjects overlap those of the younger inexperienced subjects only at points distal from stimulus onset or offset. This suggests that experience influences gap thresholds in conditions where gap detectability is poorer (as indicated by larger gap thresholds).

To determine whether significant differences were present between the younger experienced ($N=5$) and inexperienced ($N=6$) subjects, a repeated measures ANOVA (SYSTAT, 1997) was performed on the mean gap thresholds with one between factor (experienced and inexperienced groups) and one within factor (10 gap placements). The main effects of group [$F(1,9)=8.692$, $p < 0.016$] and gap placement [$F(9,81)=12.374$, $p < 0.001$] and the interaction between gap location and group [$F(9,81)=4.802$, $p = 0.017$] were significant. *Post hoc* contrasts revealed significant decreases ($p < 0.05$) in mean gap thresholds as the gap delays increased from 11 to 13 ms, 15 to 20 ms, and 20 to 40 ms. Mean gap thresholds increased significantly as the gap delay was increased from 60 to 110 ms, and 120 to 130 ms.

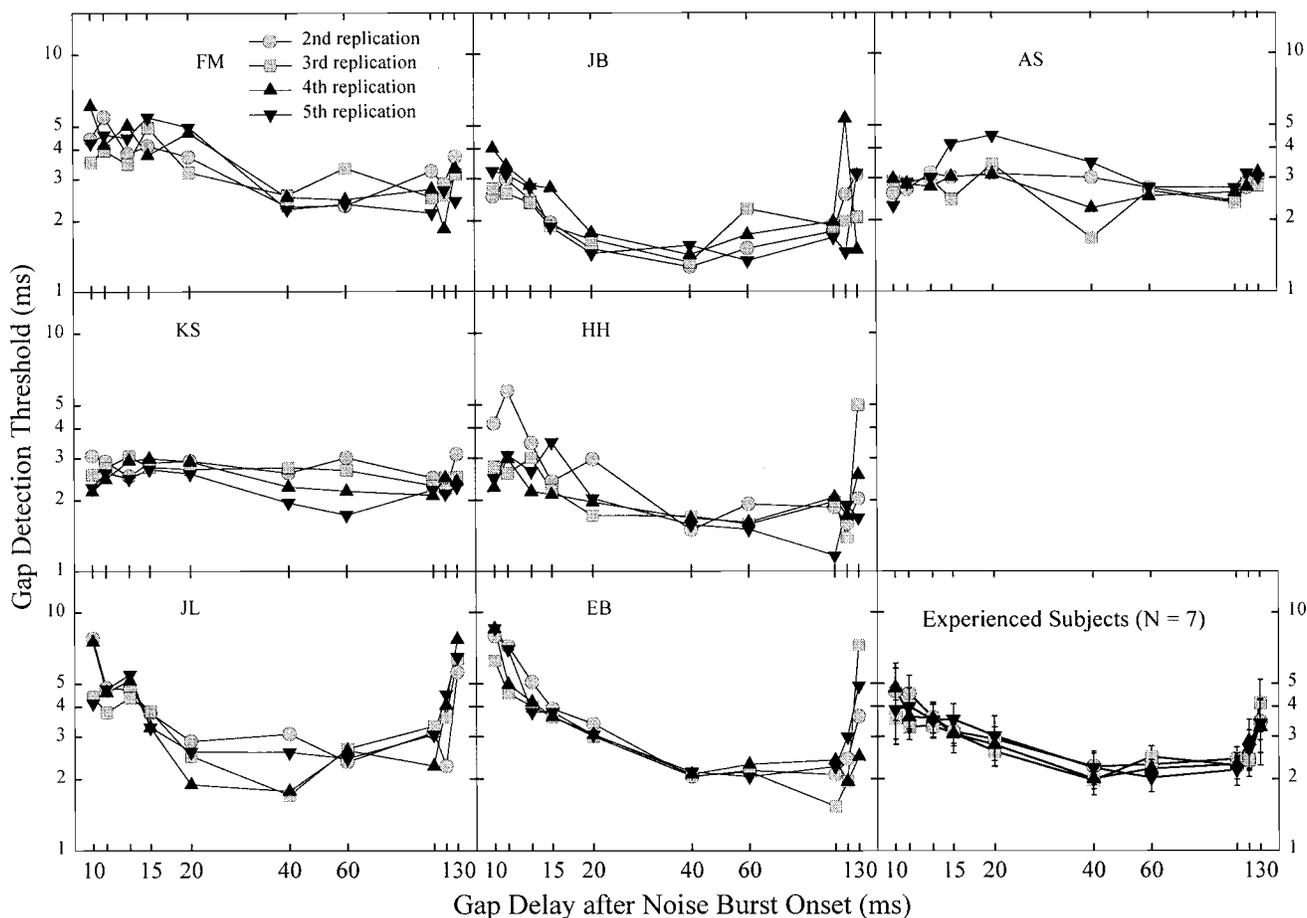


FIG. 2. Gap detection thresholds (ms) as a function of the temporal placement of the gap. The gap was parametrically displaced (delayed) from the onset of the stimulus noise burst. Replication is the parameter. The lighter symbols represent thresholds obtained during earlier (the second and third) replications; the darker symbols represent those obtained during the last two (the fourth and fifth) replications. The top three panels contain individual data for three young, experienced subjects. The middle two panels contain the individual data for the two authors and the left and middle lower panels contain individual data for two older, experienced subjects. The rightmost lower panel contains the mean data for the experienced subjects.

Tukey's honestly significant difference (HSD) test was used to explore the significant interaction between group and delay and the results indicated that the two younger groups differed significantly ($p < 0.05$) only at gap placements of 10, 11, 13, and 130 ms.

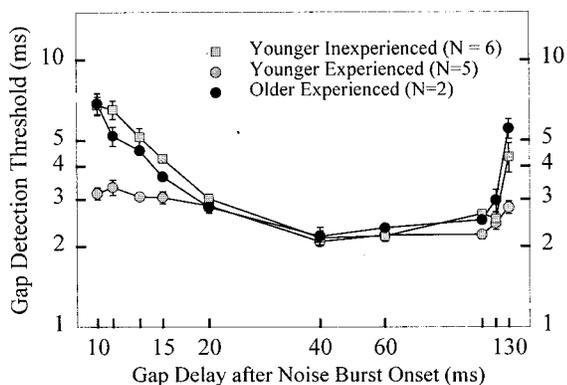


FIG. 3. Gap detection thresholds (ms) as a function of the temporal placement of the gap. The mean thresholds of younger, inexperienced subjects ($N=6$) are indicated by gray squares, those of younger, experienced subjects ($N=5$) by gray circles, and those of older, experienced subjects by black circles. Error bars indicate plus and minus one standard error.

III. DISCUSSION

A. The effect of gap placement

The temporal course of gap detection is similar to that of simultaneous masking. In most inexperienced listeners, thresholds decrease as the gap is delayed relative to the onset of the noise burst. They increase again near noise burst offset. Among experienced listeners, some listeners show an obvious gap placement effect and in others it is small. At a stimulus level commensurate with that used in this study, Penner (1977) reported that gap thresholds of three subjects (one of whom was the author) did not vary (Fig. 3, p. 554) when the trailing burst was increased from 2 to 200 ms if the leading burst were held constant at 200 ms. That is, 3-ms gap thresholds were obtained when the gap was close to the offset (of a 202-ms noise burst) and also when the gap was placed midpoint (in a 400-ms noise burst). This result is in agreement with our finding of no significant increase in gap threshold at stimulus offset for the experienced younger listeners. In the same study, Penner reported that increasing the leading burst from 2 to 20 to 200 ms while holding the trailing burst constant at 2 ms decreased gap thresholds of two listeners (one of whom was the author) from roughly 3, to 2.5, to 0.7 ms. Again, these results are roughly in agree-

ment with ours except for the unusually small thresholds (0.7 ms) obtained when both the leading and trailing markers were 2 ms. As Forrest and Green (1987) have previously suggested, it may be that these thresholds reflected detection of an increment in duration (0.7 ms/4 ms = increment of 18%) rather than detection of the gap itself.

Forrest and Green (1987) explored the effects of placement on gap detectability for three experienced (five hours of practice) subjects. They used a 100-ms noise burst and gap placements of 10, 30, 50, 70, and 90 ms after the onset of the waveform. The overall presentation level of each burst was varied randomly over a 10-dB range to minimize the use of intensity as a detection cue. They reported little influence of gap location except that the gap thresholds were “slightly better” when the gap was located about 30 ms after the noise burst onset. As can be seen in their Fig. 3 (1987, p. 1935), between-subject variability was also reduced at this placement. Both of these findings are consistent with the minima seen in our figures at delays of 40 and 60 ms. That thresholds were not elevated at a displacement of 10 ms appears inconsistent with our findings. However, their waveforms were not shaped at onset whereas ours were shaped by 10-ms cosine-squared envelopes. It may be more appropriate to compare our thresholds at 20 ms (10 ms after the first point of full amplitude on the waveform) with their thresholds at 10 ms. This adjustment improves the agreement between studies.

More recently, Eggermont (1995) reported that in the cat auditory cortex, the coding of gaps was poorer when they occurred close to stimulus onset rather than 500 ms later. Phillips *et al.* (1997) examined the effects of placement on gap detection in three young adults, two of whom were experienced subjects. They reported no systematic effect of gap placement in experimental conditions similar to those in this study. This result differs from those of a second study (Phillips *et al.*, 1998), in which ANOVA revealed a significant effect of placement in similar experimental conditions. The effect was described as “small,” but this characterization was likely due in part to comparison with the large effects seen in other experimental conditions that were dissimilar to those used in this study.

B. The effects of experience

Our finding that gap detection thresholds do not improve significantly during the second through fifth replications differs somewhat from observations by Phillips *et al.* (1997). They noted that for experimental conditions comparable to those in this study, learning curves for gap detection were “very short” and that performance reached a plateau in as few as five threshold determinations. This suggests that the gap thresholds of some of their subjects may have improved, perhaps significantly, across the first several replications.

Neff *et al.* (1982) reported that even after roughly 150 hours of practice, individual and mean patterns present in the first block of three replications were similar to those seen in the last four replications. This is consistent with our results shown in Figs. 1 and 2 and the statistical analyses. The results of the statistical analyses suggest that for both inexperienced and experienced subjects, the effect of gap placement is significant but the effect of replication is not. Our findings

differ, however, in that we find a significant effect on temporal acuity of long-term experience. That is, although the overall patterns are similar in that the smallest thresholds for both groups are found for placements distal from stimulus onset and offset, the mean thresholds of the more experienced younger subjects are significantly smaller at noise burst onset and offset.

C. The effect of age

As illustrated in Fig. 3, age-related differences in gap thresholds between younger and older experienced subjects appear to increase as gap placements approach stimulus onset and offset. There the gap thresholds of the older experienced subjects are more comparable to those of the younger inexperienced subjects than they are to those of the younger experienced subjects. At gap placements closer to the center of the burst, the gap thresholds of all three groups are similar. It may be that extended experience is less effective in improving the gap thresholds of older subjects. Alternatively, gap detection may be poorer at all delays in older inexperienced subjects but improve relatively more with practice at midpoint placements. At any rate, it is clear that the gap thresholds of even highly experienced, older subjects are poorer than those of younger, experienced subjects in conditions where gap detectability is poorer. These results are in general agreement with those recently reported by Schneider and Hamstra (1999). They found that gap thresholds in subjects (presumably inexperienced) were influenced by both marker duration and age. Gap detection thresholds of older adults were higher than those of younger adults for marker durations of less than 250 ms, and age-related increases in gap thresholds were inversely proportional to marker duration. Since the gap placement approaches stimulus onset and offset as marker duration is shortened, these results are broadly consistent with our results despite differences between studies in stimuli and absolute sensitivity of subjects. The present results are similar to results from He *et al.* (1999). They examined the effect of gap location using young and older subjects and 400-ms noise burst carriers to generate psychometric functions in a “yes/no” paradigm. None of their subjects had previous experience in temporal resolution studies. They reported that mean gap thresholds were shortest at the temporal center of the noise burst and increased at stimulus onset and offset, and that the effect was more marked for older subjects, especially at stimulus offset.

D. Physiological bases of the placement effect

The effects of maturation on gap detection have been explored by Trehub *et al.* (1995) using short sinusoidal carriers. They reported a smaller difference in gap detection thresholds for infants and adults than had been previously found with longer duration carriers. To explain this smaller difference, they suggested that infants were more sensitive to the effects of adaptation, showing more profound adaptation and slower recovery from adaptation. Further, they noted that a proportionately decreased relative response of adapted fibers should worsen gap detection thresholds since the overshoot at the onset of the trailing carrier would be less pro-

nounced. It follows that gap thresholds would be elevated with stimuli producing greater adaptation, such as longer duration leading carriers. Because they used very short carriers, this could explain their relatively smaller maturational effect.

Although Trehub *et al.* (1995) did not directly measure the effects of gap placement, they argued that in general, the detectability of a gap should worsen as the duration of the carrier before the gap increased. Phillips *et al.* (1997) later noted that following the general arguments of Plomp (1964) and Penner (1977) one would predict longer gap thresholds for longer leading carriers. Our results show the opposite to be the case. Detectability of a gap improves as the duration of the leading carrier increases (until the gap begins to approach the stimulus offset), the opposite of what one would expect if adaptation of primarylike (PL) auditory nerve fibers were responsible for the effect. However, desynchronization of PL fibers at stimulus onset could contribute to poorer detectability at stimulus onset. Alternatively, other types of units, namely those that show greatest adaptation at stimulus onset and offset, may be responsible.

As Walton *et al.* (1997) reported in a study of the neural correlates on gap detection in the young mouse, phasic on, on-off, and off-type (ON, ON-OFF and OFF, respectively) units discharge synchronously to the onset and offset of gaps as well as to onsets and offsets of noise bursts. Of all unit types, a higher proportion of phasic ON- and ON-OFF types will be in a refractory state shortly after stimulus onset than units of any other type. While PL neurons as well as phasic ON- and ON-OFF types have minimal gap thresholds within or below behavioral thresholds, the PL neurons (5%) are outnumbered by the ON- and ON-OFF types (60%) in the inferior colliculus (IC), an obligatory synapse in the ascending auditory pathway. ON and ON-OFF units that respond to the onset of the leading marker are less likely to respond to the onset of a gap placed proximal to the noise burst onset. Thus refractoriness, or adaptation, in the response of phasic ON-type units encountered in the inferior colliculus may be the limiting factor in gap detection when the gap is placed close to onset.

The decreased detectability at stimulus offset can not be directly attributed to adaptation in phasic OFF responses, since presumably the response to the onset of the gap precedes the response to the carrier offset. Also, while the phasic OFF units discharge synchronously to mark gaps, Walton *et al.* (1997) reported that their mean gap threshold is 12.67 ms, far longer than the mean behavioral gap thresholds reported in the same species. Rather, the offset effect resembles backward masking, in that the response of those neurons coding the offset of the gap appear to be masked by their response to the offset of the trailing carrier. In fact, the placement effect at offset is just as strong as it is at onset, if not more so for the inexperienced subjects. As shown in Fig. 3, the gap threshold at a gap placement of 15 ms for younger inexperienced subjects is roughly 4 ms. At a comparable displacement from signal offset (a gap placement of 130 ms since the total duration is 150 ms and placement is with reference to the onset of the gap), the threshold is slightly larger, roughly 4.5 ms. This resembles nonsimultaneous masking in that at very short intervals between masker and

signal, backward masking can produce greater threshold shifts than forward masking.

When the gap more closely resembles waveform characteristics, i.e., when the amplitude of the envelope (aside from the gap itself) is rapidly changing, as it does at both offset and onset of the stimulus, attention may influence thresholds. If this were the case, one might expect experienced subjects to be less influenced by gap placement. This is true for the younger experienced listeners. However, as seen in Fig. 3, older experienced subjects show an effect of placement similar to that of the inexperienced subjects. A parsimonious explanation might attribute onset and offset effects on gap detection primarily to refractoriness in phasic units in the inferior colliculus. Other phenomena, however, such as desynchronization of PL fibers, attention, and backward masking at offset, may have influenced our results.

IV. CONCLUSION

The effects of placement on gap detection follow a time course like that seen in overshoot studies. Gap detectability is poorest at noise burst onset where gap thresholds are two to three times larger than when the gap is moved to the middle of the stimulus waveform. We suggest that the decreased detectability for these gaps may primarily reflect refractoriness and desynchronization in the neural response of phasic ON-type units resulting from the earlier response to the onset of leading noise burst. Similarly, mean gap thresholds are two to three times longer when the gap is moved from the middle of stimulus waveform to near the stimulus offset.

Neither experienced nor inexperienced subjects showed significant effects of short-term practice. However, younger, experienced subjects show increases in gap thresholds for placements close to stimulus offset as well as stimulus onset. Attention and backward masking, as well as desynchronization and adaptation in the neural offset response, may contribute to decreased detectability when the gap is near the stimulus offset.

ACKNOWLEDGMENTS

This article is based in part on a paper presented at the 129th Meeting of the Acoustical Society of America in 1995. We are grateful to Craig Champlin, Sid Bacon, and an anonymous reviewer for their insightful comments. This research was supported by the Rochester International Center for Hearing and Speech Research and a grant from NIA (Grant No. AG09524).

- ANSI (1989). ANSI S3.6-1989, "Specifications for audiometers" (American National Standards Institute, New York).
- Bacon, S. P. (1990). "Effect of masker level on overshoot," *J. Acoust. Soc. Am.* **88**, 698-702.
- Eggermont, J. J. (1995). "Neural correlates of gap detection and auditory fusion in cat auditory cortex," *NeuroReport* **6**, 1645-1648.
- Elliott, L. L. (1965). "Changes in the simultaneous masked threshold of brief tones," *J. Acoust. Soc. Am.* **38**, 738-746.
- 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.

- He, N., Horwitz, A. R., Dubno, J. R., and Mills, J. H. (1999). "Psychometric functions for gap detection in noise measured from young and aged subjects," *J. Acoust. Soc. Am.* **106**, 966–978.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- McFadden, D. (1989). "Spectral differences in the ability of temporal gaps to reset the mechanisms underlying overshoot," *J. Acoust. Soc. Am.* **85**, 254–261.
- Neff, D. L., Jesteadt, W., and Brown, E. L. (1982). "The relation between gap discrimination and auditory stream segregation," *Percept. Psychophys.* **31**, 493–501.
- Penner, M. J. (1977). "Detection of temporal gaps in noise as a measure of the decay of auditory sensation," *J. Acoust. Soc. Am.* **61**, 552–557.
- Phillips, D. P., Hall, S. E., Harrington, I. A., and Taylor, T. L. (1998). "'Central' auditory gap detection: A spatial case," *J. Acoust. Soc. Am.* **103**, 2064–2068.
- Phillips, D. P., Hall, S. E., Taylor, T. L., Carr, M. M., and Mossop, J. E. (1997). "Detection of silent intervals between noises activating different perceptual channels: Some properties of 'central' auditory gap detection," *J. Acoust. Soc. Am.* **101**, 3694–3705.
- Plomp, R. (1964). "Rate of decay of auditory sensation," *J. Acoust. Soc. Am.* **36**, 277–282.
- Schneider, B. A., and Hamstra, S. J. (1999). "Gap detection thresholds as a function of tonal duration for younger and older listeners," *J. Acoust. Soc. Am.* **106**, 371–380.
- SYSTAT for Windows: Statistics, Version 7.0 edition, SPSS, Inc., Chicago, IL, 1997.
- Trehub, S. E., Schneider, B. A., and Henderson, J. L. (1995). "Gap detection in infants, children, and adults," *J. Acoust. Soc. Am.* **98**, 2532–2541.
- Walton, J. P., Frisina, R. D., Ison, J. R., and O'Neill, W. E. (1997). "Neural correlates of behavioral gap detection in the inferior colliculus of the young CBA mouse," *J. Comp. Physiol. A* **181**, 161–176.
- Zwicker, E. (1965). "Temporal effects in simultaneous masking by white-noise bursts," *J. Acoust. Soc. Am.* **37**, 653–663.

The role of perceived spatial separation in the unmasking of speech

Richard L. Freyman^{a)} and Karen S. Helfer

Department of Communication Disorders, University of Massachusetts, Amherst, Massachusetts 01003

Daniel D. McCall and Rachel K. Clifton

Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01003

(Received 21 December 1998; revised 4 June 1999; accepted 13 August 1999)

Spatial separation of speech and noise in an anechoic space creates a release from masking that often improves speech intelligibility. However, the masking release is severely reduced in reverberant spaces. This study investigated whether the distinct and separate localization of speech and interference provides any perceptual advantage that, due to the precedence effect, is *not* degraded by reflections. Listeners' identification of nonsense sentences spoken by a female talker was measured in the presence of either speech-spectrum noise or other sentences spoken by a second female talker. Target and interference stimuli were presented in an anechoic chamber from loudspeakers directly in front and 60 degrees to the right in single-source and precedence-effect (lead-lag) conditions. For speech-spectrum noise, the spatial separation advantage for speech recognition (8 dB) was predictable from articulation index computations based on measured release from masking for narrow-band stimuli. The spatial separation advantage was only 1 dB in the lead-lag condition, despite the fact that a large perceptual separation was produced by the precedence effect. For the female talker interference, a much larger advantage occurred, apparently because informational masking was reduced by differences in perceived locations of target and interference. © 1999 Acoustical Society of America. [S0001-4966(99)01412-5]

PACS numbers: 43.66.Pn, 43.66.Qp, 43.66.Dc [DWG]

INTRODUCTION

The perception of speech is improved when the source of speech is separated spatially from the source of interference, whether that interference is a steady or fluctuating noise, a second talker, or a group of talkers (e.g., Hirsh, 1950; Koenig, 1950; Kock, 1950; Dirks and Wilson, 1969; MacKeith and Coles, 1971; Plomp, 1976; Plomp and Mimpen, 1981; Bronkhorst and Plomp, 1988, 1992; Koehnke and Bessing, 1996; Yost *et al.*, 1996; Peissig and Kollmeier, 1997; Hawley *et al.*, 1999). However, our understanding of the size and underlying basis of this improvement as it occurs in an anechoic environment is well ahead of our understanding of this improvement in the reverberant situation. For an anechoic environment, Zurek (1993) has summarized the research on the benefits that occur when speech and noise originate from separate locations. The improvement in speech recognition is assumed to be due to the reduction in masking that occurs when signal and noise are spatially separated. The reduction is as large as 16 dB for broadband and bandlimited signals (Saber *et al.*, 1991; Good *et al.*, 1997).

Two basic factors are involved in the improved signal detection: head shadow and binaural interaction. The head shadow effect arises because the source-to-ear transformations vary depending on source location (e.g., Shaw, 1974). Assuming a condition in which the source of the target is directly in front at 0 degrees azimuth, and a noise source is 60 degrees to the right, then at the left ear (away from the

noise) the head shadow will have the effect of attenuating the noise more than the target, improving the signal-to-noise ratio at that ear relative to when target and noise come from a common location. The head shadow effect is frequency dependent, being much larger at higher than at lower frequencies due to the short wavelengths of high-frequency sounds. Zurek (1993) assumed that a listener could make full use of the ear with the more favorable signal detectability in each frequency band.

The second effect, binaural interaction, is mostly a low-frequency phenomenon, in which the binaural auditory system takes advantage of differing interaural time delays created by signal and noise to produce less masking than in conditions in which there is identical time delay. Levitt and Rabiner (1967a, 1967b) demonstrated how reductions in masking due to binaural interaction, in combination with articulation theory (French and Steinberg, 1947), could be used to predict speech intelligibility advantages for antiphase conditions under headphones. Zurek's (1993) synthesis combined the binaural masking release with head shadow advantages to create a predictive model of the benefit of spatial separation on speech recognition in an anechoic sound field.

In comparison to the anechoic environment, the benefits of spatial separation in a room with reflections are reduced (e.g., Hirsh, 1950; MacKeith and Coles, 1971; Plomp, 1976; Bronkhorst and Plomp, 1988; Koehnke and Bessing, 1996), and the underlying bases of the remaining benefits are not immediately obvious. Assume now that the anechoic room has been modified so that the left wall is a reflective surface. The target is still at 0 degrees and the noise is at 60 degrees to the right. The attenuation of the noise at the left ear due to

^{a)}Electronic mail: rlf@comdis.umass.edu

the head shadow effect will be partially negated by the reflection off the left wall, potentially reducing the advantage in signal-to-noise (S-N) ratio at that ear. The reflection will also obscure the simple differences in interaural phase between two separate source locations, resulting in smaller release from masking (see Koenig *et al.*, 1977). As a result of the reduction of head shadow and binaural interaction advantages, detection of the target would not be expected to be improved much by its spatial separation from a competing noise.

This conclusion about rooms with reflections runs counter to our intuition and our experience that even in such rooms, there is a large perceptual benefit when speech and noise are separated spatially. The purpose of the current research is to identify and quantify the contributions of an additional potential benefit that is perceptual in nature, namely that when speech and noise sources are physically separated, they also “seem” to be in different places, as Hirsh (1950) pointed out. Even in strongly echoic spaces, different sound sources are perceived as separate auditory events in their respective spatial locations. This phenomenon, wherein a signal and its reflections are gathered into a single image perceived near the location of the original source, is known as the precedence effect (see Zurek, 1987, and Gilkey and Anderson, 1997, for reviews). Hirsh (1950) suggested that the precedence effect, by preserving localization of speech and noise in reverberant environments, may facilitate speech understanding in noise in such environments. When applied to the target/masker situation in a reflective environment, the precedence effect may bestow the benefits of true spatial separation, even though head shadow and binaural interaction advantages are reduced.

The extent to which Hirsh’s idea is correct may well depend on whether the masker produces mostly “energetic” or “informational” masking. Energetic masking is the classic conceptualization of masking, where a signal, such as a pure tone, is inaudible in the presence of a noise because the neural elements that would normally respond to the signal are either suppressed or swamped by the masker. With informational masking, a distracting sound makes it difficult to attend to the target and perceptually disentangle it from the interference (see, for example, Watson *et al.*, 1976; Leek *et al.*, 1991; Kidd *et al.*, 1994, 1995, 1998; Doll and Hanna, 1997). It is assumed that higher level cognitive processes are involved in analyzing signals in the presence of an informational masker. Studies with nonspeech stimuli suggest that spatial separation is especially effective in reducing the informational type of masking. Kidd *et al.* (1998) investigated the effect of spatial separation on the identification of tone patterns in the presence of informational and energetic types of masking. The informational masker was a sequence of complex tone patterns that interfered with listeners’ ability to recognize the target tone pattern, even in conditions where little energetic masking should have occurred. The energetic masker was a broadband noise. Kidd *et al.* (1998) showed that with the informational type of masker, the advantage of spatial separation of target and masker was in some cases greater than 30 dB, much larger than the spatial separation advantage for the broadband noise. Large spatial separation

TABLE I. Loudspeaker configurations for presentation of stimuli. In the text the conditions are either referred to by number (1–6) or by locations of target and interference. For example, F-F means target and interference are both from the front loudspeaker. FR-RF means that the front loudspeaker leads the right loudspeaker for the target, and the right loudspeaker leads the front for the interference.

Condition	Physical signal		Expected perceived image	
	Target	Interference	Target	Interference
1 F-F	Front	Front	Front	Front
2 F-R	Front	Right	Front	Right
3 F-FR	Front	Front-Right	Front	Front
4 F-RF	Front	Right-Front	Front	Right
5 FR-FR	Front-Right	Front-Right	Front	Front
6 FR-RF	Front-Right	Right-Front	Front	Right

advantages for this masker were maintained even when head shadow advantages were accounted for.

Given these results with nonspeech signals and maskers, it is reasonable to assume that the degree to which the understanding of speech is improved by its spatial separation from a masker will also depend greatly on the nature of the masker. This may be particularly true in a reverberant room because, as discussed above, reverberation may minimize the release from energetic masking gained by spatial separation. On the other hand, the precedence effect creates a *perceived* separation of target and masker in a reverberant room that may aid performance. We hypothesize that only informational masking would benefit from perceived separation that, because of reflections, is not accompanied by substantial head shadow and binaural interaction advantages.

In the current study we investigated spatial separation advantages in speech recognition with two different types of maskers: one that we assume produced only energetic masking and one that we assume produced both energetic and informational masking. The target speech stimuli were nonsense sentences spoken by a female talker. One masker, a steady speech-spectrum noise, was considered to produce purely energetic masking. The other masker, the speech of a second female talker producing similar sentences, was assumed to create informational masking in addition to energetic masking. In the main conditions, simulated reflections were used to minimize the energetic masking advantage gained by spatial separation. For such conditions, it was predicted that a significant advantage of spatial separation would occur only for the female talker masker, not for the speech-shaped noise. Predictions for anechoic conditions are less clear, because in the case of the female talker masker, there is likely to be release from both energetic and informational masking. It is not known how these two benefits might add.

I. EXPERIMENT I: SPEECH RECOGNITION

A. Methods

1. Experimental conditions

A total of six loudspeaker presentation conditions were used for the target and interference stimuli, as shown in Table I. Target and masker were either single source or were

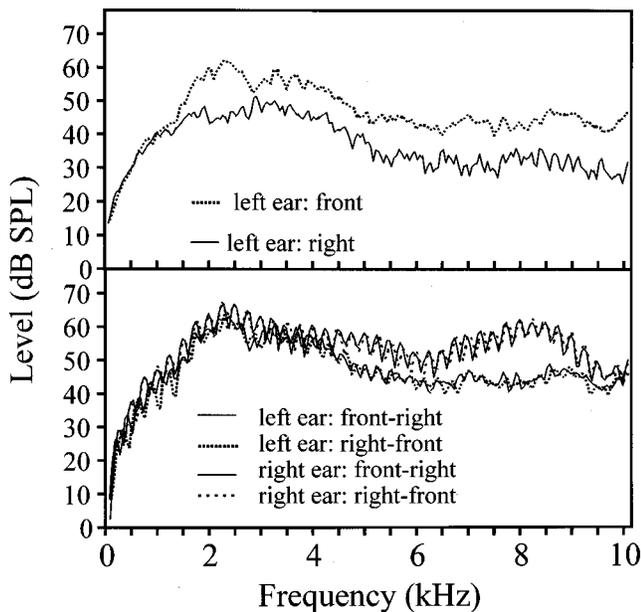


FIG. 1. Spectral measurements from the ears of KEMAR for a white-noise stimulus. Top panel: the spectrum at the left ear when the noise is from 0 degrees (front) or 60 degrees to the right (right). The head shadow effect is seen above 1.5 kHz. Bottom panel: the spectrum at the two ears for two-source noises with a 4-ms delay. The spectrum is different at the two ears, but does not depend significantly upon whether the front (front-right) or right (right-front) loudspeaker is the lead. Except for some of the periodicity seen in the bottom panel, the response for the front condition in the top panel is highly similar to the left-ear responses in the bottom panel.

presented from two loudspeakers with a 4-ms delay to one loudspeaker. In conditions 1–4, the target was presented only from the front loudspeaker while the loudspeaker configuration of the interference varied. In condition 1 the interference was presented from the front loudspeaker only and condition 2 from the right only. The spatial separation of target and noise in condition 2 should improve audibility of the target due to head shadow and binaural interaction. In Fig. 1 (top panel), the SPL of white noise at the left ear of the KEMAR manikin placed at the listener’s position in an anechoic chamber is plotted as a function of frequency when delivered from the front loudspeaker (condition 1) and the right loudspeaker (condition 2). The effect of head shadow is clearly seen in the high frequencies. Also, there should be an overall detection advantage for low frequencies in condition 2 due to binaural interaction. These advantages for detection are expected to lead to benefits for speech recognition as summarized in Zurek’s (1993) model.

In conditions 3 and 4, the target was again presented only from the front loudspeaker but the interference was presented from both loudspeakers. In condition 3 the right loudspeaker was delayed by 4 ms, and in condition 4 the front loudspeaker was delayed by 4 ms. This difference between conditions 3 and 4, a reversal of the lead and lag loudspeaker of the masker, has less predictable consequences for speech recognition than does the conditions 1 versus 2 difference. KEMAR measurements at the left and right ears for conditions 3 and 4 are shown in Fig. 1 (bottom). The spectral peaks show a periodicity of 250 Hz due to the 4-ms delay, but the differences in level between the two masker conditions at either ear are quite small. While this figure does not

display interaural phase delays that might lead to differences in masking in the low frequencies, it does indicate that the considerable head shadow advantage seen in the top panel of the figure is essentially eliminated with this masker configuration. On the other hand, there are clear perceptual differences between the two maskers. Because of the precedence effect, in condition 3 the perceived location of the masker should be fairly close to the lead (front) location, which is the also the location of the target. In condition 4, the perceived location of the masker should be to the right, well separated perceptually from the target location.¹ The purpose of the experiment is to determine whether condition 4 is therefore a better environment for speech recognition than condition 3.

Conditions 1–4 constitute a set of conditions in which the different configurations of interference can be directly compared for a fixed single source target. As a consequence, in conditions 3 and 4 the target was single source even though the masker consisted of a source and a simulated reflection. In real rooms, both target and interference would have reflections. To create a more realistic simulation, two additional conditions were included (5 and 6 in Table I) in which the target and masker were *both* lead-lag pairs. The target was front-right, but the masker was either front-right (condition 5) or right-front (condition 6). Because of the added “reflection,” the differences in energetic masking between conditions 5 and 6 should be much less than in the 1 versus 2 comparison. However, like condition 2, condition 6 will create a large perceptual separation between target and interference. Performance on conditions 1 and 2 will be contrasted with conditions 5 and 6, with the expectation that any difference between the latter two conditions would be due to a release from informational masking due to perceived location separation.

2. Apparatus, stimuli, and procedures

The experiments were conducted in an IAC anechoic chamber measuring $4.9 \times 4.1 \times 3.12 \text{ m}^3$. The walls, floor, and ceiling are lined with 0.72-m foam wedges. Subjects sat in the center of the room in front of a semicircular arc constructed of wood and covered with foam. Two Realistic Minimus 7 loudspeakers were positioned on the arc, one at 0 degrees (directly in front of the listener) and one at 60 degrees to the right. The loudspeakers were angled to face the listener at distance of 1.9 m from the center of the listener’s head and at a height of 1.4 m from the wire mesh floor of the anechoic chamber, ear height for the average seated subject.

Speech stimuli were 320 “nonsense” sentences developed by Helfer (1997). The sentences are correct syntactically but are not meaningful. Examples are, “The *thorn* can wake the *kettle*” and “His *hand* should doubt the *line*.” Italicized words are key words that are scored during speech recognition testing. This type of speech material was used because the sentences have the flow of connected speech but each key word must be recognized individually and cannot be determined from the semantic context of the sentence. The sentences were spoken in an audiometric sound room (IAC 1604) by a college-aged female native speaker of standard American English. They were recorded onto digital au-

TABLE II. Long-term average one-third-octave spectra for the target, the talker interference (“Talker”), and the speech-spectrum noise interference (“SSN”). Values in the table are given in decibels relative to the A-weighted level.

	One-third octave center frequency (Hz)														
	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300
Target	-13	-12	-1	-6	-5	-8	-9	-14	-17	-20	-17	-21	-24	-25	-25
Talker	-7	-12	-3	-4	-5	-9	-12	-14	-13	-16	-18	-18	-22	-25	-30
SSN	-15	-15	-13	-12	-11	-11	-10	-10	-10	-10	-11	-12	-14	-16	-18

diotape and then replayed, low-pass filtered at 8.5 kHz, and sampled at 20 kHz using a 16-bit analog-to-digital converter (TDT AD1). The resulting digital waveforms were displayed on a computer monitor and were examined visually and auditorily for artifacts such as excessive noise or peak clipping that would require replacement of the sentence. The sentences were divided on an arbitrary basis into 16 lists of 20 sentences (60 scoring words) each.

Two types of interference were recorded on digital audiotape. One was speech-spectrum noise recorded from an audiometer (GS 16). The other was a recording by a second female college student of a different set of 25 nonsense sentences. These speech waveforms were subjected to broadband automatic gain control at the time of recording to minimize variations in peak SPLs when they were presented in the anechoic room. The taped waveforms were low-pass filtered and digitized in a manner similar to the target waveforms described above. Periods of silence before and after each sentence were removed so that the final waveform consisted of an uninterrupted stream of 25 sentences. A continuous 30-min repetition of this stream was recorded on digital audiotape for later playback during the experiments. A long-term one-third-octave band analysis of the target and two types of interferers was conducted using a spectrum analyzer (HP 3569A). For the analysis of the target, 25 sentences, chosen arbitrarily from among those used in the experiments, were concatenated in a similar manner to the talker interference. The results of the analysis, shown in Table II, reveal a reasonable similarity in long-term spectrum between the target and the talker interference, while the speech-spectrum noise interference had a flatter spectral envelope than either of the two talkers.

Target sentences were presented through 16-bit D/A conversion at 20 kHz (TDT DA1), low-pass filtering at 8.5 kHz, programmable attenuation (TDT PA3), summation with the masker when necessary (TDT SUM3), and power amplification (NAD 2100), before delivery to the Realistic loudspeakers. When two-channel presentation was used (conditions 5 and 6) a delayed version of the waveform was presented from a second channel. This was created by pad-

ding 80 zeros (4 ms at 20 kHz) at the beginning of the waveform file. The interference was presented from a single channel of the DAT, attenuated, and fed to a delayer (Klark Teknik DN716), one output of which was delayed by 4 ms relative to the other. The delayer outputs were mixed with the target channels using the TDT SUM3.

Calibration of target and interference was made with a microphone placed at the location of the center of the listeners’ head with the listener absent. The speech stimuli were calibrated based on peak needle movement measured on a sound level meter (B&K 2204) using A weighting and a “fast” rms meter response. The speech-spectrum noise was calibrated to the steady rms A-weighted level. The level of the target was measured to be 47 dBA from either the 0- or 60-degree loudspeaker. Small adjustments in attenuation were applied to the individual sentences to maintain the same presentation level throughout. The 47-dBA presentation level was fixed within and across blocks of 20 sentences. The interference was varied in level from block to block to create four different signal-to-noise ratios for each type of masker. The nominal values of S-N ratio reported in the text and tables refer to the values measured for single sources of target and interference (conditions 1 and 2) and also apply directly to conditions 5 and 6 because a delayed copy of *both* stimuli was added. For conditions 3 and 4, the interference (but not the target) was presented from two loudspeakers, increasing the relative level of the interference. However, no adjustment was made in the labeling of S-N ratios for these conditions.

Table III indicates how the conditions were grouped to form four individual subexperiments with different sets of subjects. Within a subexperiment, the type of interference was fixed, and the loudspeaker conditions either were 1–4 or 5 and 6. In the cases where conditions 1–4 were tested, the combination with four S-N ratios yielded 16 total conditions. A within-subjects Latin square design was used in which each subject listened to all 16 conditions once in a random order, with a different list for each condition. The assignment of lists to conditions was different for each subject. For example, in the first subexperiment listed in Table III, subject 1

TABLE III. Conditions used in experiment 1.

Loudspeaker conditions	Type of interference	S-N ratios (dB)	Conditions/ subjects	Lists per condition per subject	Total scored key words per condition
1,2,3,4	Speech noise	-9,-6,-3,0	16/16	1	960 (20×3×16)
5,6	Speech noise	-9,-6,-3,0	8/8	2	960 (40×3×8)
1,2,3,4	Female talker	-12,-8,-4,0	16/16	1	960 (20×3×16)
5,6	Female talker	-12,-8,-4,0	8/16	2	1920 (40×3×16)

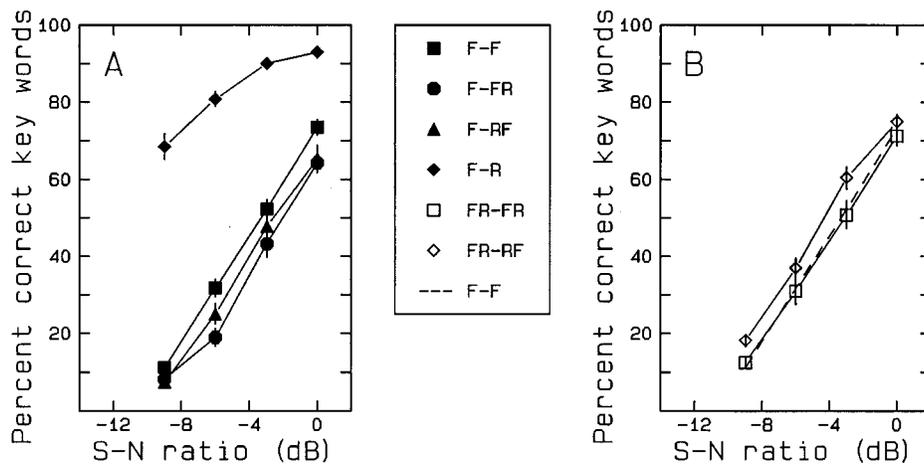


FIG. 2. Mean percent correct and \pm one standard error for the speech-spectrum noise interference in experiment I. Panel (a) shows the results for the front-only target, panel (b) for the FR target.

might have initially been presented with list 4 in loudspeaker condition 1 with a -6 -dB S-N ratio. The next block could have been list 5 presented in loudspeaker condition 4 at a 0 -dB S-N ratio, and so on. This design eliminated the difficulties inherent in creating lists that are reliably equivalent across stimulus conditions, because the data set for each condition consisted of all 320 sentences, or 960 scored items (320 sentences \times 3 key words per sentence). A similar design was used when conditions 5 and 6 were presented, although there were eight total stimulus conditions (2 loudspeaker conditions \times 4 S-N ratios). Each of eight subjects listened to two 20-sentence lists for each of the eight conditions. However, for the female talker interference the entire experiment was repeated with a different set of eight subjects for a total of 16 because variability seemed to be higher.

Listeners were a total of 56 young college students with self-reported normal hearing. Their participation consisted of one listening session lasting approximately 1 h. Following initial instructions they were given 15 practice trials with several loudspeaker conditions in order to become familiar with the task. During the experiment subjects listened to 16 blocks of 20 sentences with a short break after eight blocks. Individual trials were initiated by a listener's button press. The interference was gated on first followed by a single sentence spoken by the target talker which began 0.6 to 1.2 s later. This delay provided a basis for the subject to attend to the target in conditions where female talker interference was presented from the same location(s) as the target. Because the female talker interference was presented from a continuously running tape, the initial words heard were variable, occurring at any point during any of the 25 sentences. The target and interference ended simultaneously. The listener was instructed to then repeat the target sentence to the best of his/her ability. An experimenter seated approximately 5 ft behind the listener scored the key words.

B. Results

1. Speech-spectrum noise interference

The percentage of correct key words for the speech-spectrum noise interference, displayed in Fig. 2(a), demonstrates the effects of the first four loudspeaker conditions. For the F-F (masker from front) condition, only 11% of words were correctly identified at the poorest S-N ratio. As S-N

ratio increased, performance grew approximately linearly with a slope of about 7% per dB. By contrast, when the masker was presented from the right loudspeaker (the F-R condition), 69% of words were correctly identified at the -9 -dB S-N ratio, and performance grew to above 90% correct at the higher S-N ratios. Using a criterion of 70% correct, the advantage of the masker being presented from the right as opposed to the front was approximately 8.2 dB. This 8.2-dB advantage is consistent with comparable data and predictions in the literature. For example, it is within 1 dB of the 9-dB advantage measured by Bronkhorst and Plomp (1988) with the target speech at 0 degrees and a speech-spectrum noise interference at 60 degrees. It is also within 1 dB of the predictions by Zurek (1993, Figure 15.4) for a 60-degree separation. This adds additional support for Zurek's (1993) model; i.e., the improvement in speech recognition is due to increases in the articulation index which result from greater in-band detectability created by head shadow and binaural interaction effects.

Also displayed in Fig. 2(a) is the percent correct performance for the conditions with the two source maskers (F-FR and F-RF). Both functions fall slightly below the front-only masker data (F-F), probably because of the additional masker energy resulting from the presentation of the masker from two loudspeakers (which did not affect the reported S-N ratio, as described in Sec. A.2). There was little difference in results for the RF and FR maskers, suggesting that the perceived location of the masker in relation to the signal was not important. Thus, although the RF masker was heard well to the right of the target (due to the precedence effect), while the FR masker appeared to be close to the target, this had no major effect on speech recognition.

Figure 2(b) displays the results for conditions 5 and 6 where the target talker was FR and the speech-spectrum interference was either FR or RF. Not surprisingly, the FR-FR data are highly similar to the F-F data from the left panel, which are replotted in the right panel as a dashed line. Compared to the FR-FR data, the results for FR-RF are 5% to 10% better at each S-N ratio, resulting in a horizontal shift of less than 1 dB. Thus, despite the fact that the FR masker is perceived to be right on top of the target and the RF masker is perceived to be well separated from the target, these two configurations do not create large differences in masking, at

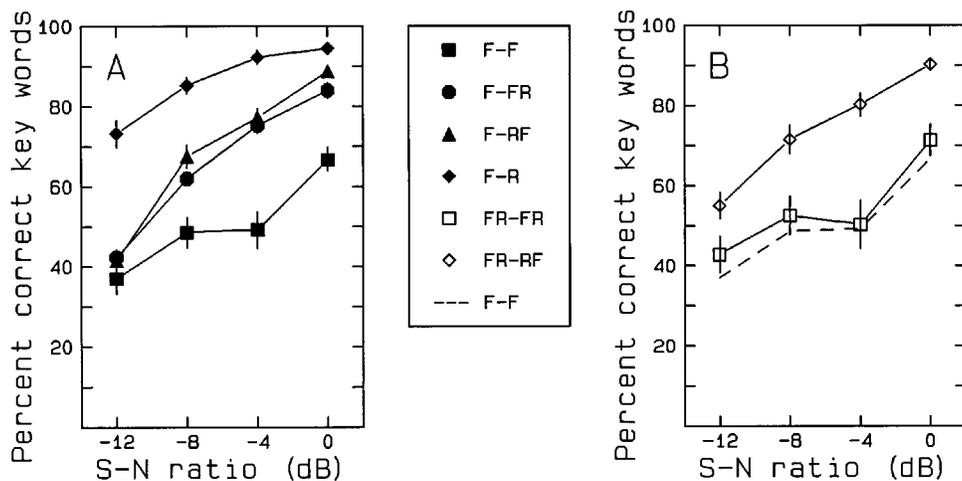


FIG. 3. Mean percent correct and \pm one standard error for the female-talker interference in experiment I. Panel (a) shows the results for the front-only target, panel (b) for the FR target.

least when the target is speech and the masker is a steady noise. This condition may therefore serve as a baseline for considering perceptual separation advantages with maskers that contain a substantial informational component and are highly confusable with the target.

2. Female talker interference

The speech recognition data for the female talker interference are substantially different from the data obtained for the speech-spectrum interference. The data for the front-only targets, shown in Fig. 3(a), show shallow growth in performance with increasing S-N ratio. For example, in the F-F condition, the percentage of words correctly identified improved by only approximately 30 percentage points over a 12-dB range, in contrast to the 60 percentage points over a 9-dB range obtained for the speech-spectrum interference. It follows that recognition performance, which is reasonably similar at the 0-dB S-N ratio across the two types of interference (comparing Figs. 2 and 3), is much better for the talker interference at low S-N ratios. The more gradual slope obtained with the talker interference is consistent with earlier data on speech or speechlike maskers (Dirks and Bower, 1969; Festen and Plomp, 1990). In the case of Dirks and Bower (1969), the psychometric function for speech recognition in single talker interference extended down to -40 -dB S-N ratio for some conditions. The difference in the slopes of the talker and noise interference functions in the current study is also consistent with differences between maskers found for nonspeech stimuli by Kidd *et al.* (1998). They found much shallower masking functions for their informational masker than for their energetic masker.

Performance in the F-R condition was again much better than in the F-F condition. The difference in S-N ratio at 70% correct was at least 12 dB; with small linear extrapolations of both the F-F and F-R functions it is estimated to be 13.7 dB. This is in contrast to the 8.2-dB difference found for the speech-spectrum interference. The larger effect of spatial separation, we propose, is due to the fact that the female talker produces informational masking as well as energetic masking, and spatial separation creates an additional release from this type of masking. When the target speech is separated spatially from this type of interference, the listener is

assumed to benefit from head shadow and binaural interaction advantages and, in addition, from the fact that the interference is *perceived* to be well separated from the target.

The similarity of the data for the FR and RF maskers, also shown in Fig. 3(a), seems inconsistent with the above interpretation, at least on first examination. Due to the precedence effect, the FR masker is perceived to be near the front, close to the location of the target, while the RF masker appears well to the right. Yet there seems to be little or no advantage of this wider perceptual separation of target and masker. One clue to this inconsistency may be found in the fact that both the FR and RF maskers produced considerably *less* masking of speech performance than the F masker, despite the fact that the overall level of the masker was higher in the two-loudspeaker conditions. The additional overall energy provided by the second loudspeaker decreased performance slightly for the speech-spectrum noise interference, as seen in Fig. 2(a), but just the opposite was seen for the female talker interference. This difference may be explained by considering that the existence of the precedence effect does not imply that the delayed sound has no effect. First, the FR masker is probably heard nearly but not exactly from directly in front. Using values of the weighting of lead and lag reported by Shinn-Cunningham *et al.* (1993), the centroid of the image is expected to be shifted 6 to 12 degrees to the right due to the small influence of the delayed right loudspeaker. Even small shifts have been shown to improve speech recognition in previous studies (Speith *et al.*, 1954; Dirks and Wilson, 1969). The addition of the sound from the delayed loudspeaker also affects timbre and produces a bigger, more spacious image than that which occurs with a single source sound (Blauert, 1983; Bech, 1998). Any of these subtle differences might have been useful to the subject in segregating the target and interference. Carhart *et al.* (1969) reported that interaurally delayed maskers that produced distinct images and clear intracranial separation between target speech and talker interference produced no more, and sometimes less release from masking than interaurally phase-reversed maskers that created diffuse images and less pronounced spatial separation from the target. In the current study, it appears that the FR masking condition provides sufficient differences between target and interferer to

allow optimal performance and the greater spatial separation in the RF condition does not enhance performance. These perceptual differences are apparently not useful for the speech-spectrum noise interference, which is assumed to produce no informational masking.

The two conditions which employed the FR target provide a better test of whether the difference in perceived location of the FR and RF interference was important. This is because the FR-FR condition clearly did not create the differences in timbre, spaciousness, and location of the auditory image that are assumed to have occurred for the F-FR condition. The results for the FR target, displayed in Fig. 3(b), show a large, 4 to 9 dB, improvement in required S-N ratio for the RF masker in comparison to the FR masker. This is in stark contrast to the mere 1-dB difference obtained for the same conditions with the speech-spectrum noise interference [see Fig. 2(b)]. The dashed line in Fig. 3(b) is a replotting of the F-F results from the left panel. Because one would expect F-F and FR-FR to be the same, the similarity again demonstrates the reliability of the measurement (see also the dashed line in Fig. 2(b)). This also suggests that the lack of change in performance that occurred between -8 and -4 dB for these conditions is more than a random occurrence. One possible explanation is that informational masking was actually less at -8 dB S-N ratio than at -4 dB and canceled the effect of greater energetic masking. At the poorer S-N ratios, the difference between the levels of the target and interference could have helped listeners segregate the two talkers (see also Egan *et al.*, 1954; Dirks and Bower, 1969). In summary, the FR-FR vs FR-RF loudspeaker presentation conditions that produced little difference in intelligibility for the speech-spectrum interference produced large differences for the female talker interference. This suggests that the benefits of spatial separation of target and masker can be large even in an environment with reflections, if a masker produces informational as well as energetic masking.

The design of the experiments and interpretations of the results to this point have made the assumption that the speech-spectrum noise interference produces purely energetic masking. Because the speech and noise are so dissimilar, the speech is not confused with the noise, only masked by it. When target and interference are separated, binaural interaction and head shadow effects produce an advantage, but there is presumably no additional advantage to having the target and interference appear to be in different places. This type of continuous noise interference can therefore serve as a baseline for studying maskers with informational components. The assumption is supported by the agreement between the size of the advantage for F-R vs F-F [Fig. 2(a)] and Zurek's (1993) predictions based on head shadow and binaural interaction. However, there is uncertainty about precisely what is predicted for the FR and RF maskers with either the F or FR targets. Predicting masking level differences for such stimuli is possible, in theory, but is not straightforward and is untested. We determined that it was necessary to measure the detectability of signals in the presence of these maskers so that direct predictions of the data based on energetic masking could be made.

II. EXPERIMENT II: FREE FIELD DETECTION THRESHOLDS

In this experiment, detection thresholds for narrow-band noises with 15 center frequencies were obtained in the presence of the speech-spectrum noise used in experiment I. Thresholds were obtained for all six of the loudspeaker conditions used in the speech recognition experiments.

A. Methods

The signals were bursts of digitally synthesized narrow-band noise, 200 ms in duration (including 20-ms linear rise/fall periods), and one-third octave in bandwidth, with center frequencies (CFs) of 250, 315, 400, 500, 630, 800, 1000, 1250, 1600, 2000, 2500, 3150, 4000, 5000, and 6300 Hz. The bursts were created by low-pass filtering two independent Gaussian noises with a cutoff frequency of $0.1155 \times \text{CF}$ (one-half of the third-octave bandwidth), multiplying one of the noises by a sine waveform at the CF and the other by a cosine waveform at the CF, and summing the resultants. The spectra of the signals created this way were verified using a spectrum analyzer. Ten independent samples of each signal were stored on a computer. The masker was the speech spectrum noise used in experiment I.

During the collection of the threshold data, the speech spectrum noise masker was presented continuously at an overall level of 47 dBA per loudspeaker as measured with a microphone at the position at the center of the listeners' head (with the listener absent). Calibration of the signals followed the same procedure. Detection thresholds for the signals were estimated using a four-alternative-forced-choice adaptive tracking procedure. On each trial, the signal was presented in one of four temporal intervals marked by lights on a response box. The signal interval was chosen randomly from trial to trial. After recording their responses, subjects were provided with feedback on each trial. The signal level varied adaptively according to a two-hits down, one-miss up stepping rule, which tracks 70.7% correct performance (Levitt, 1971). The initial presentation level was well above threshold and the initial step size was 16 dB. The step size decreased with subsequent reversals in the adaptive tracking, reaching a minimum of 2 dB. Threshold was taken as the average of the SPLs existing on the last six of ten total reversals. Each of three normal-hearing subjects (confirmed by audiometry) obtained three such threshold estimates for each of the 15 CFs.

The six loudspeaker conditions used in experiment I were also used here. For the purposes of data collection the conditions were paired as shown in Table I, with conditions 5 and 6 run first, followed by 1 and 2 and then 3 and 4. In a block of trials all 15 frequencies were tested in a different random order for each block. Three blocks were run for each loudspeaker condition, six for each pair of conditions. Within the six blocks required for each pair, the two conditions were interspersed (e.g., 5, 6, 6, 5, 6, 5). These orders varied randomly from subject to subject.

B. Results

The mean data from the three subjects are displayed in Table IV. As expected with detection of narrow-band signals

TABLE IV. Mean thresholds from three subjects for one-third-octave bands of noise in the presence of a speech spectrum noise background. The change in articulation index relative to the F-F condition was computed for each of the other conditions by (1) subtracting each threshold from the corresponding threshold for the F-F condition, (2) dividing the difference by 30 dB, (3) multiplying the resultant by the articulation index weights for "average speech" from Pavlovic (1987), and (4) summing the resultants across 15 frequency bands. The approximate S-N ratio advantage for each condition was then computed by multiplying the change in articulation index by 30 dB.

One-third octave center frequency	Mean thresholds in dB SPL by loudspeaker presentation condition						Average speech weights
	F-F	F-R	F-FR	F-RF	FR-FR	FR-RF	
250	18.97	14.77	22.73	22.32	17.75	18.21	0.0150
315	22.18	15.91	22.25	23.37	22.69	18.28	0.0289
400	24.10	15.87	21.27	21.51	24.30	19.24	0.0440
500	25.90	15.58	23.43	24.98	24.20	18.65	0.0578
630	28.24	20.32	27.08	27.28	29.30	24.54	0.0653
800	33.99	25.95	31.05	32.69	33.71	31.14	0.0711
1000	33.99	27.48	34.17	34.32	33.20	32.08	0.0818
1250	36.20	29.66	37.13	36.49	34.51	34.75	0.0844
1600	38.24	29.58	40.48	42.03	37.26	36.70	0.0882
2000	37.53	24.17	37.82	36.44	37.30	35.65	0.0898
2500	33.84	20.46	34.69	34.36	34.46	32.61	0.0868
3150	31.61	20.72	33.42	33.73	31.31	30.24	0.0844
4000	28.58	17.35	28.04	28.35	28.03	26.23	0.0771
5000	21.17	11.61	22.57	23.02	24.86	23.76	0.0527
6300	21.03	12.84	21.65	23.31	21.24	21.26	0.0364
Computed AI <i>re</i> F-F (dB)		0.303	-0.003	-0.015	0.006	0.068	
Predicted S-N ratio <i>re</i> F-F		9.09	-0.10	-0.44	0.17	2.04	

in broadband noise, good consistency was observed across the three runs for each subject and across subjects. Standard errors (not shown) were low, typically in the 1- to 2-dB range. Not surprisingly, the F-F versus F-R difference was the largest by far, as large as 13 dB at some frequencies, and encompassed the entire frequency range. In comparison, the FR-FR versus FR-RF difference was much less, averaging about 5 dB in the low frequencies through 630 Hz, and approximately 1 dB at 1000 Hz and above. The purpose of these latter conditions was to create perceptual separation while minimizing the head shadow and binaural interaction advantages. It is clear from the results that detection advantages were reduced considerably with this set of conditions, although they were not eliminated. The small differences in high-frequency thresholds suggest that intensity differences due to head shadow were reduced to near zero (see also Fig. 1), but some phase differences at low frequencies were sufficient to create release from masking of 4 to 6 dB. This release from masking was not as large as what was obtained for F-R versus F-F conditions. The F-RF thresholds were only minimally different from the F-FR thresholds. Both were also very close to the F-F data. The similarity of the thresholds for the F-F, F-FR, and F-RF conditions is consistent with the finding that the speech recognition scores were very similar across these conditions for the speech spectrum interference (see Fig. 2).

The thresholds obtained at the 15 frequencies were used to predict changes in S-N ratio required for criterion performance using the articulation index (AI) (Kryter, 1962). With the thresholds for the F-F condition as the baseline, thresholds for the other conditions were used to calculate the change in articulation index for each condition with the speech-spectrum interference.² Weights for each one-third

octave were taken from values for "average speech" from Pavlovic (1987).³ The difference in AI was multiplied by 30 dB to convert it to an approximate S-N ratio difference relative to the F-F condition for an equivalent level of performance. In addition to the mean masked thresholds described above, Table IV also includes the AI weights, predicted AI differences, and predicted S-N ratio differences. It is worth noting that the 9-dB advantage predicted for F-R versus F-F is very close to Zurek's (1993) predictions for the 60-degree separation, and the predicted advantage is reduced to 2 dB when simulated reflections are added for both signal and masker (FR-RF).

Figure 4 displays the predicted spatial separation advan-

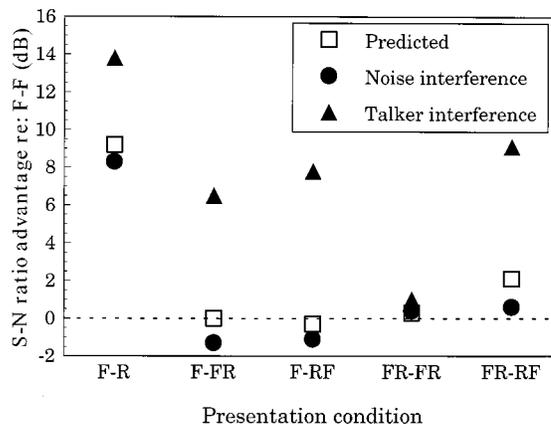


FIG. 4. Predicted and actual advantages in S-N ratio for criterion speech recognition performance. Predictions are based upon thresholds obtained in experiment II in combination with articulation index weighting of frequency bands, and are also shown in Table IV. Actual advantages in S-N ratio are based on 70% correct performance using linear interpolation (and small linear extrapolation, when necessary) of the functions in Figs. 2 and 3.

tages taken from the bottom row of Table IV together with the actual advantages measured from the data. These actual advantages were estimated by linearly interpolating and extrapolating the data in Figs. 2 and 3 to find the S-N ratio corresponding to 70% correct for each condition and subtracting the S-N ratio obtained for the F-F condition from each of the others. This was done separately for the speech spectrum interference and the female talker interference. The criterion of 70% correct was used because the majority of the functions passed through or near that performance level, so the need to extrapolate was minimal. For the speech-spectrum noise interference, the good correspondence between the predictions and the data demonstrates the utility of the articulation index, in that the difference in performance across conditions with the speech spectrum interference can be accounted for by the differences in binaural detection thresholds that presumably result from head shadow and binaural interaction. For comparison, the filled triangles show the actual differences relative to the F-F data for the female talker interference. The S-N ratio advantages relative to the F-F condition are 5–8 dB better than they are for the speech spectrum noise interference except for the FR-FR condition, where there were no spatial separation cues.

III. DISCUSSION

The findings of the current study are in agreement with previous studies which show that the benefit of spatially separating speech and a *steady* background noise is severely reduced in a reverberant environment relative to the anechoic situation (e.g., Koehnke and Besing, 1996). The advantage of separating the original sources of target speech and speech-spectrum noise interference by 60 degrees, measured as the difference in S-N ratio required for 70% correct, was reduced from 8 dB to 1 dB or less by adding a single simulated ‘‘reflection’’ for target and masker. The reflection disrupted the interaural differences that produce release from masking in an anechoic environment. As shown in Fig. 4, the reduced advantage of spatial separation on speech recognition was predictable from the reduction in release from energetic masking observed in experiment II. There appeared to be no additional advantage from the fact that speech and noise were perceived to originate from different locations.

Relative to the speech-spectrum noise interference, the advantages of spatial separation were considerably greater when the interference was a single talker of the same sex. In the anechoic condition, the spatial separation advantage for the female talker interference was approximately 14 dB, compared with 8 dB obtained for the speech-spectrum interference. In the simulated reflection condition, the spatial separation advantage was approximately 9 dB for the female talker interference compared with less than 1 dB for the speech-spectrum noise interference.

Substantial differences between the two types of interference complicate the interpretation of the differences in results. In addition to the fact that one was noise and the other was speech, the interferers were different in their long-term average spectrum and, perhaps more importantly, in their short-term spectral and amplitude fluctuations. However, the previous literature indicates that maskers with fluc-

tuations are not necessarily associated with unusually large advantages of spatial separation in speech recognition (e.g., Duquesnoy, 1983; Bronkhorst and Plomp, 1992). For example, using a noise masker modulated by the envelope fluctuations of a single talker, Bronkhorst and Plomp (1992) found only an 8-dB advantage in speech recognition threshold when the masker was moved from 0 degrees (the target position) to a location 90 degrees to the right. This 8-dB difference obtained with a fluctuating noise is well predicted by Zurek’s (1993) model, even though the model was intended for steady background noise conditions.

While the considerable differences between the maskers in this study make it impossible to be certain about the reasons for the differences in the data, we hypothesize that the larger spatial separation advantages observed for the talker interference were because that type of interference produced informational as well as energetic masking. When target and interference were both speech, spatial separation helped to segregate the target message from the interfering message. This interpretation is consistent with the substantial benefit of spatial separation found with nonspeech informational maskers (Kidd *et al.*, 1998). There may be practical implications if the larger benefit of spatial separation with informational maskers turns out to be a general finding. Specifically, it may be sensible to consider using such maskers to evaluate technology intended to improve spatial hearing (e.g., binaural and multi-microphone hearing aids).

The generality of the finding of a larger benefit of spatial separation with the talker interference is a question that must be answered by future research. The results with the female talker interference are an example, not a global description, of how spatial separation of target and interference can improve speech recognition. Presumably, there exist talkers whose speech is more difficult to segregate from the target, and other talkers whose speech is less disruptive. The advantages of spatial separation are likely to vary depending on talker characteristics. Using male target and interfering talkers, Plomp (1976) did not find a larger advantage of spatial separation compared with a condition where speech-spectrum noise was used as the interference. However, the 0-degree target, 0-degree interference condition used two separate loudspeakers, so there was a small amount of spatial separation even in that condition. This might be an important difference from the current study. The effect of two or more interfering talkers should be considered as well. Yost *et al.* (1996) found that spatial separation advantages were greater when there were a total of three talkers than when there were two.

The sex of the target talker relative to the interference is also likely to be an important variable. Duquesnoy (1983) used a female talker as the target and a male talker or a continuous noise as the interference. Both masker conditions benefitted from spatial separation in an anechoic environment. However, unlike the current data from our F-R condition, Duquesnoy (1983) did not find a greater effect of spatial separation for the speech masker than for the continuous noise masker. For the speech masker, that study found that the S-N ratio required for 50% correct sentence recognition in the F-F condition was -17.6 dB, which is considerably

lower than the current result of approximately -8 to -4 dB (see Fig. 3). Perhaps the cues to distinguish between the two talkers were so obvious that it was easy to attend to the target in the presence of the interference and, therefore, there was little release from informational masking in the spatial separation conditions.

Interestingly, Duquesnoy (1983) found that presenting the talker interference backwards had no effect on the amount of masking for either the F-F or spatially separated conditions. Hygge *et al.* (1992) reported a similar finding with a female talker target and spatially separated male talker interference. However, Dirks and Bower (1969) and Sperry *et al.* (1997) showed that presenting the masker backwards in monaural conditions did have an effect when target and interference contained talkers of the same sex. In the Dirks and Bower (1969) study, target and interference were the same male talker. When the interference was presented normally, but not backwards, a plateau was observed in the performance versus intensity function in the region just below 0-dB S-N ratio. Speech recognition performance did not improve as the S-N ratio increased from -10 to 0 dB. With similar conditions, Egan *et al.* (1954) observed the same type of discontinuity in the region of -8 to -4 dB S-N ratio. These results are similar to the plateau in the current experiment between the -8 and -4 -dB S-N ratio when target and interference were not spatially separated (F-F and FR-FR in Fig. 3). It appears that segregation of two talkers without spatial separation is particularly difficult when the two are similar in level, but this difficulty disappears when other cues are available to distinguish the target from the interference (e.g., different sex of target and interference, or interference presented backwards).

Why does perceived spatial separation appear to only aid performance when the stimuli are highly confusable? When two streams of information come in simultaneously, the subject must decide which words to concentrate on and decipher. Unless some additional cue is provided, the two messages are difficult to segregate. In our study the cue was a spatial separation of the two talkers; what is remarkable is that the separation could be an illusion, created by the precedence effect. In a recent study, Driver (1996) furnished an illusory visual cue, using the ventriloquist effect. Driver (1996) showed that it was easier to distinguish target and competing words spoken simultaneously by the same talker if a visual display of the talker producing the target words was separated spatially from the loudspeaker delivering both sets of words. In this case a release from informational masking occurred in the absence of acoustic differences because the displaced visual display pulled out the target words to its spatial location. As in the current experiment, the benefit was created by perceived differences in target and distracter locations. Both experiments illustrate how an illusory spatial separation can help direct attention to the designated target sounds.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute of Deafness and other Communicative Disor-

ders (DC01645), and a Research Scientist Award (MH00332) to RKC. The authors would like to thank Carrie Band, Amie Bavosi, and Meghan Phelan for their assistance in data collection for this project.

¹Informal listening by the experimenters to all the conditions verified that the localizations of target and interference were in the vicinity of the front and right loudspeakers as specified in the right columns of Table I. Formal testing may have revealed small shifts in perceived location toward the lag loudspeaker for the precedence effect stimuli, as predicted by Shinn-Cunningham *et al.* (1993) and Litovsky and Macmillan (1994).

²The computation of changes in the articulation index based on changes in threshold requires the assumption that all of the audible speech spectrum is within the dynamic region of the AI, i.e., within the 0- to 30-dB range relative to the noise (-12 to $+18$ dB for rms-to-rms comparison). We used the long-term spectrum measurements shown in Table II to check the validity of this assumption at the point of comparison (70% correct). Out of 90 possibilities (15 frequencies for six conditions), the S-N ratio fell out of this range only twice, and then only by 1 dB. Therefore, we believe that our assumptions concerning the use of the AI are reasonable.

³The "average speech" AI computations include also the frequencies of 160, 200, and 8000 Hz. However, the weightings of these extreme frequencies are very low. For efficiency we limited the threshold testing to 15 frequencies, for which the articulation index weightings sum to more than 0.96 out of a possible 1.0.

- Bech, S. (1998). "Spatial aspects of reproduced sound in small rooms," *J. Acoust. Soc. Am.* **103**, 434-445.
- Blauert, J. (1983). *Spatial Hearing* (MIT, Cambridge, MA).
- Bronkhorst, A. W., and Plomp, R. (1988). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," *J. Acoust. Soc. Am.* **83**, 1508-1516.
- Bronkhorst, A. W., and Plomp, R. (1992). "Effect of multiple speech like maskers on binaural speech recognition in normal and impaired hearing," *J. Acoust. Soc. Am.* **92**, 3132-3139.
- Carhart, R., Tillman, T. W., and Greetis, E. S. (1969). "Release from multiple maskers: Effects of interaural time disparities," *J. Acoust. Soc. Am.* **45**, 411-418.
- Dirks, D. D., and Bower, D. R. (1969). "Masking effects of speech competing messages," *J. Speech Hear. Res.* **12**, 229-245.
- Dirks, D. D., and Wilson, R. H. (1969). "The effect of spatially separated sound sources on speech intelligibility," *J. Speech Hear. Res.* **12**, 650-664.
- Doll, T. J., and Hanna, T. E. (1997). "Directional Cuing effects in auditory recognition," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. B. Anderson (Erlbaum, Hillsdale, NJ).
- Driver, J. (1996). "Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading," *Nature (London)* **381**, 66-68.
- Duquesnoy, A. J. (1983). "Effect of a single interfering noise and interfering speech upon the binaural sentence intelligibility of aged persons," *J. Acoust. Soc. Am.* **74**, 739-743.
- Egan, J. P., Carterette, E. C., and Thwing, E. J. (1954). "Some factors affecting multi-channel listening," *J. Acoust. Soc. Am.* **26**, 774-782.
- Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725-1736.
- French, N. R., and Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90-119.
- Gilkey, R. H., and Anderson, T. B. (1997). *Binaural and Spatial Hearing in Real and Virtual Environments* (Erlbaum, Hillsdale, NJ).
- Good, M. D., Gilkey, R. H., and Ball, J. M. (1997). "The relation between detection in noise and localization in noise in the free field," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. B. Anderson (Erlbaum, Hillsdale, NJ).
- Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (1999). "Speech intelligibility and localization in a multi-source environment," *J. Acoust. Soc. Am.* **105**, 3436-3448.
- Helper, K. S. (1997). "Auditory and auditory-visual perception of clear and conversational speech," *J. Speech Lang Hear Res.* **40**, 432-443.
- Hirsh, I. J. (1950). "Relation between localization and intelligibility," *J. Acoust. Soc. Am.* **22**, 196-200.

- Hygge, S., Ronnberg, J., Larsby, B., and Arlinger, S. (1992). "Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds," *J. Speech Hear. Res.* **35**, 208–215.
- Kidd, Jr., G., Mason, C. R., and Rohla, T. L. (1995). "Binaural advantage for sound pattern identification," *J. Acoust. Soc. Am.* **98**, 1977–1986.
- Kidd, Jr., G., Mason, C. R., Rohla, T. L., and Deliwala, P. S. (1998). "Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns," *J. Acoust. Soc. Am.* **104**, 422–431.
- Kidd, Jr., G., Mason, C. R., Deliwala, P. S., Woods, W. S., and Colburn, H. S. (1994). "Reducing informational masking by sound segregation," *J. Acoust. Soc. Am.* **95**, 3475–3480.
- Kock, W. E. (1950). "Binaural localization and masking," *J. Acoust. Soc. Am.* **22**, 801–804.
- Koenig, W. (1950). "Subjective effects in binaural hearing," *J. Acoust. Soc. Am.* **22**, 61–62.
- Koenig, A. H., Allen, J. B., and Berkley, D. A., and Curtis, T. H. (1977). "Determination of masking-level differences in a reverberant environment," *J. Acoust. Soc. Am.* **61**, 1374–1376.
- Koehnke, J., and Besing, J. M. (1996). "A procedure for testing speech intelligibility in a virtual listening environment," *Ear Hear.* **17**, 211–217.
- Kryter, K. D. (1962). "Methods for calculation and use of the Articulation Index," *J. Acoust. Soc. Am.* **34**, 1689–1697.
- Leek, M. R., Brown, M. E., and Dorman, M. F. (1991). "Informational masking and auditory attention," *Percept. Psychophys.* **50**, 205–214.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Levitt, H., and Rabiner, L. R. (1967a). "Binaural release from masking for speech and gain in intelligibility," *J. Acoust. Soc. Am.* **42**, 601–608.
- Levitt, H., and Rabiner, L. R. (1967b). "Predicting binaural gain in intelligibility and release from masking for speech," *J. Acoust. Soc. Am.* **42**, 820–829.
- Litovsky, R. Y., and Macmillan, N. A. (1994). "Sound localization precision under conditions of the precedence effect: effects of azimuth and standard stimuli," *J. Acoust. Soc. Am.* **96**, 752–758.
- MacKeith, N. W., and Coles, R. R. A. (1971). "Binaural advantages in the hearing of speech," *J. Laryngol. Otol.* **85**, 213–232.
- Pavlovic, C. V. (1987). "Derivation of primary parameters and procedures for use in speech intelligibility predictions," *J. Acoust. Soc. Am.* **82**, 413–422.
- Peissig, J., and Kollmeier, B. (1997). "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," *J. Acoust. Soc. Am.* **101**, 1660–1670.
- Plomp, R. (1976). "Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise)," *Acustica* **34**, 200–211.
- Plomp, R., and Mimpfen, A. M. (1981). "Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech reception threshold for sentences," *Acustica* **48**, 325–328.
- Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., and Perrott, D. R. (1991). "Free field release from masking," *J. Acoust. Soc. Am.* **90**, 1355–1370.
- Shaw, E. A. G. (1974). "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," *J. Acoust. Soc. Am.* **56**, 1848–1861.
- Shinn-Cunningham, B. G., Zurek, P. M., and Durlach, N. I. (1993). "Adjustment and discrimination measurements of the precedence effect," *J. Acoust. Soc. Am.* **93**, 2923–2932.
- Speith, W., Curtis, J. F., and Webster, J. C. (1954). "Responding to one of two simultaneous messages," *J. Acoust. Soc. Am.* **26**, 391–396.
- Sperry, J. L., Wiley, T. L., and Chial, M. R. (1997). "Word recognition performance in various background competitors," *J. Am. Acad. Audiol.* **8**, 71–80.
- Watson, C. S., Kelly, W. J., and Wroton, H. W. (1976). "Factors in the discrimination of tonal patterns: II. Selective attention and learning under various levels of stimulus uncertainty," *J. Acoust. Soc. Am.* **60**, 1176–1185.
- Yost, W. A., Dye, R. H., and Sheft, S. (1996). "A simulated 'cocktail party' with up to three sound sources," *Percept. Psychophys.* **58**, 1026–1036.
- Zurek, P. M. (1987). "The precedence effect," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York).
- Zurek, P. M. (1993). "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, 2nd ed., edited by G. A. Studebaker and I. Hockberg (Allyn and Bacon, Needham Heights, MA).

Auditory localization of nearby sources. III. Stimulus effects

Douglas S. Brungart^{a)}

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 29 March 1999; revised 2 July 1999; accepted 13 August 1999)

A series of experiments has examined the auditory localization of a nearby (<1 m) sound source under four conditions: (1) a fixed-amplitude condition where loudness-based distance cues were available; (2) a monaural condition where the contralateral ear was occluded by an ear-plug and muff; (3) a high-pass condition where the stimulus bandwidth was 3 Hz to 15 kHz; and (4) a low-pass condition where the stimulus bandwidth was 200 Hz to 3 kHz. The results of these experiments were compared to those of a previous experiment that measured localization performance for a nearby broadband, random-amplitude source [Brungart *et al.*, *J. Acoust. Soc. Am.* **106**, 1956–1968 (1999)]. Directional localization performance in each condition was consistent with the results of previous far-field localization experiments. Distance localization accuracy improved slightly in the fixed-amplitude condition relative to the earlier broadband random-amplitude experiment, especially near the median plane, but was severely degraded in the monaural condition. Distance accuracy was also found to be highly dependent on the low-frequency energy of the stimulus: in the low-pass condition, distance accuracy was similar to that in the broadband condition, while in the high-pass condition, distance accuracy was significantly reduced. The results suggest that low-frequency interaural level differences are the dominant auditory distance cue in the proximal region. [S0001-4966(99)01712-9]

PACS numbers: 43.66.Qp, 43.66.Pn [DWG]

INTRODUCTION

The problem of identifying the location of a sound source in space from auditory information has been studied extensively, but until recently little attention has been given to the unique aspects of localization in the immediate vicinity of the head. In the “distal region,” which we define as the region greater than 1 m from the head, the acoustic localization cues represented by the head-related transfer function (HRTF) are roughly independent of distance. In contrast, in the “proximal region” (within 1 m of the head), these localization cues change systematically with distance. This study extends a previous study of proximal-region localization with a broadband source (Brungart, Durlach, and Rabinowitz, 1999) to a variety of conditions, including fixed stimulus amplitude, monaural listening, and high-pass and low-pass filtered stimuli. The results are compared to the broadband condition, and to previous results from far-field localization experiments under similar conditions.

I. BACKGROUND

The important role that the frequency content of the source plays in auditory localization has been recognized from the earliest days of auditory localization research. Lord Rayleigh’s famous duplex theory of localization (1907) is based on the idea that different mechanisms of localization are dominant depending on the frequency spectrum of the source. At low frequencies (below 1500 Hz), where head shadowing is minimal, the interaural time delay (ITD) is the

primary localization cue for horizontal-plane sources. At higher frequencies, where the phase difference between the ears is ambiguous, interaural level differences caused by head shadowing dominate horizontal-plane localization. More recently, researchers have noted that the complex pattern of folds in the outer ear provides high-frequency spectral localization cues which are important in determining the elevation of a source (Batteau, 1967; Roffler and Butler, 1968; Gardner and Gardner, 1973; Butler *et al.*, 1980) and resolving front-back confusions (Musicant and Butler, 1984b; Oldfield and Parker, 1984).

The different localization mechanisms function in different frequency ranges, so the spectral content of a stimulus has a strong influence on localization accuracy. When the source contains only low-frequency energy, localization relies almost entirely on interaural time delays (ITDs) (Wightman and Kistler, 1992). Under these conditions, the ITD provides accurate information about the lateral position of the source, but cannot be used to distinguish between sources at the same lateral position in the front and rear hemispheres or to determine the elevation of a sound source. Musicant and Butler (1984b) found that the number of front-back reversals increased substantially when the stimulus was low-pass filtered at 4 kHz and dramatically when the stimulus was low-pass filtered at 1 kHz. Hebrank and Wright (1974) found that subjects were unable to accurately judge the elevation of a sound source in the median plane (fewer than 50% of the responses were within 45 degrees of the source location) without spectral content in the stimulus above 8 kHz. When the source contains only frequencies below 4 kHz or so, localization performance is poor in elevation and the number of front-back reversals increases substantially.

^{a)}Currently at Human Effectiveness Directorate, Air Force Research Laboratory. Electronic mail: douglas.brungart@he.wpafb.af.mil

In contrast, localization accuracy is only slightly degraded when the source contains only high-frequency energy and the source is not narrow band. Hebrank and Wright (1974) found that elevation judgments were quite accurate when the stimulus was high-pass filtered above 4 kHz (100% of responses were within 45 degrees of the actual source location and 60% were within 15 degrees). Musicant and Butler (Musicant and Butler, 1984b) reported no significant differences between the accuracy of azimuthal localization with a broadband noise and a noise high-pass filtered above 4 kHz. There is some evidence that subjects are able to obtain interaural time delay information from the envelopes of high-frequency sounds (Henning, 1974; McFadden and Pasanen, 1976; Trahoitis and Bernstein, 1986; Yost *et al.*, 1971). In addition, Musicant and Butler (1984b) argued that the directional characteristics of the pinnae play an important role in azimuthal localization, and found that localization accuracy decreased substantially when the folds of the pinnae were filled with putty.

When binaural information is removed from the stimulus by occlusion or impairment of one of the ears, localization is seriously degraded in azimuth but only slightly impaired in elevation. Azimuthal localization judgments are extremely inaccurate on the side of the occluded ear (Wightman and Kistler, 1997; Slattery and Middlebrooks, 1994) and exhibit a strong lateral bias on the side of the normal ear (Slattery and Middlebrooks, 1994; Butler, 1987). In contrast, elevation accuracy is only slightly degraded under monaural listening conditions (Oldfield and Parker, 1986; Slattery and Middlebrooks, 1994).

Source characteristics also have an influence on distance perception. In particular, the perceived distance of an unfamiliar source can be increased by low-pass filtering the stimulus (Little *et al.*, 1992). While this effect is often explained by the greater absorption of high-frequency sound energy by the atmosphere, a more realistic explanation is that low-frequency sound propagates around obstacles in a room more efficiently than high-frequency sound and consequently relatively more low-frequency energy reaches a listener from a distant source in an enclosed space.

Examining auditory localization under a variety of source conditions has significantly advanced our understanding about the mechanisms of spatial hearing. To date, however, no studies have examined the influence of source spectrum or monaural listening conditions on the localization of sound sources near the head. Although researchers have recognized that interaural level differences are substantially larger in the immediate vicinity of the head than in the distal region (Hartley and Fry, 1921), and have noted the possible importance of proximal-region ILDs as distance cues (Coleman, 1963), only in the past few years has proximal-region localization received serious attention. Recent studies have calculated the proximal-region head-related transfer function for a rigid sphere model of the head (Brungart and Rabinowitz, 1996) and have verified the results with measurements on the surface of a bowling ball (Duda and Martens, 1998). Another study has measured proximal-region HRTFs on a KEMAR manikin with a compact acoustic source (Brungart and Rabinowitz, 1999). The changes that occur in the HRTF

in the proximal region can be summarized as follows:

- (i) The interaural level difference increases dramatically for all frequencies as the source approaches the head.
- (ii) The interaural time delay is roughly independent of distance.
- (iii) Diffraction by the head slightly low-pass filters stimuli located close to the head.
- (iv) The high-frequency features of the HRTF which are dependent on elevation are relatively insensitive to source distance.
- (v) The high-frequency features of the ipsilateral HRTF occur at more lateral locations when the source is near the head than when the source is distant due to an acoustic parallax effect caused by the discrepancy between the angle of the source relative to the head and the angle of the source relative to the ear.

The current study is based on a previous psychoacoustic experiment which examined proximal-region localization for a broadband source (Brungart and Durlach, 1999). In that study, listeners were asked to place a response sensor at the perceived location of a noise burst stimulus generated by a compact, nondirectional source at a random location in the right hemisphere. The major conclusions of that study are summarized below:

- (i) Azimuthal localization accuracy decreased only slightly as the source approached the head.
- (ii) Elevation localization accuracy did not change consistently with source distance.
- (iii) Distance localization was accurate and relatively unbiased for lateral sources but inaccurate in the median plane.

The previous study focused on a general overview of proximal-region auditory localization, and provided a map of localization accuracy as a function of the azimuth, elevation, and distance of the source. This study extends the previous results by measuring relative localization performance under four different listening conditions exploring the consequences of fixing the source amplitude, occluding one ear with an ear-plug and muff, and high-pass and low-pass filtering the stimulus. The results in each of these conditions are compared to results from the broadband condition and to previous far-field experiments under similar conditions. The results are also discussed in relation to the proximal-region characteristics of the HRTF.

II. METHODS

A. Overall procedure

The procedure used in experiment I was identical to the one used in the earlier study (Brungart and Durlach, 1999), but it is summarized here for convenience. This study used the same four right-handed, normal-hearing male subjects who participated in the earlier experiment. During the experiment, the subject was seated on a stool in MIT's anechoic chamber and used a chin-rest to immobilize his head. The experimenter stood in a fixed location approxi-

mately 1.5 m to the right of the subject, and manually placed the sound source prior to each trial. The source, which consisted of an electromagnetic horn driver connected to a long section of tubing, was specifically designed to approximate the compact, nondirectional properties of an acoustic point source (Brungart and Rabinowitz, 1999). The end of the source had a curved shape to allow the experimenter to place the source anywhere in the right hemisphere of the subject without moving from a stationary position. Random numbers, read to the experimenter through an earphone prior to each trial, guided the placement of the source over a range of positions from 0 to -180 degrees in azimuth, -60 to 90 degrees in elevation, and 10 to 100 cm in distance.

A Polhemus 3 space tracker recorded the stimulus and response positions in each trial. The Polhemus source was rigidly mounted on the chin-rest to maintain a constant position relative to the subject's head. One of the two position sensors, mounted near the tip of the point source, measured the stimulus location during each trial. The second position sensor, which was mounted on the end of a 20-cm wooden wand, was moved to the perceived location of the stimulus by the subject and measured the response location. This *direct-location* response method was described fully in an earlier assessment of proximal-region response paradigms (Brungart *et al.*, 1999).

Under most conditions, the stimulus amplitude was normalized to eliminate amplitude-based distance cues. The normalization factor was calculated from the source position prior to each trial according to the formula:

$$\frac{1}{\frac{50}{\text{Distance to left ear (cm)}} + \frac{50}{\text{Distance to right ear (cm)}}}$$

This normalization factor emphasizes the distance from the source to the right ear when the source is close to the head, but includes the contribution of both ears when the source is farther away. In addition to this normalization, the amplitude of the stimulus in the random-amplitude conditions was roved randomly over a 15-dB range (from 0 to 15 dB of attenuation in 1-dB steps).

The overall procedure for each trial can be summarized as follows:

- (1) The subject closed his eyes while the experimenter moved the source to a random location according to the three random numbers generated by the control computer.
- (2) The control computer recorded the location of the point source, determined the proper amplitude normalization factor, and generated the stimulus.
- (3) The experimenter moved the source to a neutral position, and told the subject to open his eyes and move the response sensor to the perceived location of the sound.
- (4) When the subject finished the response, the experimenter pressed the response switch to tell the computer to record the response location, and the subject closed his eyes in preparation for the next trial. Note that the subject was not provided with any feedback.

TABLE I. Stimulus conditions. In the monaural stimulus condition, binaural cues were eliminated by occluding the subject's left ear with an ear-plug and an ear-muff modified to cover only the one ear. The Max SPL is the maximum sound pressure level generated by the stimulus when the roving attenuation level was at 0 dB, measured when the distance to the head was 1 m.

Stimulus	Pass-band (kHz)	Binaural	Roved	Gating	Max SPL (dBA)
Broadband (BB)	0.2–15	y	y	Rect.	59
Fixed amplitude (FA)	0.2–15	y	n	Rect.	51 ^a
Monaural (MO)	0.2–15	n	y	Rect.	59
High-pass (HP)	3–15	y	y	cos ²	51
Low-pass (LP)	0.2–3	y	y	cos ²	66

^aIn the fixed-amplitude condition roving was not used, but the signal was attenuated 8 dB to place its amplitude in the middle of the roved range. The value given was measured at 1 m and, since the amplitude was not normalized for distance, the effective amplitude was greater at locations closer than 1 m.

B. Stimulus

The stimulus consisted of five short (150 ms) bursts of noise, separated by 30-ms intervals. The waveforms were constructed from white Gaussian noise, which was first filtered to flatten the irregular spectral response of the point source within ± 2 dB. The stimulus was then band-pass filtered with the pass-band cutoffs determined by the stimulus condition in Table I and sharp (120 dB/decade) rolloffs in the stop-bands. Finally, the waveforms were low-pass filtered with a 6 dB/octave rolloff above the low-frequency cutoff in order to maximize the output of the point source.¹

In the baseline stimulus conditions, rectangular gating was used, while in the high-pass and low-pass conditions, 50-ms cos² ramps were used to reduce transients. In each stimulus condition, the maximum amplitude (prior to normalization) was chosen to generate the greatest possible non-distorted output from the point source. The source produced low-frequency sound more efficiently than high-frequency sound, resulting in relatively lower output in the high-pass condition. The details of all five stimulus conditions are provided in Table I.

C. Data collection

The data were collected in blocks of 100 trials, with each block requiring approximately 20 min. Four or five blocks of trials were recorded in each 2-h session. The data were collected consecutively for each of the five conditions. First, each subject participated in a few trial blocks to become familiar with the experimental procedure. Next, 2000 trials per subject were collected in the broadband condition of the previous study. Then 500 trials per subject were collected consecutively in the broadband, monaural, high-pass, low-pass, and fixed-amplitude conditions.² The subjects were informed that the source level was constant in the fixed-amplitude condition and, of course, understood the nature of the monaural condition, but only subject DSB was aware of the composition of the high-pass and low-pass stimuli.

TABLE II. Summary of directional localization accuracy in each condition. The overall angular errors were calculated separately for each subject and averaged together. For the azimuth and elevation errors, the data were divided into three bins in azimuth (<-120 degrees, -120 to -60 degrees, and >-60 degrees), three bins in elevation (<-20 degrees, -20 to 20 degrees, and >20 degrees), and three bins in distance (<25 cm, 25 to 50 cm, and >50 cm). The signed errors and BCRMS errors in azimuth were calculated separately for each subject in each distance and elevation bin and averaged together. Similarly, the signed errors and BCRMS errors in elevation were calculated separately for each subject in each distance and azimuth bin and averaged. BB = broadband; FA = fixed amplitude; MN = monaural; HP = high-pass; LP = low-pass.

	BB		FA		MN		HP		LP	
Angular (degrees):	16.6	(0.1)	17.0	(0.2)	29.7	(0.4)	22.6	(0.3)	30.2	(0.6)
Azimuth:										
Signed (degrees)	1.7	(0.2)	2.7	(0.5)	0.3	(0.8)	5.0	(0.5)	1.0	(0.5)
BCRMS (degrees)	12.7		12.2		14.9		13.3		14.7	
Reversals	10%		11%		16%		21%		36%	
Elevation:										
Signed (degrees)	2.5	(0.2)	1.7	(0.3)	0.9	(0.6)	3.7	(0.5)	-10.4	(0.8)
BCRMS (degrees)	13.2		10.5		14.7		15.5		15.2	

Note: Standard errors are shown in parentheses. Trials where front-back reversals occurred were omitted from the calculations of the angular errors and the azimuth errors, but were included in the calculations of elevation errors. The BCRMS errors were averaged together by taking the square root of the mean-squared BCRMS errors. In other words, the average BCRMS value was the square root of the average bias-corrected mean squared errors.

D. Coordinate system

When dealing with locations close to the head, the coordinate system must be chosen carefully. This experiment used a coordinate system based on the geometry of the subject's head. At the beginning of each block of trials, the subject placed his head in a comfortable position in the chinrest and the locations of three reference points were recorded using the response sensor: the opening of the left ear canal, the opening of the right ear canal, and the tip of the nose. These locations were used to correct for stimulus distance and to define a vertical spherical coordinate system based on the subject's head, with its origin at the midpoint of the left and right ears, its "horizontal plane" defined by the locations of the left and right ears and the nose, and its median plane perpendicular to the interaural axis and passing as close as possible to the location of the nose (Brungart *et al.*, 1999). In this coordinate system, azimuth is the angle around the vertical axis, set at 0 degrees directly in front of the head and increasing in the counter-clockwise direction, and elevation is the angle above (positive values) or below (negative values) the horizontal plane.

III. DIRECTIONAL LOCALIZATION

A. Data analysis

Three measures were used to evaluate directional localization accuracy in each of the five stimulus conditions: the overall angular error, the mean signed error, and the bias-corrected root-mean-square error (BCRMS). The angular error is simply the angle of the great-circle arc from the stimulus location to the response location. The mean signed error measures the overall bias in the responses. The BCRMS error attempts to measure the precision of the subject's responses when the bias is direction dependent: it measures the root-mean-squared value of the difference between each sub-

ject's response and the second-order polynomial best fitting the subject's responses to the stimulus locations (Brungart *et al.*, 1999).

An additional measure was used to characterize localization performance in azimuth: front-back confusions. Usually front-back reversals are assumed to occur whenever the stimulus location and the response location are on different sides of the frontal plane. In order to reduce the number of reversals in the vicinity of -90 degrees in azimuth, a relatively conservative definition of front-back confusions was used in this experiment. Trials were considered reversals only if the response location was at least 10 degrees closer to the mirror image of the source across the frontal plane than to the actual source location. The number of front-back reversals was calculated in each condition, and the trials where front-back confusions occurred were eliminated from the analyses of overall angular error and of the errors in azimuth.

B. Overall angular error

The overall angular error, which includes both the effects of systematic response biases and of response variability, clearly divides the five stimulus conditions into three groups (Table II). The best performing group included the broadband and fixed-amplitude conditions. Both exhibited a mean angular error around 17 degrees. The worst performing group included the monaural and low-pass filtered conditions. The error for these two conditions was 30 degrees, nearly double the error of the best performing group. The error in the high-pass filtered condition, approximately 23 degrees, was significantly larger than in the broadband and fixed-amplitude conditions, and significantly smaller than in the monaural and low-pass conditions ($p < 0.01$, one-tail t -tests).

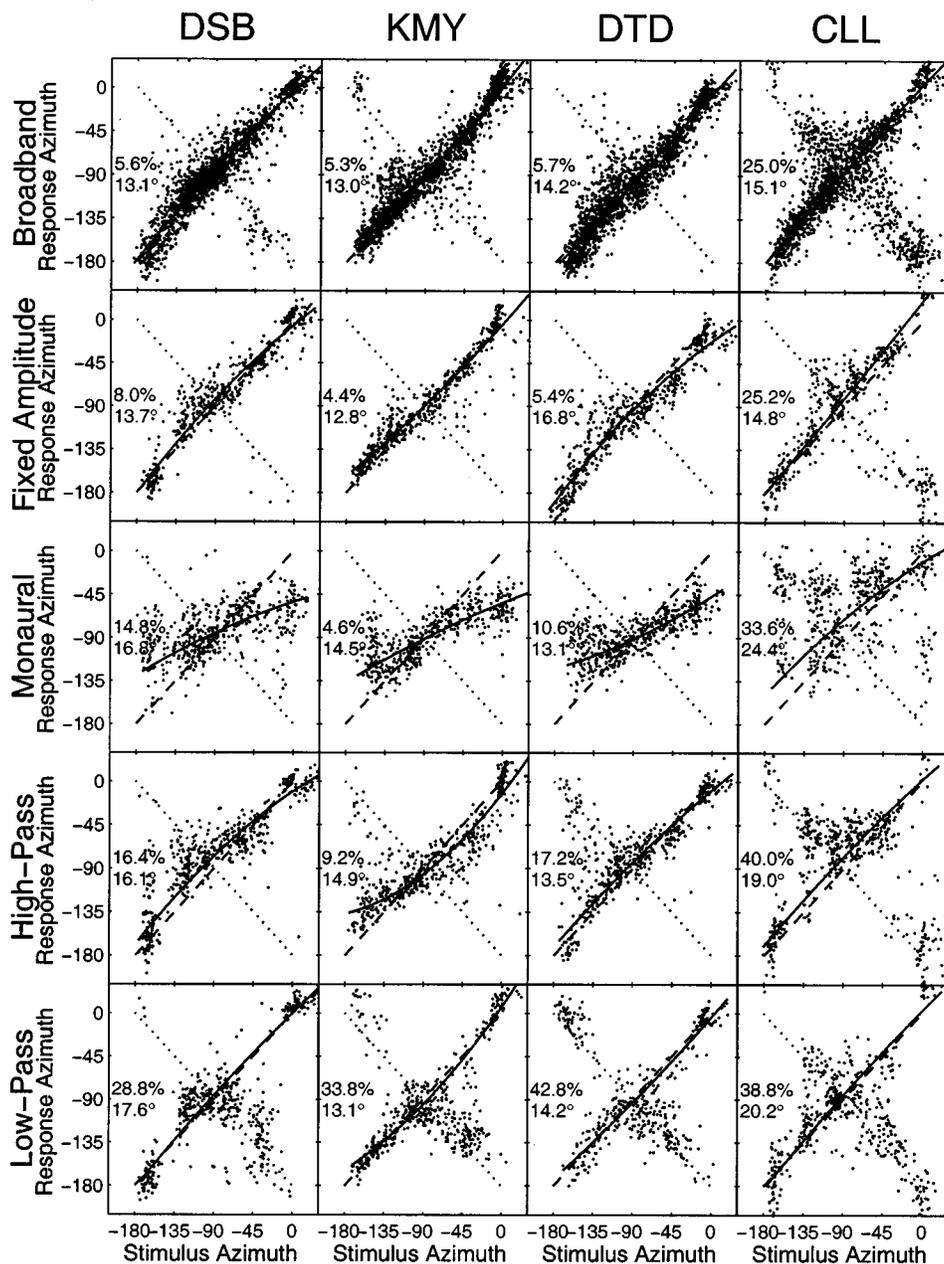


FIG. 1. Raw stimulus-response data in azimuth for each of the four subjects (columns) in each of the five conditions (rows). In each panel, the dashed line shows the location of “ideal” responses and the dotted line shows the location of perfect reversals. The solid black line is the second-order polynomial fit of the response location to the stimulus location. The percentage of reversals is the top number at the left of each panel, and the bias-corrected rms error (BCRMS) is the bottom number at the left of each panel (see text for details).

C. Azimuth localization performance

Overall azimuth localization performance was roughly comparable across all the conditions except the monaural condition. The signed azimuth and BCRMS errors were slightly higher in the high-pass and low-pass conditions than in the broadband conditions, but the differences across these four conditions were minor (Table II). The major difference in azimuth performance is in the monaural condition, as can be seen in the raw stimulus-response azimuth data (Fig. 1). From the raw data, it is evident that the responses in the monaural condition exhibited a strong lateral bias for three of the four subjects. The responses were clustered around -90 degrees and the slope of the best fitting stimulus-response curve is much less than 1. The responses for the fourth subject were almost random in the monaural condition.

Another important measure of azimuth localization is the percentage of front-back confusions. The overall number

of reversals was similar in the broadband and fixed-amplitude conditions ($\approx 10\%$), and increased sequentially in the monaural (16%), high-pass (21%), and low-pass (36%) conditions (ordering is significant at $p < 0.005$ from one-tailed t -tests). From Fig. 1, note that subject CLL exhibited substantially more reversals than any of the other subjects in every condition except the low-pass condition, where his reversals were probably constrained by a ceiling effect at pure chance (50%).

Of particular interest in this proximal-region localization study is the relationship between directional localization accuracy and source distance. Overall, azimuth localization was slightly but significantly better when the source was relatively distant (>50 cm) than when it was close (<25 cm) (Fig. 2). This relationship holds for the BCRMS error and the reversal percentage in all of the binaural conditions and for the BCRMS error in the monaural condition. While lo-

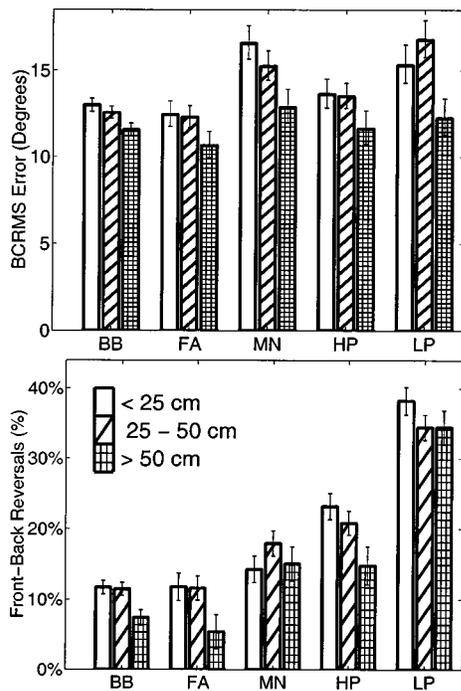


FIG. 2. Azimuth performance as a function of source distance. The top panel shows the bias-corrected rms error in azimuth at each of three distances. The signed bias-corrected errors were calculated separately for each subject at each of three elevation bins (<20 degrees, -20 to 20 degrees, and >20 degrees). The BCRMS errors in the top panel are the rms values of all the signed bias-corrected errors in each distance bin. The bottom panel shows the mean number of reversals in each condition at each distance. The error bars mark the 95% confidence intervals: BB=broadband; FA=fixed amplitude; MN=monaural; HP=high-pass; LP=low-pass. Note: The 95% confidence intervals in the BCRMS are derived from the χ^2 distribution with the assumption that the residual errors (signed azimuth error minus the mean signed azimuth error or signed azimuth error minus the second-order polynomial fit to azimuth error) are normally distributed.

calization in azimuth does depend on distance and is generally less accurate at close distances, the decrease in performance for nearby sources is relatively modest.

D. Elevation localization performance

In terms of the BCRMS error, elevation performance was best in the fixed amplitude condition, intermediate in the broadband condition, and worst in the monaural, high-pass, and low-pass conditions. (Table II). However, from the raw data (Fig. 3) it is clear that overall performance was substantially worse in the low-pass condition than in any other condition due to a strong bias toward the horizontal plane in all responses. Also, it is evident that overall there was a tendency for the subjects to underestimate high elevations.

Localization accuracy in elevation did not change consistently with source distance (Fig. 4). The ordering of the response errors with distance varied with condition.

E. Directional response bias

Directional biases are displayed graphically in Fig. 5, and are also represented by the mean signed azimuth and elevation errors in Table II. In the broadband conditions (broadband and fixed-amplitude), the response biases are relatively modest, and the general patterns of bias are similar

across each of the three distance bins. In the high-pass condition, the response biases tend to be larger than in the broadband conditions, but they continue to be roughly constant across distance. The most interesting response biases are in the monaural and low-pass conditions. In the monaural condition, the responses are strongly biased in the direction of the unoccluded ear. Such a pattern of bias has been reported in previous studies of monaural localization. In the low-pass condition, virtually all of the responses are clustered around the horizontal plane, independent of the actual source location. This pattern of response bias provides further evidence that the subjects were receiving little or no information about the elevation of the source when the stimulus was limited to frequencies below 3 kHz.

F. Discussion

The results of the broadband condition were discussed in great detail and compared to previous results from distal-region localization experiments in a previous paper (Brungart and Durlach, 1999), and those analyses will not be repeated here. The focus of this experiment was the relative accuracy of proximal-region localization accuracy in the five stimulus conditions. The results in each of the four new stimulus conditions are discussed below.

1. Fixed amplitude condition

Directional localization performance in the fixed-amplitude condition was virtually identical to the random-amplitude (broadband) condition in every way. The angular error of 16.6 degrees is nearly the same as the 17-degree error in the broadband condition, and the other error measures are also similar. The direction and magnitude of the biases were also similar in the two conditions.

The similarity between the broadband and fixed-amplitude conditions is not surprising. The only difference between the two conditions was the amplitude of the source, and under normal conditions directional localization judgments are based on interaural differences and spectral cues, and should be effectively independent of amplitude.³ The similarity between the two conditions, therefore, should be viewed simply as a measure of the repeatability of the experimental procedure.

2. Monaural condition

Performance in the monaural condition was drastically worse than in the broadband condition. The decrease in performance was most pronounced in azimuth, where response variability was significantly larger than in the broadband condition. The response biases, which were characterized by a strong pull in the direction of the unoccluded ear, were also dramatically different from the other conditions. This type of response behavior has been documented in previous studies of monaural localization (Wightman and Kistler, 1997). Perhaps more interesting is the relatively modest degradation in elevation accuracy in the monaural condition. Elevation performance was comparable to the binaural high-pass conditions and better than the binaural low-pass condition. The increase in front-back confusions (from 10% to 16%) was also relatively small, but note that the bias toward -90 de-

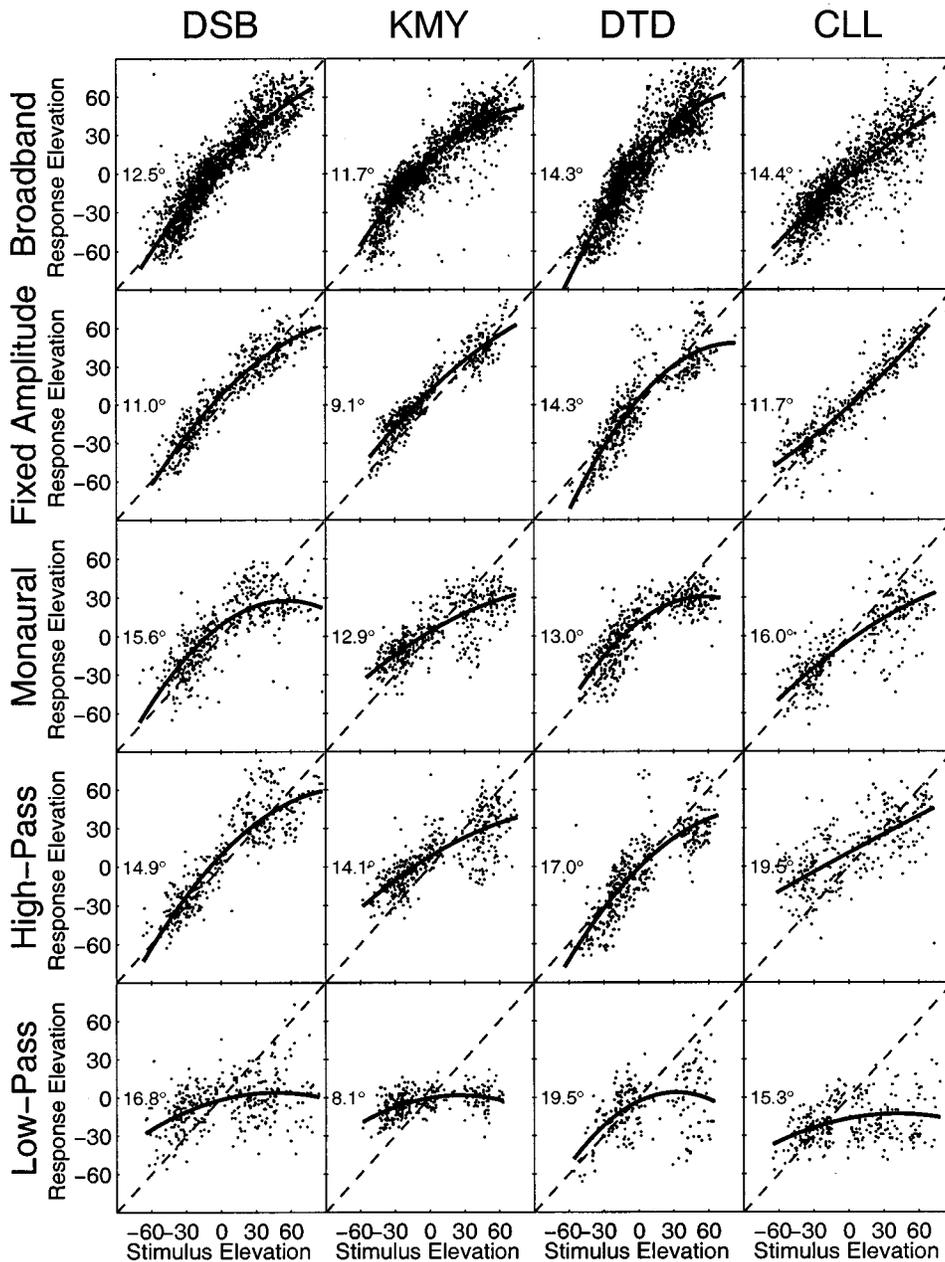


FIG. 3. Raw stimulus-response data in elevation for experiment 1. As in Fig. 1, the dotted line indicates “correct” responses, while the solid black line shows the best second-order polynomial fit of the response data to the stimulus data. The number at the left of each panel is the mean bias-corrected error.

degrees in azimuth tended to move reversed responses into the region where reversals were not 10 degrees closer to the mirror image of the source location than to the source location and were not counted as reversals. The elevation and reversal results are consistent with the widely held belief that the directional properties of the pinna at high frequencies are largely responsible for resolving the locations of sources within the “cone of confusion” where interaural differences are ambiguous. The available cues from one ear are apparently sufficient to allow subjects to make reasonably accurate judgments about elevation and correctly resolve front-back confusions, even though they are unable to accurately judge azimuth without binaural difference cues.

3. High-pass condition

In the high-pass condition, localization accuracy was slightly worse than in the broadband condition both in azimuth and in elevation. These results were consistent with

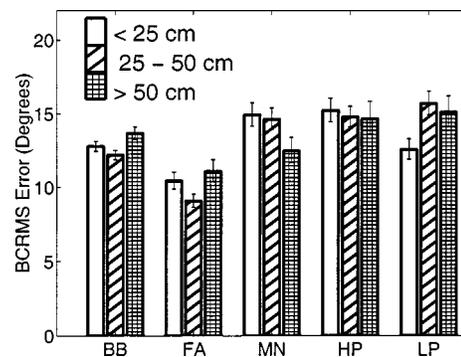


FIG. 4. Elevation performance as a function of source distance. The bias-corrected signed elevation errors were calculated separately for each subject in each of three azimuth bins (< -120 degrees, -120 to -60 degrees, and > -60 degrees) at each of the three distances, and combined into a single BCRMS value as in Fig. 2. The error bars mark the 95% confidence intervals: BB=broadband; FA=fixed amplitude; MN=monaural; HP=high-pass; LP=low-pass.

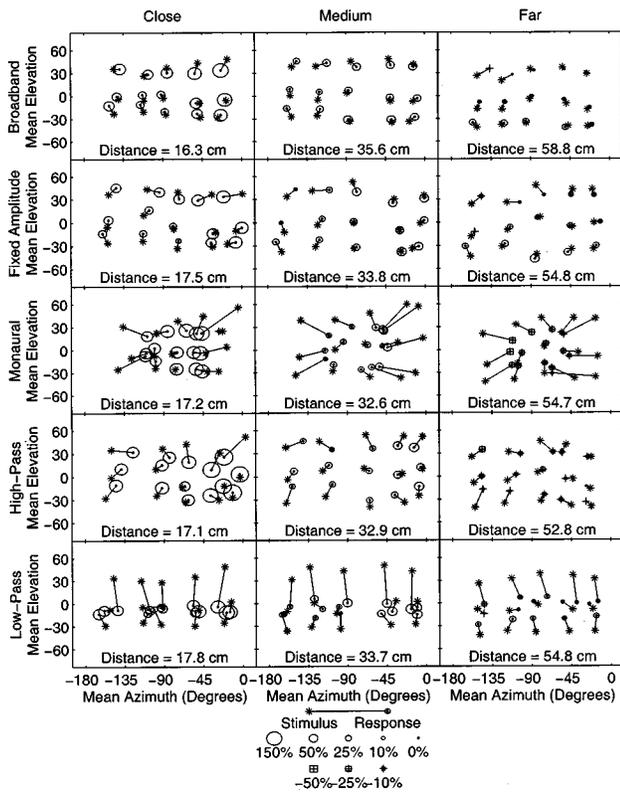


FIG. 5. Direction of response bias as a function of source location. In this figure, the trials have been sorted into three equal-sized non-overlapping bins in elevation, five equal-sized non-overlapping bins in azimuth, and three equal-sized bins in distance. The asterisk is the mean stimulus location within each bin, while the dot is the mean response location. The circles indicate that the subject overestimated distance (see legend), while the squares surrounding crosses indicate the subject underestimated distance.

those of Musicant and Butler (1984a), who measured localization performance for noise high-pass filtered at 4 kHz. Musicant and Butler did, however, report a much smaller number of front-back confusions (approximately 0.5% versus 21% in this experiment), possibly because their 5-s stimulus may have allowed some small head motions. The results of the high-pass localization condition are particularly interesting in the proximal region because the interaural intensity difference, which, according to the duplex theory of localization, dominates azimuth localization at high frequencies, varies substantially with distance across all frequencies. Since the ILD does not correspond to a unique lateral position in the proximal region, one might expect listeners to confuse the distance and direction of a proximal-region source containing only high frequencies. In particular, it seems likely that subjects would be unable to distinguish between a close source near the median plane and a more distant source at a more lateral position with the same average ILD. At the very least, one would expect this ambiguity to result in a very strong lateral bias for sources near the head. The results do not, however, conform to these expectations. Although there is a relatively strong lateral bias in the high-pass condition (Fig. 5), the bias is only slightly larger at close distances than at the farthest measured distances. Furthermore, an examination of the raw azimuth data (Fig. 1) reveals that only one subject (KMY) exhibited a

strong tendency to compress his responses around -90 degrees in the high-pass condition (note that almost all of his responses fall within the lateral quadrants defined by the dotted and dashed lines).

The most probable explanation for the failure of the subjects to exhibit significant lateral bias in the high-pass experiment is that the subjects were able to obtain salient interaural timing information from either the onset or the envelope of the stimulus. The 50-ms ramps used to gate the high-pass stimulus may not have been sufficient to eliminate all onset cues. Furthermore, researchers have demonstrated that subjects are sensitive to envelope delays in bandpass noise (Trahtois and Bernstein, 1986). The availability of accurate interaural timing information from either of these sources would explain the reasonably accurate localization judgments in the high-pass condition and the absence of a strong lateral bias for nearby sources.

4. Low-pass condition

In the low-pass condition, subjects were able to make reasonably accurate judgments about the lateral position of the source, but they were almost completely unable to estimate source elevation or resolve front-back confusions. The subjects tended to respond near the horizontal plane, independent of the true source elevation. Three of the subjects experienced more than twice as many front-back reversals in the low-pass condition as in any other condition. Again, these results illustrate the importance of the high-frequency directionality of the pinna in determining the elevation of a source and distinguishing between sources in the front and rear hemisphere. Since these high-frequency directional cues only occur above approximately 4 kHz, they were not available in the low-pass condition. Consequently, the subjects exhibited large numbers of reversals and very poor elevation accuracy.

The overall directional accuracy in each of the five stimulus conditions, and in particular the relative performance in each of the five conditions in azimuth, elevation, and front-back confusions, was consistent with the results of previous localization studies conducted in the distal region. This consistency both validates the procedure used in the experiment and tests the conjecture, based on measured head-related transfer functions in the proximal region, that directional localization performance is similar in the proximal and distal regions. This experiment also provided the opportunity to evaluate exactly how directional accuracy changes with distance in the proximal region. The results were mixed, but in general performance was slightly worse both in azimuth and in elevation for very near sources (<25 cm) than for more distant sources (>50 cm). In all cases, however, the distance dependence was relatively weak. Again, this confirms the observation that directional localization is relatively similar in the proximal and distal regions.

IV. DISTANCE LOCALIZATION

A. Results

While there does not appear to be a large difference in directional localization performance between sources in the

proximal and distal regions, the accuracy of auditory distance perception is dramatically different in the proximal region. As reported in the second paper in this series (Brungart and Durlach, 1999), proximal-region distance perception with a broadband, random-amplitude source is significantly more accurate for lateral sources than for sources near the median plane. Furthermore, distance judgments for lateral sources are highly correlated with source location (log-stimulus log-response correlation coefficients were in excess of 0.85) and are relatively unbiased. Although it is difficult to make direct comparisons between these results and those of previous far-field distance localization experiments, it appears that proximal-region distance perception for lateral sources is significantly better than has been reported in far-field experiments in the absence of room reverberation and amplitude cues.

In this experiment, overall distance performance was best in the fixed-amplitude condition, and only slightly less accurate in the broadband and low-pass conditions (Fig. 6). Performance in the high-pass condition was substantially worse than in the low-pass condition, and performance in the monaural condition was exceptionally poor. As has been reported in previous localization experiments, the subjects (with the exception of DSB) tended to compress their responses over a narrower range than the range of stimulus locations (generating a linear fit with a slope less than one). In most cases, this caused the subjects to significantly overestimate close distances.

As in the earlier proximal-region experiment, distance performance was highly dependent on source azimuth (Figs. 7 and 8). In these figures, the data were sorted by azimuth into 13 overlapping bins, and the correlation coefficients between the log-stimulus and log-response distances are shown as a function of the mean azimuth in each bin. In each of the five conditions, the correlation coefficient was lowest near 0 degrees azimuth and increased systematically as the source moved toward -90 degrees. Performance in the fixed-amplitude condition, where all subjects consistently performed best, showed the least degradation at locations near the median plane. Consequently, fixed-amplitude distance performance was only slightly better than broadband or low-pass distance performance near -90 degrees, but was considerably better near 0 and -180 degrees. Performance in the broadband and low-pass conditions was virtually identical everywhere except behind the listener, where the correlation coefficients were slightly higher in the broadband condition. Performance in the high-pass condition was substantially worse than in the broadband, fixed-amplitude, or low-pass condition. Performance in the monaural condition was extremely poor, with mean correlations never exceeding 0.4. Note that the increase in performance for lateral source locations and the relative ordering of the five stimulus conditions were essentially identical for each of the four subjects.

B. Discussion

The relative distance performance in these five stimulus conditions provides some clues about the mechanisms involved in proximal-region distance perception. The most

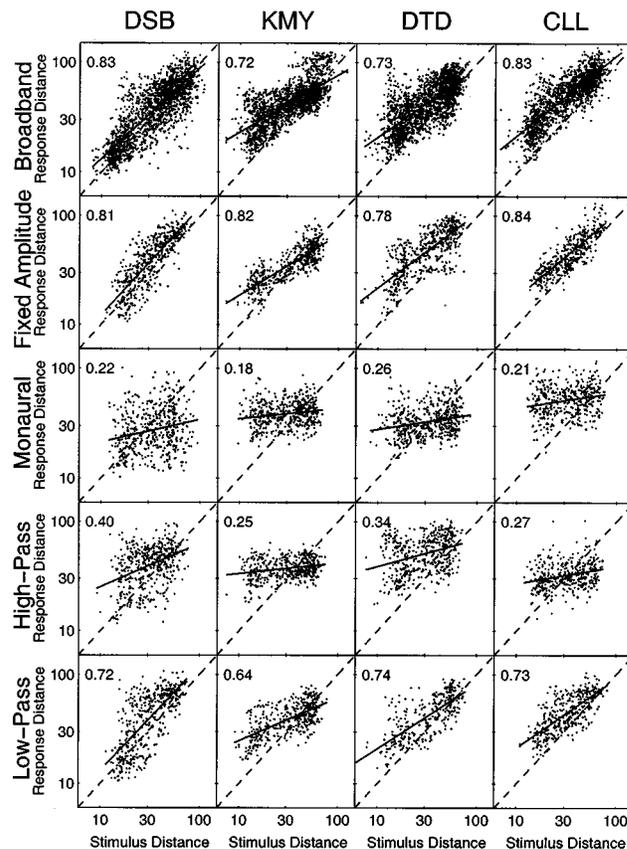


FIG. 6. These stimulus-response pairs for distance (in centimeters) include all azimuth and elevation locations. The errors in distance increased proportionately with stimulus distance, so the data are shown on a log-log scale. The solid line in each panel of the figure represents the best linear fit of the stimulus location to the response location, and the number at the top left of each panel is the correlation coefficient between the log-stimulus and log-response distances.

compelling indication is the importance of binaural cues to the localization of nearby sources. Several factors contribute to this conclusion:

- (i) In all four binaural conditions, distance judgments were considerably less accurate for medial sources than for lateral sources. Since binaural cues are largest near the interaural axis and smallest near the median plane, this is consistent with a binaural distance cue.
- (ii) In the monaural condition, distance judgments are exceptionally poor. Without binaural information, the subjects were unable to accurately judge the distance of the source.
- (iii) The availability of amplitude cues in the fixed-amplitude condition provided the most benefit when the source was near the median plane. While performance in this condition was better than in the random-amplitude condition at all locations, the increase in performance was greatest in the median plane (the correlation coefficient is more than twice as large in the fixed-amplitude condition than in the broadband condition at 0 degrees). This indicates that the binaural distance cues are nearly as potent as the amplitude cue for lateral sources, but that the listeners rely primarily on the amplitude cue in the median plane where the binaural cues are weak.

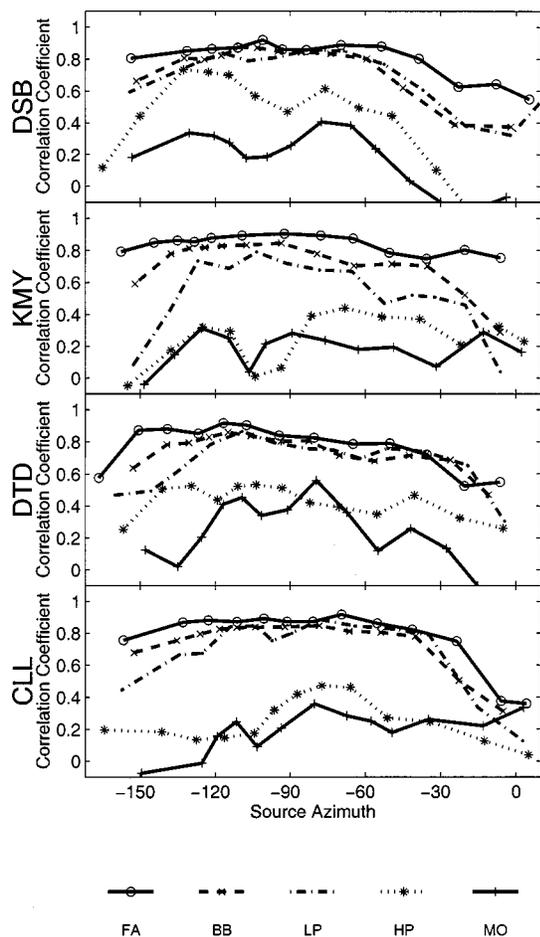


FIG. 7. Stimulus–response correlation in distance as a function of source azimuth for each subject. The data for each subject in each condition were sorted by azimuth into 13 bins, each containing 14% of the total trials and overlapping with 50% of the trials of each adjacent bin. The correlation coefficient between the log of the stimulus location and the log of the response location in each bin is plotted as a function of the mean azimuth location of each bin.

In addition, there is substantial evidence indicating that low-frequency interaural intensity cues dominate proximal-region distance perception:

- (i) The results in Fig. 8 show that distance localization performance is comparable for a broadband source and a source low-pass filtered below 3 kHz. In contrast, distance performance degrades substantially for a source high-pass filtered above 3 kHz. These results indicate that low-frequency components, below 3 kHz, are essential to proximal-region distance perception.
- (ii) Measurements of the proximal-region HRTF on a KEMAR manikin (Brungart and Rabinowitz, 1999) indicate that, in terms of human perception, the interaural level difference at low frequencies increases substantially as a nearby source approaches the head, while the interaural time delay is roughly independent of distance.⁴

Since proximal-region distance perception appears to depend on binaural information below 3 kHz, and the ITD is roughly

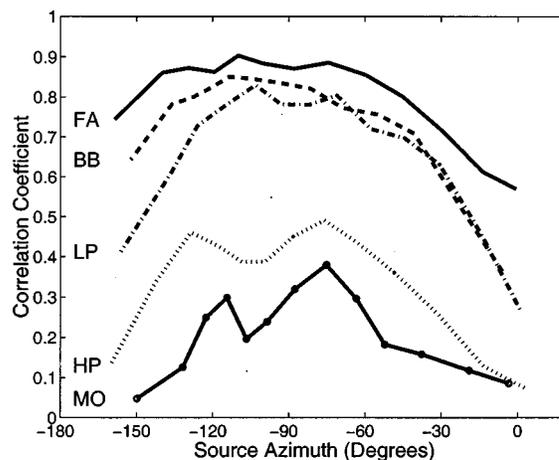


FIG. 8. Mean overall stimulus–response correlation in distance. The data for each subject in Fig. 7 have been averaged together by plotting the mean correlation coefficient (averaged using the Fisher transform) in each of the 13 bins versus the mean azimuth value of each bin.

independent of distance, by elimination it appears that the low-frequency ILD must be the dominant factor in proximal-region distance perception. This explanation is particularly appealing because it is known that, for far-field sources at least, the low-frequency ITD dominates the perception of the azimuth of a sound source (Wightman and Kistler, 1992). If this result also applies to proximal-region localization (which is not guaranteed, because low-frequency ILDs are much larger in the proximal region than in the distal region), it would indicate that relatively orthogonal acoustic cues, ITD and ILD, are used to determine the direction and distance of a nearby sound source.

Two other comments should also be made about the importance of low-frequency ILDs to proximal-region distance perception. First, it should be noted that substantial low-frequency ILDs in general exist only in the proximal region. The very presence of a large, low-frequency ILD is indicative of a close source. Thus, a distance perception mechanism based on low-frequency ILD could be highly specialized for determining the distance of nearby sources. In contrast, ITDs, high-frequency ILDs, and spectral shape are all integrated into the general process of auditory localization, and all can contribute to the perception of azimuth, elevation, and distance. Second, a distance-perception mechanism based on low-frequency ILDs may help to explain the sensitivity of human listeners to small changes in relatively large low-frequency ILDs that would not occur for a distant source. Previous research (Hershkowitz and Durlach, 1969; Hafter *et al.*, 1977) has shown that listeners are sensitive to changes in ILD of less than 1 dB over a 30-dB range. At distances of 1 m or more, however, the ILD generated by a source in a free field would never exceed 5 dB at 500 Hz, and would be even smaller at lower frequencies. If only far-field listening conditions are considered, the ability to detect small changes in the low-frequency ILD over a wide range of ILDs does not appear to serve any useful purpose. But this ability would be vital if low-frequency ILDs are used to determine distance in the proximal region.

One surprising outcome of this experiment is that high-frequency ILDs appear to provide so little information about the distance of a sound source. As discussed in the earlier papers in this series, the most basic model for proximal-region auditory distance perception assumes that the ITD, which is roughly independent of distance, is used to determine the lateral position of the source, and that the ILD, which varies systematically as a function of distance for a fixed source direction, is used to determine the distance of the source. Since the ILD increases across all frequencies as the distance of a proximal-region source decreases, and actually increases more at high frequencies than at low frequencies, it is not clear why listeners cannot judge the distance of a source from the high-frequency ILD. The most obvious explanation, that the subjects were simply unable to accurately determine the direction of the high-pass filtered source, is not supported by the directional localization data. The results indicate that the directional localization was reasonably accurate with the high-pass filtered source, perhaps because the subjects were able to use high-frequency envelope delays to estimate the ITD. Yet distance localization with the high-pass filtered source was poor. It is puzzling that the subjects were able to determine the direction of the source, but were unable to use the distant-dependent changes in the high-frequency ILD to estimate its distance. Two factors may explain these results:

- (1) The proximal region ILD at high frequencies may be so large that the auditory system is no longer able to detect small changes in the ILD. At low frequencies, the ILD is relatively small when the source is in the distal region, and it does not increase above 20–30 dB as the source approaches the head. At 500 Hz, for example, the ILD increases from 4 to 20 dB as a source at 90 degrees moves from 1 to 12 cm (Brungart and Rabinowitz, 1999). At high frequencies, however, the ILD can be large in the distal region and becomes extremely large in the proximal region. At 6 kHz, the ILD increases from 30 to 45 dB as a source at 90 degrees moves from 1 to 12 cm. Most studies examining the ILD have reported that sensitivity to increases in the just-noticeable difference (jnd) does not change significantly as the initial ILD varies over a wide range (Hershkowitz and Durlach, 1969; Hafter *et al.*, 1977). Nevertheless, it is likely that if the ILD becomes sufficiently large, sensitivity to changes in the ILD will decrease; a sufficiently large ILD will cause the signal at the contralateral ear to fall below the threshold of hearing. Since the ILDs at high frequencies are large even in the distal region, they may become so large when the source approaches the head that the auditory system saturates and can no longer detect changes in the ILD. This saturation would explain the inability of listeners to determine the distance of a high-frequency proximal-region source.
- (2) A second possible explanation for poor high-frequency distance localization follows from the assumption of the duplex theory that interaural timing information is salient primarily at low frequencies. If one assumes that the mechanisms of the auditory system that determine

the ITD of a low-frequency signal [which are known to dominate the perception of direction when the stimulus contains low-frequency energy (Wightman and Kistler, 1992)] are fundamentally different from those that determine the ITD of a high-frequency signal (from the envelopes of the signal, for example), then it is possible that our ability to determine the distance of a source from its direction (determined from the ITD) and its ILD can only use low-frequency interaural timing information. In other words, it is possible that the distance perception mechanism receives input only from the low-frequency ITD detector, while the directional localization mechanism receives information from both the low- and high-frequency ITD.

The inability of listeners to determine the distance of high-pass filtered signals is one of the more interesting results of these proximal-region localization experiments. While it is clear that low-frequency ILDs play an important role in proximal-region localization, further investigation is required before the mechanisms of auditory distance perception are fully understood.

V. EXPERIMENT 2: DISTANCE LOCALIZATION ALONG THE INTERAURAL AXIS

The results of the first experiment provide strong evidence that proximal-region distance perception is based on the low-frequency interaural intensity difference. However, there were variations in the overall amplitude of the different stimuli in the first experiment, and it was not clear how those variations might have influenced overall localization performance. In particular, the high-pass stimulus was substantially quieter than the low-pass or broadband stimulus. Although the subjects were clearly able to localize the high-pass stimulus (based on the directional localization data), it is not clear whether the amplitudes of the different stimulus types influence distance localization.

In addition, performance in some of the conditions (particularly the monaural condition) was so poor that it is difficult to verify that the subjects were not using extraneous information to assist them in their localization responses (either from “peeking” or from sounds made by the experimenter in the placement of the source). The validity of the first experiment was tested by collecting 100 trials per subject where the stimulus was placed normally but no sound stimulus was generated. The results showed that the subjects were receiving almost no extraneous information about the direction of the source (overall angular errors exceeded 50 degrees) but that they were extracting some information about the source distance. The average log-distance correlation coefficient was 0.16, and it was as high as 0.33 for one subject. This small amount of distance information is largely irrelevant in the conditions where performance was relatively good (low-pass, broadband) and the mean overall correlation was in excess of 0.75, but it complicates the interpretation of the results in conditions where measured performance was poor (monaural, high-pass). Further investigation showed that the extraneous information was a combination of moving shadows which were visible through closed eyelids, a

small amount of amplifier noise from the source, and the perception of air currents on the face from the moving source.

In order to ensure that overall stimulus amplitude and extraneous distance information did not significantly influence the distance results in the first experiment, a second experiment was performed to test distance accuracy for sources at -90 degrees azimuth and 0 degrees elevation.

A. Methods

The procedure for experiment 2 was similar to the procedure in the first experiment described, with three exceptions:

- (1) Source locations were limited to the interaural axis (-90 degrees in azimuth, 0 degrees in elevation). During each trial, the computer produced only a single number, ranging from 1 to 6, which the experimenter used to determine the approximate distance of the source from the head. In order to ease the placement of the source lateral to the listener, the experimenter stood behind and to the right of the listener, rather than directly to the right as in the first experiment.
- (2) The same signals from the first experiment (broadband, low-pass at 3 kHz, high-pass at 3 kHz) were used in experiment 2, but amplitudes of the stimuli were adjusted to produce the same maximum sound pressure level (48 dB SPL at 1 m) in each condition. The fixed-amplitude condition was dropped, and two more high-pass conditions were added, one with a cutoff frequency of 1.5 kHz and one with a cutoff frequency of 0.75 kHz. The randomization of the stimulus amplitude in each trial was also modified slightly. The range over which the normalized stimulus amplitude was roved was increased 5 dB, from 15 to 20 dB, by limiting the maximum amplitude at distances greater than 0.5 m. Thus the maximum amplitude was slightly lower at 1 m than at 0.5 m in this condition (but the difference was small relative to the range of the amplitude rove).
- (3) Several steps were taken to eliminate extraneous distance information. A blindfold eliminated any possible visual cues.⁵ A fan blew a stream of air across the subject's face during the experiment in order to eliminate any air currents caused by the moving source, and a masking signal, generated by a loudspeaker located on the floor of the anechoic chamber 1 m in front of the subject, obscured any audible amplifier noise. The masking signal consisted of USASI noise, low-pass filtered at 10 kHz, and its intensity was 49 dBA SPL at the location of the subject's head. The masker signal was continuously generated throughout the experiment. Note that, although the masker was louder than the stimulus at the location of the head, the spatial separation of the signal and masker ensured that the stimulus was audible.

B. Results and discussion

The results of experiment 2 verify the overall results of the first experiment. Performance was best in the low-pass and broadband conditions and was degraded significantly in

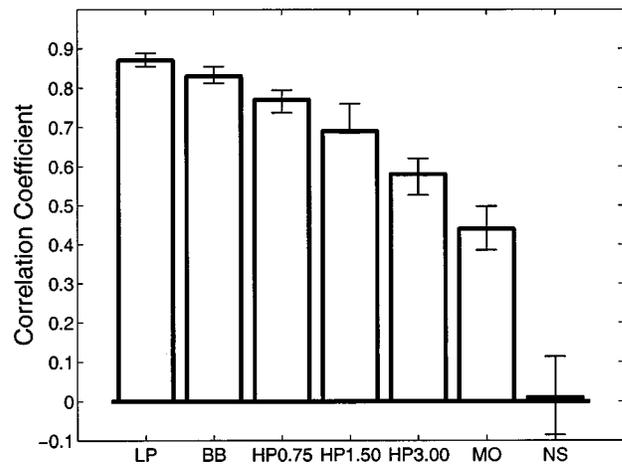


FIG. 9. Correlation coefficients between log-stimulus distance and log-response distance in experiment 2. The results have been averaged across the four subjects using the Fisher transform. The bars represent the 95% confidence intervals in each condition: LP=low-pass; BB=broadband; HP0.75=high-pass, 750-Hz cutoff; HP1.5=high-pass, 1.5-kHz cutoff; HP3.00=high-pass, 3-kHz cutoff; MO=monaural; NS=no sound.

the high-pass and monaural conditions (Fig. 9). Thus it does not appear that the lower overall amplitude of the stimulus in the high-pass condition was responsible for poor distance performance. It is interesting, however, that the stimulus-response correlation is higher in the low-pass condition than in the broadband condition. Since the stimuli all had equal overall amplitude, the low-pass stimuli contained more low-frequency energy than the broadband stimuli, and thus were not masked as effectively as the broadband condition. It is further testimony to the importance of low-frequency energy to proximal-region distance perception that this increase in low-frequency energy allowed subjects to perform slightly but significantly better in the low-pass condition than in the bandpass condition in the presence of masking noise.

The results of experiment 2 also confirm the existence of monaural proximal-region distance cues. The mean stimulus-response distance correlation was similar in the monaural conditions of experiments 1 and 2 (approximately 0.45 at -90 degrees), but in the first experiment it was not clear how much extraneous information contributed to this correlation (since subject KMY was able to achieve a distance correlation of 0.33 with no stimulus). In experiment 2, where the blindfold, fan, and masking noise rendered the subjects unable to make accurate judgments about source location in the no-stimulus condition, they clearly were still able to obtain some distance information in the monaural condition. A surprising result from the raw data (Fig. 10) is that the monaural distance correlation does not seem to be purely a result of correctly identifying very near sound sources. Subject DSB's responses were substantially more accurate when the source was within 15 cm of the head, but the other subjects' responses were not. All four subjects' mean responses increased slightly with source distance up to 1 m. This is surprising, because previous analysis of proximal-region head-related transfer functions indicated that spectral localization cues, primarily in the form of a slight emphasis on low frequencies for nearby sources, were most pronounced in the region within 25 cm of the head

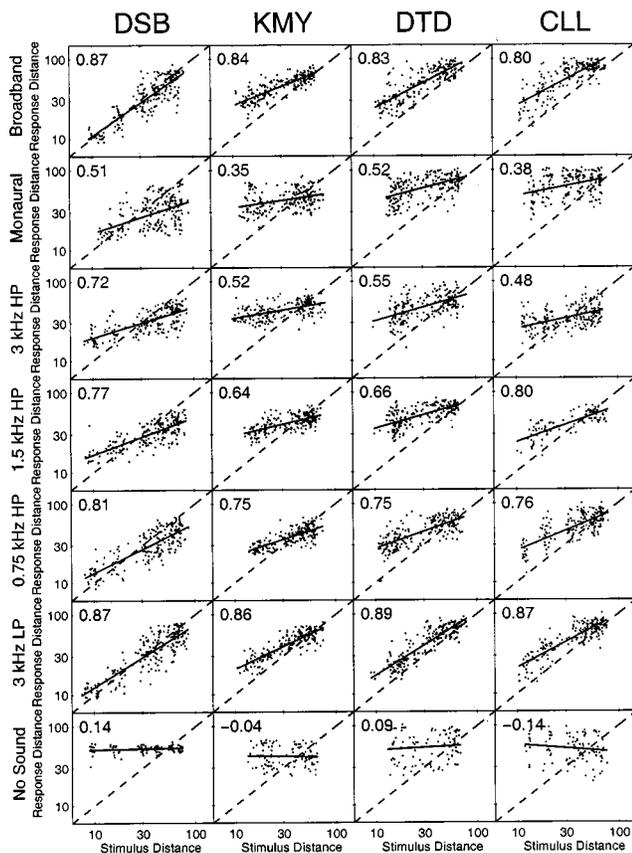


FIG. 10. Raw distance stimulus–response data for experiment 2. The data are for sources positioned along the interaural axis. The dashed lines represent “correct” responses, while the solid line is the best linear fit of the stimulus data to the response data. The number at the top left of each panel is the linear correlation coefficient.

(Brungart and Rabinowitz, 1999). Further research, including the measurement of monaural distance jnd’s and the evaluation of monaural distance performance with a variety of stimulus spectra, is necessary to fully understand the mechanisms that provide some distance information under monaural listening conditions. It is evident, however, that monaural distance cues are weaker than binaural cues and that binaural distance cues dominate when sources are located outside the median plane.

A final important result of experiment 2 is the relative performance in the three high-pass conditions, with cutoff frequencies of 3.0, 1.5, and 0.75 kHz. Although overall amplitude was the same in each of the three conditions, the distance correlation coefficient increased systematically as the cutoff frequency decreased for each of the four subjects, and the performance was better in the broadband condition than in the 0.75-kHz high-pass condition. This result provides further evidence that low-frequency energy is required for accurate proximal-region distance perception. It also indicates that some useful distance information occurs in the frequency range from 1.5 to 3 kHz, and that a significant portion of the useful distance information in a broadband stimulus occurs below 750 Hz. Thus it appears that even very low-frequency interaural level differences are important to auditory depth perception.

VI. CONCLUSIONS

The results from these studies can be summarized as follows:

- (i) Directional localization performance is similar in the proximal and distal regions. Although there was a slight decrease in the accuracy of directional localization for nearby sources, it appears that directional localization accuracy does not depend strongly on source distance, even when the source is near the head.
- (ii) Directional accuracy in each of the five stimulus conditions was consistent with the auditory localization literature. Performance in the fixed-amplitude and broadband conditions was almost identical. In the monaural condition, subjects were able to determine whether sources were in the front or rear hemisphere and make reasonable guesses about the elevation of the source, but were unable to accurately determine source azimuth. In the high-pass condition, localization performance was modestly degraded both in azimuth and elevation. In the low-pass condition, accuracy was only slightly degraded in azimuth, but the number of front–back confusions increased dramatically and elevation localization was severely degraded. These results conform to the widely held theory that interaural time and intensity differences allow subjects to determine the lateral position of a sound source, while high-frequency monaural pinna cues allow listeners to determine source elevation and to resolve front–back confusions.
- (iii) Some interaural timing information was apparently available to the subjects in the high-pass condition. Although the interaural level difference for a fixed source direction increases dramatically in the proximal region, the subjects did not appear to confuse source distance and source direction in the high-pass condition. They were not able to judge distance accurately, but they were able to judge direction reasonably well. Furthermore, they did not show a systematically larger lateral bias in direction at close distances, as would be expected if the subjects were relying solely on interaural level differences to determine source direction. This result suggests that the subjects were able to obtain interaural timing information, either from envelope delays or onset cues, that allowed them to accurately judge source direction.
- (iv) Distance perception is reasonably accurate for proximal-region sources outside the median plane, and the distance perception in this region appears to depend primarily on low-frequency binaural cues. In the broadband condition, the correlation between the log of the stimulus distance and the log of the response distance was as high as 0.85. This indicates distance perception superior to that found in far-field experiments under anechoic conditions with a random-amplitude source. Distance performance was poor in the monaural condition, and was worse near the median plane than near the interaural axis. This

pattern of performance indicates the importance of binaural cues to distance perception. Distance perception was also significantly worse in the high-pass condition than in the low-pass condition, which suggests the importance of low frequencies to distance perception. Measurements of the HRTF in the proximal region have indicated that interaural time delay is relatively independent of distance in the proximal region, so by elimination it appears that low-frequency interaural level differences are responsible for distance perception in the proximal region. The results of experiment 2 indicate that distance performance is enhanced by the addition of frequency content from 750 Hz to 3 kHz, but that additional distance information is provided by lower frequencies (below 750 Hz).

- (v) Amplitude cues are more important to auditory distance perception at medial locations than at lateral locations. In the median plane, binaural cues are relatively weak and amplitude cues dominate distance perception. At lateral locations, however, the binaural cues provide most of the distance information. Note that distance perception is better near the interaural axis when the amplitude of the source is randomized than in the median plane when the amplitude of the source is fixed. This indicates that proximal-region binaural distance cues are more salient than the amplitude-based distance cue which has been found to dominate distance judgments in the distal region.

ACKNOWLEDGMENTS

The author would like to thank Nat Durlach, Bill Rabinowitz, Steve Colburn, and William Peake for their assistance in the preparation of this manuscript, as well as the two anonymous reviewers who provided exceptionally detailed and helpful comments for improving this paper.

¹Due to the peculiarities of its design, the acoustic point source is capable of producing low-frequency sound more efficiently than high-frequency sound. Therefore, in order to maximize the output of the source, a "pink" stimulus was chosen.

²Due to a technical problem, only 400 trials were collected in the low-pass condition with subject DTD.

³Note that under unusual listening conditions, e.g., a unilateral hearing impairment, directional judgments may be based, in part, on the overall amplitude of the source.

⁴While the ITD can increase by as much as 100 μ s as a sound source moves from the distal region to within a few centimeters of the head, most of this increase occurs at lateral regions where listeners are relatively insensitive to changes in ITD (Brungart and Rabinowitz, 1999). From previously measured estimates of the just-noticeable differences (jnd's) in ITD and ILD for a 500-Hz tone (Hershkowitz and Durlach, 1969), we can calculate that the ILD increases by 15–20 or more jnd's as a source moves from a meter away to within a few centimeters of the head, while the ITD increases by, at most, a few jnd's. Thus we argue that the ITD is perceptually independent of distance.

⁵Note that the blindfold remained in place throughout the experiment and was not removed during the subject's responses.

- Batteau, D. (1967). "The role of the pinna in human localization," *Proc. R. Soc. London* **168**, 158–180.
- Brungart, D., Durlach, N., and Rabinowitz, W. (1999). "Auditory localization of nearby sources. II. Localization of a broadband source," *J. Acoust. Soc. Am.* **106**, 1956–1968.
- Brungart, D., and Rabinowitz, W. (1996). "Auditory localization in the near-field," in *Proceedings of the Third International Conference on Auditory Display*, Santa Fe Institute.
- Brungart, D., and Rabinowitz, W. (1999). "Auditory localization of nearby sources. Head-related transfer functions," *J. Acoust. Soc. Am.* **106**, 1465–1479.
- Brungart, D., Rabinowitz, W., and Durlach, N. (2000). "Evaluation of response methods for near-field auditory localization experiments," *Percept. Psychophys.* (in press).
- Butler, R. (1987). "An analysis of the monaural displacement of sound in space," *Percept. Psychophys.* **41**, 1–7.
- Butler, R., Levy, E., and Neff, W. (1980). "Apparent distance of sounds recorded in echoic and anechoic chambers," *J. Exp. Psychol.* **6**, 757–750.
- Coleman, P. (1963). "An analysis of cues to auditory depth perception in free space," *Psychol. Bull.* **60**, 302–315.
- Duda, R., and Martens, W. (1988). "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.* **104**, 3048–3058.
- Gardner, M. B., and Gardner, R. S. (1973). "Problem of localizing in the median plane: effect of pinnae cavity occlusion," *J. Acoust. Soc. Am.* **53**, 400–408.
- Haftner, E., Dye, R., Nuetzel, J., and Aronow, H. (1977). "Difference threshold for interaural intensity," *J. Acoust. Soc. Am.* **61**, 829–833.
- Hartley, R., and Fry, T. (1921). "The binaural location of pure tones," *Phys. Rev.* **18**, 431–442.
- Hebrank, J., and Wright, D. (1974). "Spectral cues used in the localization of sound sources on the median plane," *J. Acoust. Soc. Am.* **56**, 1820–1834.
- Henning, G. B. (1974). "Detectability of interaural delay in high-frequency waveforms," *J. Acoust. Soc. Am.* **55**, 84–90.
- Hershkowitz, R., and Durlach, N. (1969). "Interaural time and amplitude JNDs for a 500-Hz tone," *J. Acoust. Soc. Am.* **46**, 1464–1467.
- Little, A., Mershon, D., and Cox, P. (1992). "Spectral content as a cue to perceived auditory distance," *Perception* **21**, 405–416.
- McFadden, D., and Pasanen, E. (1976). "Lateralization at high frequencies based on interaural time differences," *J. Acoust. Soc. Am.* **59**, 634–1200.
- Musicant, A. D., and Butler, R. A. (1984a). "The psychophysical basis of monaural localization," *Hearing Res.* **14**, 185–190.
- Musicant, A., and Butler, R. A. (1984b). "The influence of pinnae-based spectral cues on sound localization," *J. Acoust. Soc. Am.* **75**, 1195–1200.
- Oldfield, S., and Parker, S. (1984). "Acuity of sound localization: a topography of auditory space. II: pinna cues absent," *Perception* **13**, 601–617.
- Oldfield, S., and Parker, S. (1986). "Acuity of sound localization: a topography of auditory space. III: Monaural hearing conditions," *Perception* **15**, 67–81.
- Rayleigh, L. (1907). "On our perception of sound direction," *Philos. Mag.* **13**, 214–232.
- Roffler, S., and Butler, R. (1968). "Factors that influence the localization of sound in the vertical plane," *J. Acoust. Soc. Am.* **43**, 1255–1259.
- Slattery, W., and Middlebrooks, J. (1994). "Monaural sound localization: acute versus chronic unilateral impairment," *Hearing Res.* **75**, 38–46.
- Trahoitis, C., and Bernstein, L. (1986). "Lateralization of bands of noise and sinusoidally amplitude-modulated tones: effects of spectral locus and bandwidth," *J. Acoust. Soc. Am.* **79**, 1950–1957.
- Wightman, F., and Kistler, D. (1992). "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Am.* **91**, 1648–1660.
- Wightman, F. L., and Kistler, D. J. (1997). "Monaural sound localization revisited," *J. Acoust. Soc. Am.* **101**, 1050–1063.
- Yost, W., Kistler, D. J., Wightman, F., and Green, D. (1971). "Lateralization of filtered clicks," *J. Acoust. Soc. Am.* **50**, 178–181.

Comparison of different forms of compression using wearable digital hearing aids

Michael A. Stone, Brian C. J. Moore,^{a)} José I. Alcántara, and Brian R. Glasberg
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

(Received 17 September 1998; revised 18 March 1999; accepted 2 August 1999)

Four different compression algorithms were implemented in wearable digital hearing aids: (1) The slow-acting dual-front-end automatic gain control (AGC) system [B. C. J. Moore, B. R. Glasberg, and M. A. Stone, *Br. J. Audiol.* **25**, 171–182 (1991)], combined with appropriate frequency response equalization, with a compression threshold of 63 dB sound pressure level (SPL) and with a compression ratio of 30 (DUAL-HI); (2) The dual-front-end AGC system combined with appropriate frequency response equalization, with a compression threshold of 55 dB SPL and with a compression ratio of 3 (DUAL-LO). This was intended to give some impression of the levels of sounds in the environment; (3) Fast-acting full dynamic range compression in four channels (FULL-4). The compression was designed to minimize envelope distortion due to overshoots and undershoots; (4) A combination of (2) and (3) above, where each applied less compression than when used alone (DUAL-4). Initial fitting was partly based on the concept of giving a flat specific-loudness pattern for a 65-dB SPL speech-shaped noise input, and this was followed by fine tuning using an adaptive procedure with speech stimuli. Eight subjects with moderate to severe cochlear hearing loss were tested in a counter-balanced design. Subjects had at least 2 weeks experience with each system in everyday life before evaluation using the Abbreviated Profile of Hearing Aid Benefit (APHAB) test and measures of speech intelligibility in quiet (AB word lists at 50 and 80 dB SPL) and noise (adoptive sentence lists in speech-shaped noise, or that same noise amplitude modulated with the envelope of speech from a single talker). The APHAB scores did not indicate clear differences between the four systems. Scores for the AB words in quiet were high for all four systems at both 50 and 80 dB SPL. The speech-to-noise ratios required for 50% intelligibility were low (indicating good performance) and similar for all the systems, but there was a slight trend for better performance in modulated noise with the DUAL-4 system than with the other systems. A subsequent trial where three subjects directly compared each of the four systems in their everyday lives indicated a slight preference for the DUAL-LO system. Overall, the results suggest that it is not necessary to compress fast modulations of the input signal. © 1999 Acoustical Society of America. [S0001-4966(99)00612-8]

PACS numbers: 43.66.Ts, 43.66.Cb, 43.71.Ky [JWH]

INTRODUCTION

People with cochlear hearing loss usually experience loudness recruitment, and the associated reduced dynamic range (Hood, 1972; Moore, 1998). With linear amplification, as provided by “conventional” hearing aids, it is not possible to restore audibility of weak sounds without intense sounds being overamplified and becoming uncomfortably loud. In principle, this problem can be alleviated by the use of automatic gain control (AGC), as proposed many years ago by Steinberg and Gardner (1937). Indeed, it is now widely accepted that AGC in hearing aids can be beneficial for people with loudness recruitment; for reviews, see Hickson (1994), Dillon (1996), and Moore (1998). Although the idea of AGC appears straightforward, in practice there are many ways of implementing it, and there is no clear consensus as to the “best” method. All AGC systems introduce some undesirable side effects, for example, distortions of the temporal envelope of sounds, and “pumping” or “breathing” sounds, and the benefits of any specific system must be weighed against these undesirable side effects.

Some AGC systems are intended to adjust the gain automatically for different listening situations. Essentially, they are intended to relieve the user of the need to adjust the volume control to deal with these situations. Usually, such systems change their gain slowly with changes in sound level; this is achieved by making the recovery time (t_r) or both t_r and the attack time (t_a), relatively long (usually t_r is between 500 ms and 20 s). These systems are often referred to as “automatic volume control” (AVC). Relatively few commercial hearing aids incorporate AVC, probably because conventional AVC systems involve a compromise between conflicting requirements. For speech with a fixed average level, the level may fluctuate markedly from moment to moment, and may drop to very low values during brief pauses in the speech. If the gain of an aid changes significantly during the speech itself, or during the pauses in the speech, then breathing or pumping noises may be heard which are objectionable to the user. These effects are particularly marked for users with mild to moderate losses when moderate levels of background noise are present, as is often the case in everyday situations. In addition to these objectionable effects, the temporal envelope of the speech may be distorted by the

^{a)}Electronic mail: bcjm@cus.cam.ac.uk

changes in gain. To avoid these problems, the gain should change relatively slowly as a function of time, i.e., t_a and t_r should be long.

On the other hand, it is important to protect the user from intense transient sounds, such as a door slamming or a cup being dropped. This requires the gain to change much more rapidly in response to an increase in sound level. This problem can be dealt with by having an AVC system with a fast attack time (t_a in the range 1–10 ms), and a moderate recovery time (t_r in the range 300 to 2000 ms). The fast attack time provides protection against sudden increases in sound level; the gain drops very quickly when the input sound level suddenly increases. However, such a system has the disadvantage that the gain of the aid drops to a low value immediately after an intense transient; the aid effectively goes “dead” for a while. A further problem is that a recovery time of a few hundred ms is not sufficiently long to prevent breathing and pumping sounds from being heard. In many commercial aids, the release time is smaller than 300 ms, which makes pumping even more audible (at least in single-channel devices). However, such aids should probably not be described as having AVC.

An AVC system developed in our laboratory provides a better solution to these problems (Moore and Glasberg, 1988; Moore *et al.*, 1991; Moore, 1993; Stone *et al.*, 1997). This system, referred to as “dual front-end AGC,” involves the use of two control systems, one with long attack and recovery times, and the other with shorter attack and recovery times. Normally, the operation of the AGC amplifier is determined by the slow-acting control system which has a high compression ratio. However, if an intense sound suddenly occurs, the fast-acting control system rapidly reduces the gain, thus protecting the user from overamplification of the intense sound. Following cessation of the intense sound, the gain returns to the original value determined by the overall level of the speech. In this way, intense transients are selectively attenuated, but the overall gain for the speech is held almost constant, except for a very brief period following the transient. A digital implementation of this system is described later in this paper. Control of overall gain while providing protection from intense transients can also be achieved using a combination of a slow-acting AVC with a low compression ratio and fast-acting compression limiting; a system similar to this is also described later in this paper.

An alternative type of AGC system, with lower compression ratios and lower compression thresholds, has been used in hearing aids in attempts to make the hearing-impaired person’s perception of loudness more like that of a normal listener (Villchur, 1973; Hohmann and Kollmeier, 1995; Kollmeier and Hohmann, 1995). Often, these systems split the signal into a number of frequency channels, and compression is applied independently in each channel. These are sometimes called “wide dynamic range” compressors, as the compression operates over a wide range of input sound levels. Typically, the compression threshold is about 45 dB sound pressure level (SPL), and the compression ratio has values up to about 3. One aim of such systems is to ensure that the weaker consonant sounds of speech, such as /p/, /t/, and /k/, are audible without the more intense sounds

(e.g., vowels) being uncomfortably loud. To achieve this goal, the compressors usually have short time constants. The attack times are typically 2–10 ms and the recovery times are typically 20–150 ms. Systems with time constants in this range are often referred to as “fast-acting compressors” or “syllabic compressors,” since the gain changes over times comparable to the durations of individual syllables in speech.

Evaluations of syllabic compressors have given mixed results, but benefits have been shown in some studies; speech in quiet can be understood over a wide range of sound levels (Lippmann *et al.*, 1981; Moore *et al.*, 1992a) and the intelligibility of speech in noise is sometimes slightly improved relative to that found with linear amplification (Moore *et al.*, 1992a; Yund and Buckles, 1995; Moore *et al.*, 1999b), although this has not always been found (Lippmann *et al.*, 1981; Bustamante and Braida, 1987).

The relative merits of the dual front-end AGC system and of multichannel syllabic compression have not previously been directly assessed. In principle, each system has advantages and disadvantages. For the dual front-end AGC system, the advantages are:

- (1) Speech is delivered at a comfortable level, regardless of the input level.
- (2) The temporal envelope of the speech is hardly distorted; envelope fluctuations at syllabic rates are preserved.
- (3) The spectral pattern of sounds is not distorted (although frequency-response shaping may be applied).
- (4) Harmonic and intermodulation distortion are minimal.
- (5) Protection is provided from intense brief transients with little effect on the long-term gain.

The disadvantages of the dual front-end AGC system are:

- (1) Loudness perception is not restored to “normal.” Indeed, since the output level is held almost constant for input levels above the compression threshold, it may be difficult for the user to judge the strength of sound sources, for example, the volume setting on a television or radio.
- (2) The system may not deal very effectively with situations where two voices alternate with markedly different levels, which is often the case when one of the voices is that of the aid wearer (although the problem in this specific case can be alleviated by the application of high-frequency emphasis prior to the AGC system, as described later).
- (3) When the user moves from a situation with high sound levels to one with lower levels (for example, when leaving a noisy room), the gain takes a second or two to reach the value appropriate for the new situation.

The advantages of fast-acting multichannel compression are:

- (1) It can, at least to a reasonable approximation, restore loudness perception to “normal,” although to achieve this might require extensive measurement of loudness growth functions.

- (2) If multiple channels are used, it can be set up to compensate for frequency-dependent changes in the degree of loudness recruitment.
- (3) It can restore the audibility of weak sounds rapidly following intense sounds.
- (4) It can give good results when two voices alternate with markedly different levels.

The disadvantages of fast-acting multichannel compression are:

- (1) It can introduce spurious changes in the temporal envelope of sounds (overshoot and undershoot effects).
- (2) It can introduce spurious changes in amplitude of sounds gliding in frequency, such as formants, as those sounds traverse the boundary between two channels.
- (3) It reduces intensity contrasts and the modulation depth of speech, which may have an adverse effect on the perception of certain speech cues (Plomp, 1988).
- (4) If many channels are used, then the availability of spectral cues, such as formant frequencies, may be reduced. In a multichannel hearing aid with fast-acting compression in many channels, the spectrum is flattened, reducing spectral contrasts.

It is not easy to decide upon the relative importance of these advantages and disadvantages, so as to choose the "best" system. In the present paper, we adopt a more empirical approach. We describe a direct comparison of four different AGC systems, including the dual front-end AGC system and a four-channel syllabic compressor, using wearable digital hearing aids. Each subject was tested using all four systems. For each system, the aid hardware was identical; only the software (the program implemented in the aid) was changed to implement the system.

The systems evaluated were as follows:

- (1) The slow-acting dual front-end AGC system, combined with appropriate frequency response equalization, with a compression threshold of 63 dB SPL and with a compression ratio of 30 for levels above this (DUAL-HI). This is similar to the system that was developed and evaluated in our previous work (Moore and Glasberg, 1988; Moore *et al.*, 1991; Stone *et al.*, 1997).
- (2) The dual front-end AGC system, combined with appropriate frequency response equalization, with a compression threshold of 55 dB SPL and with a compression ratio of 3 for levels above this. This was intended to give a more accurate impression of the levels of sounds in the environment (DUAL-LO).
- (3) Fast-acting full dynamic range compression in four channels. The compression was designed to minimize envelope distortion due to overshoots and undershoots (FULL-4).
- (4) A combination of (2) and (3) above, where each applied less compression than when used alone; the compression ratio for the dual front-end AGC was reduced to 1.7 and the compression threshold was set to 45 dB SPL (DUAL-4).

I. IMPLEMENTATION OF THE COMPRESSION SYSTEMS

The compression systems were implemented using Audallion digital hearing aids manufactured by Audiologic. Each aid consisted of two behind-the-ear earpieces, connected by wires to a small chest-worn processor, containing a Motorola DSP56009 DSP chip. The earmolds were fitted with Libby horns to improve the high-frequency response, and acoustic resistors were used in the tubing to damp mid-frequency resonances. The tubing was thick-walled to reduce feedback. The earmolds were hard acrylic skeletons and were deeply inserted to reduced feedback and to alleviate the occlusion effect. All had 1–2-mm vents, except those for subject VW, which had no vents. The signals picked up by the microphones in the earpieces were fed to the analog-to-digital converters (ADCs, 12-bit resolution, 15 625-Hz sampling rate) in the chest-worn part. In the versions used by us, the earpieces contained analog compression limiters with a compression threshold (at 500 Hz) of 100 dB SPL. These limiters prevented overloading of the ADCs; a full-scale input to the ADCs corresponded to an input sound level just above 100 dB SPL. Analog circuitry within the aids performed high-frequency emphasis prior to the ADC, to flatten the speech spectrum and give better coding of the higher-frequency portions of the input. After digital signal processing, the internal digital-to-analog converter (DAC, 16-bit resolution) was used to produce an analog voltage which was fed back to the amplifier and receiver in each earpiece.

Every 2 ms, the aid operating system delivered a time "frame" (a waveform segment, 4 ms in duration) to be processed. This frame contained 64 samples with a 50% overlap with the preceding frame. The data were windowed (using the positive half-cycle of a sine function) and subjected to a fast Fourier transform (FFT) to produce 32 complex values, spaced 244 Hz apart. The effects of the analog high-frequency emphasis used prior to the ADC, and of the non-flat frequency response of the microphone, were compensated for at this stage so as to give a flat frequency response between 125 and 7000 Hz. This also had the effect of reducing quantization noise introduced by the ADC.

In what follows, *bands* refers to frequency ranges used for frequency response shaping, while *channels* refers to frequency ranges used for compression. For all of the systems, eight bands were used for frequency response shaping. Also, the initial compression parameters for the two systems using four-channel compression (FULL-4 and DUAL-4) were based on calculations for an eight-channel system. Frequency "bins" resulting from the FFT occurred at intervals of 244 Hz. Bands were achieved by combining bins as follows (CF indicates center frequency):

Band 1, CF=183 Hz: dc term, +244,
 Band 2, CF=488 Hz: 488,
 Band 3, CF=849 Hz: 732+976,
 Band 4, CF=1342 Hz: 1220+1464,
 Band 5, CF=1952 Hz: 1708+1952+2196,
 Band 6, CF=2684 Hz: 2440+2684+2928,
 Band 7, CF=3660 Hz: 3172+3416+3660+3904+4148,
 Band 8, CF=5917 Hz: 4392 to 7564 (14 bins).

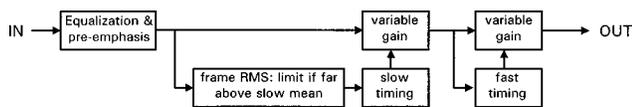


FIG. 1. A block diagram of the digital implementation of the dual front-end AGC system.

A. The dual front-end AGC systems (DUAL-HI and DUAL-LO)

As mentioned earlier, these systems used two gain control systems, one fast acting and one slow acting. The slow system ensured that, as listening conditions changed, the system gain changed only very slowly. This preserved short-term level variations while removing or reducing slower fluctuations in overall level. To protect the user from uncomfortable loudness produced by brief intense sounds, the fast system reduced the gain for such sounds. A block diagram of the operation of the dual front-end AGC system is shown in Fig. 1. The stage of equalization (to compensate for the non-flat response of the microphone and the analog high-frequency emphasis applied prior to the ADC) was described earlier. At this stage, a high-frequency emphasis, called pre-emphasis, was applied prior to the measurement of level for the slow control system. The pre-emphasis applied a gain rising at 3.3 dB octave between 500 and 4000 Hz, giving a total of 10 dB. Below 500 Hz the gain was 0 dB, and above 4000 Hz it was 10 dB. The pre-emphasis served two purposes. First, it prevented the AGC action from being dominated by intense low-frequency sounds such as car noise. Second, it prevented the AGC from being excessively activated by the user's own voice (Moore, 1993).

The root-mean-square (rms) value in a given frame after pre-emphasis was calculated, and the slow control parameter was determined by an exponential averaging of the frame rms value with previous frame values. This control parameter determined the gain applied to the output of the pre-emphasis stage. The rms value of the resulting signal was calculated for each frame, and the fast control parameter was determined by a second exponential averaging across frames. The gain determined in this way was applied in a second variable gain stage.

A schematic input-output function for the DUAL-HI system is shown in Fig. 2. For signals with long-term levels below the slow threshold of 63 dB SPL, the output was linearly related to the input. Signals with long-term levels exceeding the slow threshold were held within the range of comfortable output levels of the user, shown shaded in Fig. 2. For this system, the fast component reduced the gain only when the current frame rms value at the output of the slow system was more than 8 dB above the mean level at the output of the slow system (as determined by the exponential averaging employed to calculate the slow gain). That is, the compression threshold for the fast system tracked the mean level at the output of the slow control system. This applied whenever the compression threshold of the slow system was exceeded. For signals whose long-term mean level was below the compression threshold of the slow system, the fast system had a fixed compression threshold of 71 dB.

The DUAL-LO system was intended to give some im-

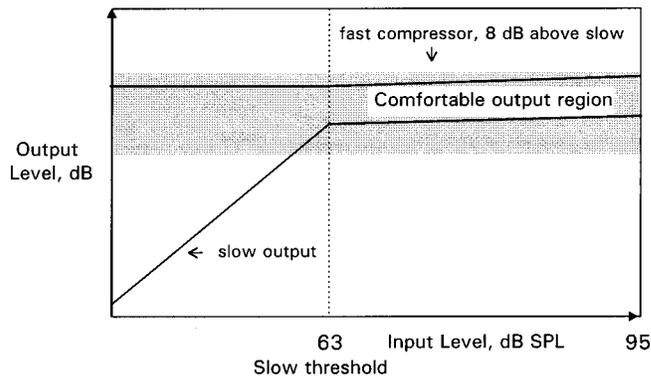


FIG. 2. Schematic illustration of the operation of the DUAL-HI system. The lower solid line shows the input-output function of the slow control system. The upper solid line indicates the output level at which the fast control system was activated. For input levels above 63 dB, the fast system was only activated when the momentary output of the slow system was 8 dB above the current mean value at the output of the slow system.

pression of the levels of sounds in the environment. The compression ratio for the slow component was 3 and the compression threshold was 55 dB SPL. An input-output function for this system is shown in Fig. 3. For the DUAL-LO system, the fast system played the role of a limiter of the output of the slow control system. The compression threshold of the fast system corresponded to the output level of the slow system when the input level was at its maximum value of 95 dB SPL (set by the analog compression limiter in the earpiece). As shown in Fig. 3, the "headroom" of the fast system decreased as the mean output level of the slow system rose. For the highest possible input level (determined by the compression limiter prior to the ADC), the headroom was 6 dB.

One feature of our original design, which was retained in the present implementation, was that, during brief intense transients, the slow gain changed little. This was achieved by limiting the value of the frame rms value if its momentary value was well above the current mean value used for calculating the slow control parameter. Thus, the gain after a transient had ended was almost the same as that before the transient. For the DUAL-HI system, the attack time for the slow component was 325 ms. The recovery actually had two components. There was a "hold" time of up to 500 ms, during

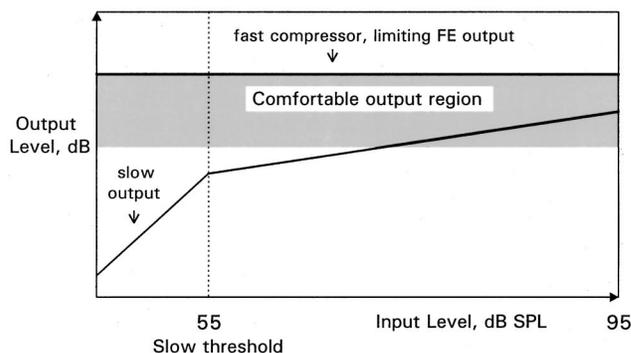


FIG. 3. As Fig. 2, but for the DUAL-LO system. The fast control system was activated when the output of the slow system exceeded a fixed value. Thus, the compression threshold for the fast control system was a variable amount above the current mean value at the output of the slow system.

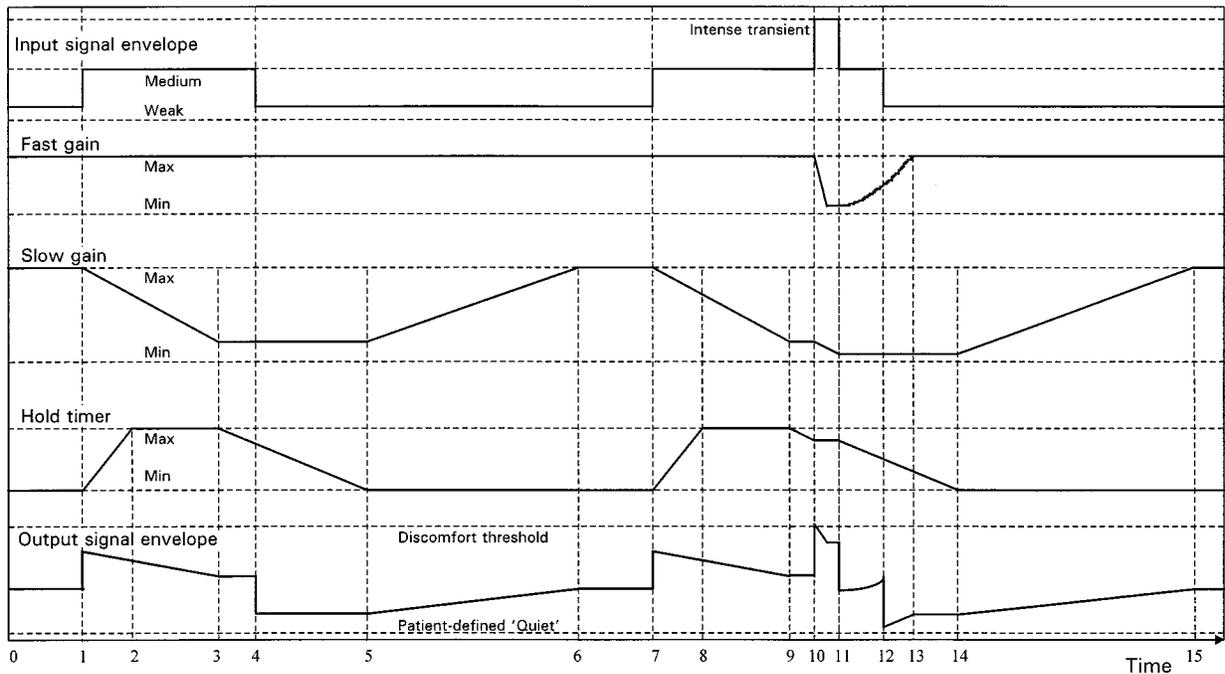


FIG. 4. An illustration of the operation of the dual front-end AGC system in response to moderate changes in input level and to a brief intense transient. The traces show (from top to bottom): the envelope of the input signal; the gain applied by the fast control system; the gain applied by the slow control system; the value of the hold timer; and the envelope of the output signal.

which the gain did not change at all, and then a recovery time of 1000 ms. The hold time prevented the slow gain control parameter from recovering during brief pauses, and so reduced pumping effects. The hold time was achieved by a hold timer, whose value was increased whenever the slow control parameter was decreasing (so as to decrease the gain), and whose value otherwise gradually declined. The slow gain was only increased when the hold timer reached its minimum value (zero).

The attack time of the fast component was chosen in the following way. Previously, we have used a short attack time of 3 ms to avoid brief overloading of the circuitry. However, overload can also be avoided by introducing a small time delay into the audio path, and this allows the use of a slightly longer attack time. This has the advantage that the level of a transient can be measured over a time scale that is of more perceptual relevance and that reduces the influence of circuitry noise. The attack time was 5 ms, and there was a 4-ms compensating delay to the audio signal. This was sufficient to prevent overloads. The recovery time was 75 ms, similar to the value found to be optimal in our earlier work (Moore *et al.*, 1991).

The attack and recovery *constants* (i.e., the time constants of the exponential averagers used to determine the slow and fast control parameters) were the same for the DUAL-LO as for the DUAL-HI system. However, because the definitions of attack and recovery times are usually referred to as the *output* settling time, a different settling time results if a different compression ratio is used. The attack time was 350 ms and the recovery time was 1080 ms.

The action of the dual front-end AGC systems is illustrated schematically in Fig. 4. Consider the action of the circuit at the times indicated in the figure.

- (0) A weak signal, below the threshold of the slow control parameter, has been present for some time. The fast and slow gain are at their maximum values. The hold timer value is zero. The signal output is at a low-to-moderate level.
- (1) A moderately intense signal is turned on. The slow gain gradually decreases and the hold timer increases. The fast gain is unaffected since the signal is below the fast threshold. The output initially increases, and then gradually decreases as the slow gain changes.
- (2) The hold timer reaches its maximum value.
- (3) The slow gain reaches a value appropriate for the signal presented. The hold timer starts to decrease since the slow gain is no longer decreasing. The signal output is now stable at a moderate level.
- (4) The moderately intense signal becomes weak. Apart from the hold timer counting down, nothing else happens; the slow gain remains stable. The signal output drops to a low level.
- (5) The hold timer reaches zero and allows the slow gain to start to recover. The signal output starts to increase back towards the original low-to-moderate level.
- (6) The slow gain reaches its maximum value. The signal output is now back at the original low-to-moderate level.
- (7) As in (1) above, a moderately intense signal is turned on. The slow gain gradually decreases and the hold timer increases. The fast gain is unaffected since the signal is below the fast threshold.
- (8) The hold timer reaches its maximum value.
- (9) As in (3) above, the slow gain reaches a value appropriate for the signal presented. The hold timer starts to decrease since the slow gain is no longer decreasing.

- (10) A brief, intense transient is presented. The fast gain starts decreasing. It reaches its minimum value between (10) and (11). The hold timer is not allowed to increase while the transient is present. There is only a small reduction in slow gain during the transient. The fast gain change causes the signal output to decrease rapidly to take the output signal level back into the patient-defined comfort region.
- (11) The transient ceases, so the fast gain starts to recover. The slow gain stops decreasing. The hold timer decreases. The signal output drops considerably due to removal of the transient, but recovers rapidly to almost the same value as before the transient.
- (12) The moderately intense signal becomes weak. The signal output is lower than at (4) as both the fast and slow gain are reduced during the transient.
- (13) The fast gain recovers to its maximum value. The signal output is still slightly lower than at (4) due to the additional slow gain reduction during the transient.
- (14) The hold timer reaches zero and allows the slow gain to start to recover.
- (15) The slow gain recovers back to its maximum value.

Both the DUAL-HI and DUAL-LO systems were followed by a stage of linear frequency-dependent gain, i.e., frequency-response shaping. The required gains were initially calculated for the center frequencies of the eight bands described earlier. Gains at the frequencies corresponding to the 32 bins in the FFT were then calculated by interpolation and extrapolation from the eight band values. The interpolation removed step changes in gain values which would otherwise have given rise to excessive oscillatory time-domain responses. An inverse FFT was performed on each modified frame and the resulting time frame was windowed (same function as before) and returned to the operating system, where successive frames were combined with 50% overlap. The resultant waveform was fed to the DAC.

B. The FULL-4 system

This system was similar to that described in Moore *et al.* (1999b). To obtain the four channels, bands were combined as follows:

- Channel 1: bands 1 and 2,
- Channel 2: bands 3 and 4,
- Channel 3: bands 5 and 6,
- Channel 4: bands 7 and 8.

For each compression channel, the signal power within a given frame was summed across all bins included in that channel; the square root of the power was taken to give an rms value. The logarithm of the channel rms value was calculated. The log value was low-pass filtered (across frames, i.e., across time) with a two-pole, Bessel-derived, filter. The step response of the filter showed no over- or undershoot for either positive or negative steps. The cutoff frequency was 24 Hz, giving a step-response delay (time to 50% of final value) of 6 ms.

The filtered envelope value for a given channel was used to calculate the gain value for that channel. The gain values for the channels were used to calculate gain values at the

center frequencies of the eight bands, as described in Moore *et al.* (1999b). Gains at the frequencies corresponding to the 32 bins in the FFT were calculated by interpolation and extrapolation for the eight band values.

The gain signals so derived were applied to the FFT of an earlier frame stored in a delay line. The delay line introduced a delay to the audio signal that matched the delay associated with the 50% point on the step response of the two-pole envelope filter. This ensured that the audio and gain signals were synchronized, reducing the signal overshoot normally associated with the response of compression systems to abrupt increases in level. A potential problem with this system is that, just before an abrupt decrease in level, the output increases (“preshoot”). However, the two-pole filter had a slightly asymmetric step response which reduced this problem. An inverse FFT was performed on the modified frame and the resulting time frame was windowed (same function as before) and returned to the operating system, where successive frames were combined with 50% overlap. The resultant waveform was fed to the DAC.

In many “fast” compression systems, the “effective” compression ratio decreases as the envelope modulation rate and the modulation depth of the input signal increase (Braidia *et al.*, 1982; Stone and Moore, 1992; Verschuure *et al.*, 1996). Such systems typically use exponential averaging of the envelope (corresponding to use of a first-order filter), with an asymmetry between the attack and release times. In contrast, the present system used a second-order filter to average the logarithm of the rms value, which gave nearly equal attack and release times. As a result, the effective compression ratio was equal to or close to the nominal compression ratio for envelope modulation rates up to the cutoff frequency of the low-pass filter used to derive the gain signal, i.e., 24 Hz, and the effective compression ratio depended relatively little on the modulation depth of the input signal. Thus, compression was effectively applied over the whole range of envelope modulations important for speech signals (Houtgast *et al.*, 1980; Plomp, 1983).

C. DUAL-4 system

This system was a combination of the dual front-end AGC system and the FULL-4 system. The slow compression ratio was 1.7, and the slow threshold was 45 dB SPL. The four-channel, full dynamic range compression system that followed was identical to the FULL-4 system. Since the FULL-4 system required a time delay of 6 ms, the 4-ms delay of the audio signal introduced for the fast component of the dual front-end AGC system was removed.

II. FITTING PROCEDURES

For each compression system, the fitting procedure had two stages: an initial fitting based on the audiogram of the individual subject, and an adaptive fine-tuning procedure based on the subject’s responses to speech stimuli. The procedures differed across systems, but they had in common the following goal for the initial fitting: a speech-shaped noise with an input level of 65 dB SPL should give a reasonably flat specific loudness pattern (equal loudness per critical

TABLE I. Values of the intercept (INT, dB) in the Cambridge formula for each audiometric frequency (kHz).

Frequency	0.125	0.25	0.5	0.75	1.0	1.5	2.0	3.0	4.0	5.0
INT	-11	-10	-8	-6	0	-1	1	-1	0	1

band) over the frequency range 500 to 4000 Hz, and the overall loudness should be similar to that evoked in a normal listener by 65-dB speech. This goal was approximated using the ‘‘Cambridge’’ formula (Moore and Glasberg, 1998), which was developed using the loudness model proposed by Moore and Glasberg (1997).

The Cambridge formula is

$$IG=HL\times 0.48+INT, \quad (1)$$

where HL stands for the absolute threshold in dB HL and INT is a frequency-dependent intercept. The value of INT is given in Table I for each audiometric frequency. Above 5 kHz, gains are equal to the value at 5 kHz, up to the upper frequency limit of the aid (7 kHz for the aids used in the present study).

The goal of achieving a flat specific loudness pattern is restricted to the frequency range 500 to 4000 Hz, since that is the frequency range that is most important for speech intelligibility. Below 500 Hz, the gains are reduced below those required to achieve a flat specific loudness pattern, in order to reduce masking of the speech by low-frequency environmental sounds, such as car noise and noise from ventilation and air-conditioning systems. Note that these gain reductions are already incorporated in the intercept values given in Table I.

The adaptive fine-tuning procedures were developed from those described by Moore *et al.* (1998, 1999a). The sentences used in these procedures were taken from the ASL lists (MacLeod and Summerfield, 1990). Two speech levels were used in the procedures, 60 and 85 dB SPL (rms levels with a flat frequency weighting, measured using a CEL 414 precision sound-level meter at the point corresponding to the center of the listener’s head after the listener was removed from the sound field). The speech at 60 dB SPL was digitally filtered so as to have the long-term average speech spectrum published by Byrne *et al.* (1994). The speech at 85 dB SPL was additionally digitally filtered to give a slight boost to the medium and high frequencies, using the filter characteristic given in Moore *et al.* (1999a). This made the spectrum more like that of speech spoken with ‘‘raised effort’’ (Pearsons *et al.*, 1976). The overall level of the speech was 85 dB SPL after taking into account this filtering.

In everyday life, speech rarely occurs in the absence of some background noise, especially when the speech has a high level. To make the stimuli more representative of everyday life, a speech-shaped background noise was presented with each sentence at a level 15 dB lower than that of the speech. The level of the noise was chosen so that it was clearly audible, but low enough to produce minimal masking of the speech. The sentences were stored on computer disk, and replayed via a 16-bit digital-to-analog converter on a card inside the PC (Soundblaster® compatible). The output of the card was fed via an amplifier to a Bowers and Wilkins DM600 loudspeaker. The listener was seated about 70 cm

from the loudspeaker in a small room (3.4×3.1×2.3 m) with moderate reverberation (reverberation time, RT₆₀=380 ms). There was a moderate level of background noise (about 45 dB SPL) generated by the computer fan. We chose this environment deliberately (rather than a sound-attenuating chamber), to ensure that the procedure would work in a setting typical of that found clinically.

A. The dual front-end AGC systems

1. Initial fitting

The fitting procedures were essentially the same for the DUAL-HI and DUAL-LO systems. These systems can be considered as a sequence of the following stages: a pre-emphasis stage, the dual front-end AGC system, and a ‘‘gainbox’’ which applied a frequency-dependent gain following the AGC system. The initial frequency response shape required for a 65-dB SPL speech-shaped noise input was calculated from the Cambridge formula. Insertion gains were initially calculated for eight frequency bands. For this purpose, the band center frequencies were rounded to the nearest audiometric value, i.e., 250, 500, 1000, 1500, 2000, 3000, 4000, and 6000 Hz.

The values were interpolated and extrapolated to 32 values, corresponding to the bins in the FFT. The gain values, GB, applied in the gainbox were adjusted to allow for the fact that high-frequency emphasis was applied prior to the dual front-end AGC, so that the correct insertion gains were obtained at each frequency. For frequencies below 0.25 kHz, insertion gains were set to the value for 0.25 kHz. For frequencies above 6.0 kHz, insertion gains were set to the value for 6.0 kHz. Linear interpolation on a log-frequency scale was used to determine insertion gains for intermediate frequencies.

2. Fine tuning using the adaptive procedure

The goal was to adjust the overall level so that speech with an input level of 85 dB SPL was presented at the ‘‘highest comfortable level’’ and had an acceptable tonal quality. Steps to achieve this were as follows:

- (1) A sentence was presented at 85 dB SPL. The following was put on the computer screen and on a piece of paper in front of the subject:

‘‘Please judge the loudness of this sentence as compared to the HIGHEST volume that you would like to listen to for a long time:

7 Far too loud,
 6 Much louder than I like,
 5 Somewhat louder than I like,
 4 The highest volume I like,
 3 Somewhat softer than the highest volume I like,
 2 Definitely softer than the highest volume I like,
 1 Far too soft.’’

The aim of this stage of the procedure was to adjust all GB values to achieve the target response of ‘‘The highest volume I like.’’ The subject responded verbally with a number, and the experimenter entered this number on the computer keyboard. The parameters of the digital hearing aid were then updated automatically by the computer, according to the following rules:

If 7 was pressed, all values of GB were decreased by 4 dB.

If 6 was pressed, all values of GB were decreased by 2 dB.

If 5 was pressed, all values of GB were decreased by 1 dB.

If 4 was pressed, values of GB were left unchanged.

If 3 was pressed, all values of GB were increased by 1 dB.

If 2 was pressed, all values of GB were increased by 2 dB.

If 1 was pressed, all values of GB were increased by 4 dB.

This was repeated, using a new sentence each time, until the target response (4) was given twice in succession. The next stage of the procedure then commenced.

- (2) A sentence was presented at 85 dB SPL. The listener was required to judge the quality of the speech on a seven-point scale, displayed on the computer screen and on a piece of paper in front of the subject:

7 Uncomfortably tinny,

6 Very tinny,

5 Tinny,

4 Neither tinny nor boomy,

3 Boomy,

2 Very boomy,

1 Uncomfortably boomy.

For this scale, the response ‘‘neither tinny nor boomy’’

(4) was deemed optimal. The response was used to adjust the relative gains at low and high frequencies in a symmetrical way. The low-frequency region was nominally centered at 750 Hz and the high-frequency region was nominally centered at 3000 Hz. Denote the rating by R . Denote the value of GB at 750 Hz by GB_{750} and the value of GB at 3000 Hz by GB_{3000} . The values of GB_{750} and GB_{3000} were adjusted according to the following rules:

If $R=7$, GB_{3000} was decreased by 3 dB and GB_{750} was increased by 3 dB.

If $R=6$, GB_{3000} was decreased by 2 dB and GB_{750} was increased by 2 dB.

If $R=5$, GB_{3000} was decreased by 1 dB and GB_{750} was increased by 1 dB

If $R=4$, the parameters were left unchanged.

If $R=3$, GB_{3000} was increased by 1 dB and GB_{750} was decreased by 1 dB

If $R=2$, GB_{3000} was increased by 2 dB and GB_{750} was decreased by 2 dB.

If $R=1$, GB_{3000} was increased by 3 dB and GB_{750} was decreased by 3 dB.

The *change* in GB was interpolated between 750 and 3000 Hz (linear interpolation of dB on a log-frequency

scale) and used to calculate new values of GB between those two frequencies. Below 750 Hz, GB values were changed by the same amount as at 750 Hz. Above 3000 Hz, GB values were changed by the same amount as at 3000 Hz. This procedure was repeated, using a new sentence each time, until the 85-dB speech was judged as ‘‘neither tinny nor boomy’’ two times in a row.

- (3) Steps (1) and (2) were repeated until the sentences were judged to have the appropriate loudness and tonal quality two times in a row. That concluded the adaptive fitting procedure.

To determine the extent to which the adaptive procedure resulted in changes in overall gain, the mean change in gain as a result of the fine tuning, averaged across the frequencies 0.25, 0.5, 1, 1.5, 2, 3, 4, and 6 kHz, was calculated for each ear of each subject. The gains here are those for a 65-dB SPL speech-shaped noise input. For the DUAL-HI system, the mean change was in the range 0 to -3 dB for 12 out of 16 ears. For subject SZ the mean gain change was -7 dB for the right ear and -5 dB for the left ear. For subject RW the mean gain change was -12 dB for the left ear. For subject GW the mean gain change was -17 dB for the left ear (but only -2 dB for the right). For the DUAL-LO system, the mean change for individual subjects was in the range 0 to -5 dB for 13 out of 16 ears. For subject SZ the mean gain change was -7 dB for the left ear. For subject RW the mean gain change was -12 dB for the right ear and -20 dB for the left ear.

Changes in frequency-response *shape* as a result of the fine tuning were generally very small. The largest change was for the right ear of subject VW with the DUAL-HI system, where the gain at high frequencies relative to low frequencies was decreased by 6 dB. On average, the relative gain change was less than 2 dB.

B. The four-channel fast compression system (FULL-4)

1. Initial fitting

Our procedure for the initial fitting of the four-channel compression hearing system was the same as described by Moore *et al.* (1999a). Briefly, it is based on two constraints. The first has been described already; for a speech-shaped noise with a level of 65 dB SPL, the insertion gain as a function of frequency should be as prescribed by the Cambridge formula. The second was that, as far as possible, speech should be audible over its entire dynamic range in each frequency channel when the overall speech level was 45 dB SPL. Taken together, these two constraints can be used to define the required gain, compression ratio, and compression threshold in each channel of the compression system. In practice, however, the second constraint is modified to take into account the empirical finding that high compression ratios (greater than 2–3) have deleterious effects on speech intelligibility in quiet and in noise (Moore *et al.*, 1992b; Plomp, 1994; Verschuure and Dreschler, 1996; Goedegebure *et al.*, 1996). To avoid such effects, the maximum compression ratio in any one channel was limited to 2.92. This had

the effect that, when the hearing loss at high frequencies was greater than 50–60 dB, the audibility of speech at 45 dB SPL was not fully restored.

The implementation of the initial fitting procedure is essentially the same as described in Moore *et al.* (1999a, 1999b), and the reader is referred to those papers for details. Briefly, to preserve accuracy in frequency-response shaping, gains and compression ratios were calculated for an eight-channel system, and the appropriate parameters for the four-channel system were derived from these.

2. Adaptive fine tuning

The adaptive fine-tuning procedure was very similar to that described by Moore *et al.* (1998, 1999a); the reader is referred to those papers for details. The goal of the procedure was that speech at 85 dB SPL should be judged as “loud,” speech at 60 dB SPL should be judged as “quiet,” and speech at both levels should have an acceptable tonal quality. The procedure involved adjusting the low-level (50 dB) gain and high-level (80 dB) gain at low frequencies and at high frequencies. Call these gains G50 (lf), G50 (hf), G80 (lf), and G80 (hf). The low-frequency region was nominally centered at 750 Hz (i.e., roughly the boundary of the two lower channels), and the high-frequency region was nominally centered at 3000 Hz (i.e., roughly the boundary of the two upper channels).

The following stages were used:

- (1) A sentence was presented at an overall level of 85 dB SPL. The target was that this should be judged as loud. If the judgment differed from the target, all post-AGC gains (i.e., the linear gains following the compression processing) were changed by the same amount; the compression ratios were not altered. The step size in gain was as described by Moore *et al.* (1998). This process was repeated until the target judgment of loud was obtained.
- (2) A sentence was presented at an overall level of 60 dB SPL. The target was that this should be judged as quiet. If the judgment differed from the target, the low-level gains, G50 (lf) and G50 (hf), were both changed by the same amount, leaving the high-level gains unaltered. This meant that the compression ratios changed. The step size in gain was as described by Moore *et al.* (1998). This process was repeated until the target judgment of quiet was obtained.
- (3) A sentence with a level of 85 dB SPL was presented. The listener was required to judge the quality of the speech on a seven-point scale going from “uncomfortably tinny” to “uncomfortably boomy.” The response “neither tinny nor boomy” was deemed optimal. The relative post-AGC gains at 750 and 3000 Hz were adjusted until the target response was achieved. For example, if the sound was rated as tinny, the post-AGC gain at 3000 Hz was decreased by 1 dB and the post-AGC gain at 750 Hz was increased by 1 dB; steps of 2 and 3 dB were used for responses that deviated more from the target response.

- (4) A sentence with a level of 60 dB SPL was presented. The listener was required to judge the quality of the speech on a seven-point scale going from “uncomfortably shrill” to “very muffled.” The response “neither shrill nor muffled” was deemed optimal. The relative low-level gains at 750 and 3000 Hz, G50 (lf) and G50 (hf), were adjusted until the target response was achieved. For example, if the sound was rated as muffled, G50 (hf) was increased by 1 dB and G50 (lf) was decreased by 1 dB; steps of 2 and 3 dB were used for responses with greater deviations from the target response.

This whole sequence was repeated until stable values were achieved and the target responses were consistently given. Generally, only one repetition was required, although two repetitions were sometimes performed. If at any time the adaptive procedure called for a compression ratio greater than 2.92 in a specific channel, the ratio was limited to 2.92.

To determine the extent to which the adaptive procedure resulted in changes in overall gain, the mean change in gain as a result of the fine tuning was calculated in the same way as for the DUAL-HI and DUAL-LO systems, for a 65-dB speech-shaped noise input. The overall average change was -1 dB. The largest decrease in gain for any individual ear was 5.5 dB, while the largest increase was 3 dB. For each ear of each subject, the mean change was subtracted from the change at each channel center frequency. Any deviations of the resulting numbers from zero indicate changes in frequency response *shape* as a result of the fine tuning. In the majority of cases, the gains deviated by less than ± 1 dB. In other words, changes in frequency-response shape were generally very small. In the most extreme case (GW, right ear), the frequency-response shape after fine tuning differed from that determined by the initial fitting procedure by about 5 dB across the range 0.25 to 6 kHz; the final gain was about 0.5 dB higher than the initial gain at 0.25 kHz and 4.4 dB lower than the initial gain at 6 kHz.

To assess the extent to which compression ratios changed as a result of the adaptive fine tuning, we calculated the ratio of the compression ratio for a given channel at the end of the fine tuning, called CR (final), and the compression ratio determined by the initial fitting procedure, called CR(start). This was done separately for each ear of each subject. Ratios close to 1 indicate that the compression ratio following fine tuning, CR (final), was close to the compression ratio determined by the initial fitting procedure, CR (start). Most (40 out of 64 cases) of the ratios were in the range 0.9 to 1.1, indicating little change in compression ratio as a result of the fine tuning. The largest ratio was 1.34 (2.92/2.17), which occurred for the right ear of subject SZ for channel 4 and for the right ear of subject GW for channel 2. The smallest ratio was 0.56 (1.48/2.63), which occurred for the right ear of subject JW for channel 4. The geometric mean ratios were 1.06 for channel 1, 1.03 for channel 2, 1.02 for channel 3, and 1.00 for channel 4.

TABLE II. Audiometric thresholds (dB HL) for each ear of the eight subjects used in the trial.

Subject/ Ear	Frequency (kHz)							
	0.25	0.5	1.0	1.5	2.0	3.0	4.0	6.0
DP-R	10	15	40	50	40	50	55	45
DP-L	5	5	25	40	40	55	50	45
ET-R	50	55	55	55	55	60	70	60
ET-L	55	60	70	65	65	70	75	70
GW-R	40	50	65	60	60	60	65	55
GW-L	35	40	60	60	55	55	60	55
RW-R	30	30	45	50	50	55	60	55
RW-L	60	60	65	65	65	60	65	60
VW-R	40	55	70	75	65	75	85	75
VW-L	45	60	70	65	70	55	50	60
PY-R	30	40	45	50	55	55	55	65
PY-L	35	40	45	50	60	60	65	70
SZ-R	40	40	35	35	40	40	45	60
SZ-L	35	35	35	45	45	45	55	70
JW-R	15	25	40	40	40	55	65	65
JW-L	15	30	40	40	45	55	65	65

C. The DUAL-4 system

1. Initial fitting

The initial fitting was based on the same two constraints as for the FULL-4 system. Details of the implementation are given in the Appendix.

2. Adaptive fine tuning

This was essentially the same as for the FULL-4 system, except that if, at any time, the procedure called for a compression ratio greater than 2, the ratio was set to 2. Overall gain changes following fine tuning (calculated for a 65-dB speech-shaped noise input, as described earlier) were in the range -5 to $+2$ dB for both ears of seven out of the eight subjects. For subject SZ, the mean gain change was -10 dB for both the right and left ears. Changes in frequency response shape (for a 65-dB speech-shaped noise input) were generally very small. The largest change was for the right ear of subject ET, where the gain at high frequencies relative to low frequencies was decreased by 6 dB. On average, the relative gain change was less than 2 dB.

III. METHOD

A. Subjects

Eight moderate-to-severely hearing impaired subjects participated. Table II shows the absolute thresholds in dB HL for each of the subjects. All were diagnosed as having bilateral cochlear hearing loss and all had active lifestyles. No subject showed any significant conductive loss (the air-bone gap was less than 10 dB). All subject showed signs of loudness recruitment as indicated by judgments of loudness for high-level sounds. Four of the subjects (DP, SZ, RW, and JW) regularly used conventional linear hearing aids (with output limiting), and the remaining four (PY, ET, GW, and VW) used two-channel, fast-acting, full-dynamic range compression hearing aids. All subjects were paid for their participation.

B. Experimental design

Each subject received each of the four compression systems in succession. The order of presentation of the systems was counterbalanced across subjects, using two repetitions of a 4×4 Latin Square design. After the fitting of a given system, subjects were asked to use that system in their everyday lives for a period of 2–3 weeks. All subjects reported that they wore the experimental aids most of the time during this time. At the end of this period, they were asked to complete the abbreviated profile of hearing aid benefit (APHAB) questionnaire of Cox and Alexander (1995) (slightly modified to allow for differences between American and British English), and were brought into the laboratory for formal tests. The latter included: (1) measures of speech intelligibility in quiet, using AB word lists (Boothroyd, 1968), presented at levels of 50 and 80 dB SPL; and (2) measures of speech reception thresholds (SRTs) using adaptive sentence lists (ASL) presented in a steady speech-shaped noise, and that same noise amplitude modulated with the envelope of speech from a single talker. Noise levels were 60 and 75 dB SPL. At the end of the evaluation, the subjects were fitted with the next appropriate compression system, and sent away for another 2–3 weeks. This procedure was continued until all four systems had been used and evaluated for each of the eight subjects. Each test session, comprising both evaluation and program fitting, typically lasted between 1 and 1.5 h.

Subjects were thoroughly instructed in the use of the hearing aid systems, including operating instructions, battery replacement, and troubleshooting procedures, prior to leaving the laboratory, and they were encouraged to get in contact, should any problems arise.

C. Fitting procedures

The fitting procedures have already been described in detail. Only one ear was fitted at a time, the earpiece in the other ear being fitted, but no signal being delivered. However, all of the aids were programed to deliver binaural stimuli, and all everyday listening and testing was done using binaural amplification. At the completion of the fitting procedure for a given system, subjects were asked to make informal qualitative judgments on the comfort of their hearing aids, in a variety of sound environments. This was done to ensure that the aids had been fitted appropriately. In two cases (SZ and RW), towards the beginning of the study, it was necessary to repeat the fitting procedure when it was clear that there were problems with the initial fit. The fitting procedure rarely took more than 5–10 min per ear, especially towards the middle to end of the study, when subjects were more acquainted with the task involved.

D. Formal evaluations

Speech intelligibility testing was carried out in a sound-attenuating chamber. Subjects were first evaluated with AB word lists, which were scored phonemically, and presented at levels of 50 and 80 dB SPL, as measured at ear level, using a CEL-414 sound-level meter. All speech material had previously been recorded onto digital audiotape (DAT), and it was played back via a Sony DTC-750 recorder/player,

Quad 306 amplifier, and Monitor Audio MA4 two-way loudspeaker. Two lists of AB words were used for each presentation level. Hence, the final score was the percent correct out of 60 phonemes.

SRTs for the ASL sentence material (MacLeod and Summerfield, 1990) were measured using an adaptive procedure. The background noise was presented at either 60 or 75 dB SPL, and the speech level was initially set to give a +5-dB speech-to-background ratio. Each sentence contained three key words. If the subject scored two or more key words correct, the level of the speech was decreased by 5 dB. If the subject scored less than two key words correct, the speech level was increased by 5 dB. After two turnpoints, the step size was decreased to 3 dB. The level of the speech was controlled by the tester using a Marconi continuously variable attenuator. Testing continued until a complete sentence list had been presented for each condition. Probit analyses (Finney, 1971) were then conducted to determine the SRT (defined as the 50%-correct point on the psychometric function) for each condition.

Subjects were given one practice list in each test condition before formal evaluations were conducted. The practice lists were used repeatedly throughout the course of the study (i.e., at the start of the formal evaluation of each compression program) due to limitations in the number of lists available. However, for the formal evaluations, the number of lists was sufficient that no repetitions were required. The order of presentation of the test lists for the formal evaluations was such that each list was used once for each hearing aid program and each testing level.

IV. RESULTS

A. Informal reports

All subjects reported that, for all of the compression systems, the aids gave acceptable loudness in most everyday listening situations. They were not unduly bothered by low-level background sounds, and intense sounds were generally not uncomfortably loud. The aids were fitted with a push-button volume control, but subjects reported that this was rarely used (they were, in fact, discouraged from using it). It was, however, sometimes used to reduce the volume in situations where there was a high level of background noise. Many normally hearing people might also want to reduce the volume in such situations!

B. APHAB test

The APHAB test (Cox and Alexander, 1995) requires subjects to rate how often they have problems in specific situations, such as “unexpected sounds, like a smoke detector or alarm bell are uncomfortable” or “when I am having a quiet conversation with a friend, I have difficulty understanding.” Response alternatives range from “always (99%)” to “never (1%).” Low percentages indicate good performance. The results are grouped into four subscales: ease of communication (EC), understanding in reverberant environments (RV), understanding in background noise (BN), and aversiveness of sounds (AV). A total score is also

given. The results from the APHAB test for the individual subjects and for the mean are shown in Fig. 5.

It is clear that there was marked variability across subjects. Some subjects showed reasonably consistent preferences for one or the other system. For example, GW reported fewer problems with the DUAL-HL system than with any of the other systems, for each subscale and for the total score. Subject SZ showed a preference for the DUAL-LO system, especially for the aversion subscale (AV). Subject DP showed a preference for the FULL-4 system, especially for the aversion subscale (AV). However, for the mean scores (bottom-right panel), there was no clear difference between the systems. A within-subjects analysis of variance (ANOVA) based on the arcsine-transformed scores, with the factors type of compression system (DUAL-LO, DUAL-HI, DUAL-4, and FULL-4) and APHAB category (EC, RV, BN, AV, and total) showed that neither main effect was significant at the 0.05 level. The two-way interaction was also not significant.

As the score for each subscale was based on the answers to six questions, it was also possible to perform ANOVAs on the arcsine-transformed data of the individual subjects, treating the individual answers as replications. Some questions were not answered as subjects had not encountered the specific situation referred to in the question; these cases were treated in the ANOVAs as missing data, and they account for the fact that the degrees of freedom associated with the error term vary across subjects. Significant effects of type of compression were found for GW ($F(3,64)=9.02, p<0.001$), DP [$F(3,69)=9.79, p<0.001$], and SZ [$F(3,73)=16.87, p<0.001$]. *Post hoc* tests (Neuman Keuls) indicated that, for GW, scores were lower (better) for DUAL-HI than for any of the other compression systems (all $p<0.005$). Scores for the remaining compression systems did not differ significantly. For DP, scores were lower for FULL-4 than for any of the other systems (all $p<0.005$). For SZ, scores were lower for DUAL-LO than for any of the other compression systems (all $p<0.001$). Scores for the remaining compression systems did not differ significantly.

It is instructive to compare the means scores with those reported by Cox and Alexander in their normative study of the APHAB test (Cox and Alexander, 1995). Two subscales are of particular interest here. The score for EC is based partly on questions about the ability to hear low-level speech in quiet situations. Cox and Alexander reported that the mean frequency of problems for the EC scale for subjects wearing “conventional analog” hearing aids was 24%. Our subjects gave similar mean scores, ranging from 17% for the DUAL-HI system to 25% for the DUAL-LO system. Note that the subjects in the study of Cox and Alexander almost certainly adjusted the volume controls on their hearing aids to deal with different listening situations; subjects probably increased the volume in situations where they were listening to low-level speech in quiet. Our subjects reported that they did not adjust the volume control in these situations, but they nevertheless gave comparable scores.

The score for AV primarily reflects aversive responses to intense environmental sounds. Cox and Alexander reported that the mean frequency of problems for the AV scale

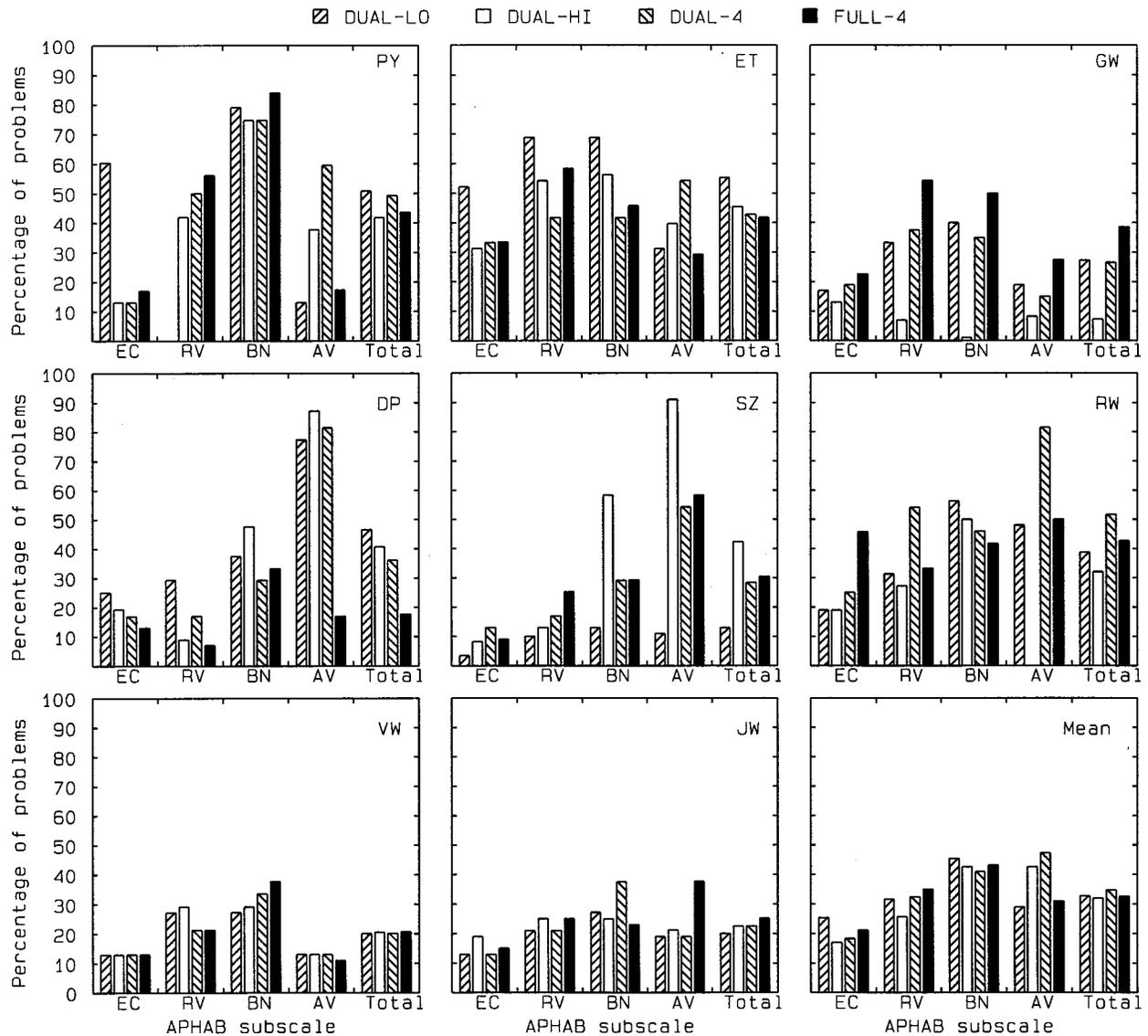


FIG. 5. Results of the APHAB questionnaire, plotted as percentage of reported problems. Each panel shows results for one subject or for the mean (bottom right panel). In each panel, results are grouped by APHAB subscale (EC, RV, BN, or AV) or the total score. Scores for the individual compression systems are indicated by the type of shading of the bars, as indicated in the key at the top.

for subjects wearing hearing aids was 55%. Our subjects gave mean scores ranging from 29% for DUAL-LO to 46% for DUAL-4, markedly lower than reported by Cox and Alexander. This indicates that the compression systems in our aids protected the users from unpleasantly loud sounds more effectively than the aids used by the subjects of Cox and Alexander (which were presumably linear aids with compression limiting or peak clipping).

In summary, the APHAB results support the idea that the compression systems led to satisfactory results in everyday life, at least with regard to the loudness of sounds; the frequency of problems in listening to low-level speech was reasonably low, and subjects were not unduly bothered by loud sounds. However, there was no clear overall difference between the different compression systems.

C. Speech intelligibility in quiet

The mean scores for the measures of speech intelligibility in quiet are shown in Fig. 6. Error bars indicate ± 1 stan-

dard deviation across subjects. The scores are very high for both levels and for all of the compression systems. A within-subjects ANOVA with the factors level and type of compression showed no significant effect of type of compression. There was a significant effect of level [$F(1,7)=8.11$, $p=0.025$], scores being slightly better at the higher level. We can conclude that all of the compression systems allow very good intelligibility of speech in quiet, over a wide range of levels.

D. SRTs in noise

The SRTs for speech in noise, averaged across subjects, are shown in Fig. 7. Scores are expressed as speech-to-noise ratios (SNR) needed for 50% intelligibility, and they are plotted with more negative numbers (indicating better performance) going upwards. Error bars indicate ± 1 standard deviation across subjects. SRTs are shown for each noise level and each noise type (steady, and modulated with the enve-

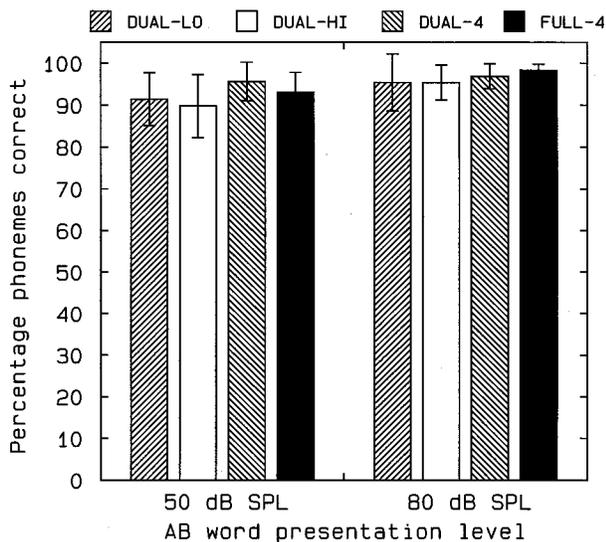


FIG. 6. Mean results for the intelligibility of AB words presented in quiet at 50 dB SPL (left four bars) or 80 dB SPL (right four bars). Scores for the individual compression systems are indicated by the type of shading of the bars, as indicated in the key at the top. Error bars indicate \pm one standard deviation.

lope of a single talker). The horizontal lines show mean scores (\pm one standard deviation) for four normally hearing subjects (mean age 49 years, s.d. 8 years) listening binaurally without hearing aids.

For the steady speech-shaped noise background, mean scores for the hearing-impaired subjects were about -7 dB, and did not vary much across level or across types of compression. The scores were about 3 dB higher (worse) than for the normally hearing subjects. For the speech-shaped noise

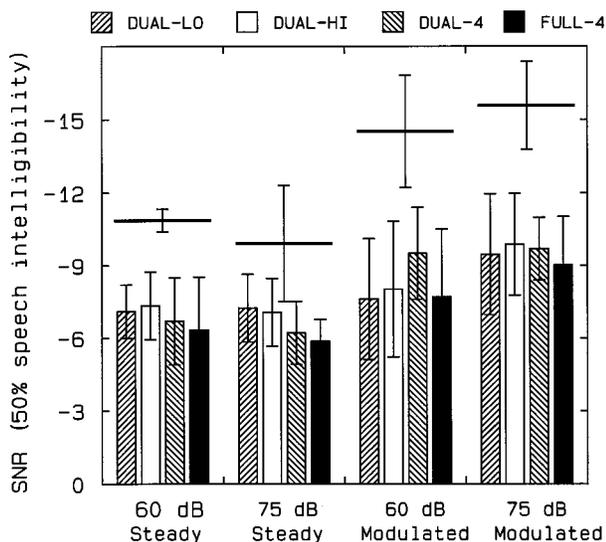


FIG. 7. Mean results for the SRTs in noise, plotted as the speech-to-noise ratio (SNR) needed for 50% of key words to be identified. The scores are plotted with more negative numbers (indicating better performance) at the top. The noise was either steady (left eight bars) or modulated with the envelope of a single talker (right eight bars). The noise level was either 60 or 75 dB SPL. Scores for the individual compression systems are indicated by the type of shading of the bars, as indicated in the key at the top. The bold horizontal lines indicate mean results for four normally hearing subjects listening unaided. Error bars indicate \pm one standard deviation.

modulated with the envelope of a single talker, the scores were slightly lower (better), averaging about -8.5 dB for the 60-dB noise and -9.5 dB for the 75-dB noise. For the lower noise level, scores were slightly better for the DUAL-4 system than for the other systems. For the modulated noise, the SRTs were about 6 dB poorer than for the normally hearing subjects. This is consistent with previous work showing that hearing-impaired subjects are less able than normally hearing subjects to take advantage of the temporal dips in a fluctuating background sound (Carhart and Tillman, 1970; Duquesnoy, 1983; Moore *et al.*, 1995; Hygge *et al.*, 1992; Takahashi and Bacon, 1992; Eisenberg *et al.*, 1995; Peters *et al.*, 1998). However, the SRTs of the hearing-impaired subjects were slightly lower for the modulated noise than for the steady noise, indicating some ability to take advantage of the temporal dips.

We had expected that the two systems including fast-acting compression (DUAL-4 and FULL-4) might give better performance than the other systems in the modulated background noise, since the fast-acting compression can selectively amplify the low-level portions of the target speech falling in the temporal dips of the background. The results for the DUAL-4 system provide some support for this idea, at least for the lower noise level, but the SRTs for the FULL-4 system were not better than those for the DUAL-LO or DUAL-HI systems, which incorporated only slow-acting compression. This finding is consistent with a laboratory study of SRTs measured in the presence of backgrounds with temporal dips, which showed only a small advantage for fast-acting compression over linear amplification (Moore *et al.*, 1999b).

The arcsine-transformed results were subjected to a within-subjects ANOVA, with the factors type of compression, type of noise (steady or modulated), and noise level (60 or 75 dB). The main effect of type of compression was not significant. The main effect of type of noise was significant; $F(1,7) = 44.5$, $p < 0.001$. The main effect of noise level was not significant. There was a significant interaction of type of noise and noise level; $F(1,7) = 60.6$, $p < 0.001$. This reflects the fact that performance was better for the modulated than for the steady noise at the 75-dB level, but not at the 60-dB level. The interaction of type of compression and type of noise was also significant; $F(3,21) = 3.3$, $p = 0.04$. *Post hoc* tests (Neuman Keuls) indicated that, in the modulated noise, SRTs were lower for DUAL-4 than for DUAL-LO and FULL-4 ($p < 0.05$). In the steady noise, SRTs were lower for DUAL-LO than for FULL-4 ($p < 0.05$).

V. DISCUSSION

Although three subjects showed clear differences in their ratings of the different compression systems (as determined by the APHAB test), these differences did not show up in the objective speech measurements for any of the three subjects. The APHAB ratings were also not clearly related to the gains and frequency responses obtained after the adaptive fine-tuning procedure for the different compression systems. For a 65-dB speech-shaped noise input, the insertion gains as a function of frequency were generally similar for the different systems, and there was no clear tendency for the preferred

system (as indicated by the APHAB test), to have higher or lower gain than the other systems, or more or less high-frequency emphasis.

One might expect that the FULL-4 system, which provided independent compression in four frequency bands, would have advantages over the other systems for subjects with sloping hearing losses; in theory, such losses would require frequency-dependent compensation for recruitment. However, the data do not provide any clear support for this idea. Subjects DP and JW both had sloping losses, but only one (DP) showed a preference for the FULL-4 system based on the APHAB results. The SRTs in steady and background noise did not reveal any benefit of the FULL-4 system relative to the other systems for subjects with sloping losses.

It is possible that the ratings given in the APHAB test were affected by changes in criteria over time. For two of the three subjects who showed statistically significant preferences for one system over the other systems (i.e., they reported lower incidences of problems for that system), the preferred system was the one that was tested first. It seems likely that all of the tested systems would have given better performance than the hearing aids that the subjects were used to wearing, as the digital aids were very carefully equalized and fitted and the AGC systems were designed to minimize distortions of various types. The contrast with their own aids may have led these subjects to report a low incidence of problems with the digital compression system that they were tested with initially. However, once they were used to this system, their criteria may have shifted, so that the compression systems tested later received more adverse scores.

Following completion of the main trial, three of the subjects, RW, DP, and SZ were fitted with digital aids in which all four compression systems were available; they could select each of the four systems by pressing a push-button on the body of the Audallion. Whenever the button was pressed, a synthesized voice produced by the Audallion indicated the system that had been selected by saying "one," "two," "three," or "four." Subjects were not told which number corresponded to which system. Subjects were asked to use each system in as many different situations as possible (although these situations were not controlled), and, after a period of several weeks, were asked to give an indication of their overall preferences for the different systems. Subject SZ, whose APHAB results in the main experiment indicated fewer problems with the DUAL-LO system, maintained this preference following the period of experience trying all four programs. Subject DP, whose APHAB results in the main experiment indicated fewer problems with the FULL-4 system, changed his preference, indicating that the DUAL-LO system was preferred. The DUAL-4 system was also liked, although background noise with that system was sometimes judged as too obtrusive. Subject RW's APHAB results did not indicate a clear preference for any specific system. Following the period of experience trying all four programs, he indicated that the DUAL-LO system was somewhat better than the DUAL-HI system or the DUAL-4 system. The DUAL-LO system was clearly preferred over the FULL-4 system.

Given these findings, we think it would be unwise to

attach too much importance to the individual differences in preference revealed by the APHAB test. These may well have been affected by changes in criteria over time. However, the counterbalanced design used in our experiment allows reasonable confidence in the mean scores for the APHAB test, which did not indicate any clear differences in frequency of problems across the compression systems. The informal subsequent trial, allowing direct comparison of the four system in every life, suggests that there may be a slight overall preference for the DUAL-LO system.

A theoretical advantage of fast-acting compression is that it makes dynamic aspects of loudness perception more nearly normal. Loudness recruitment has the effect of magnifying the perceived fluctuation in an amplitude-modulated signal (Moore *et al.*, 1996), and fast-acting compression can compensate for this effect. Recall that the FULL-4 system used in this study was effective in compressing envelope modulation for rates up to 24 Hz. If this aspect of recruitment compensation were important, we would expect better results for the FULL-4 system than for the other systems. Our results do not provide any support for this idea. The mean APHAB results, the measures of speech intelligibility in quiet and in noise, and the preferences indicated in the subsequent trial, did not show any advantage for this system over the other systems. The DUAL-HI and DUAL-LO systems applied almost no compression to the fast modulations in the envelope of speech. There was no evidence from any of the results that these systems were poorer than the other systems. Thus, it does not seem necessary to compress the fast modulations in the input signal. It is possible that different conclusions might be reached if stimuli covering a wider range of levels were used in the laboratory tests. However, this would not have been possible with the Audallion system. For the lowest speech level that we used (50 dB SPL overall), the effective level of the speech dips at high frequencies approached the level of the microphone noise and the quantization noise introduced by the ADC. For the highest level used (80 dB SPL), the short-term peaks in the speech approached the full-scale input level of the ADC.

A reasonable goal of compression systems is to give good audibility for speech over a wide range of input sound levels, while maintaining listening comfort and removing the need for frequent adjustment of a volume control. All of the compression systems studied in our trial achieved this goal. It seems that the exact method by which this goal is achieved is not critical, as long as distortion is low and side effects such as pumping and breathing sounds are minimized.

An additional goal for compression systems is that environmental sounds should be audible, but not uncomfortably loud, in most everyday listening situations. Again, all of our compression systems achieved this goal reasonably effectively. The three systems incorporating the dual front-end AGC system avoided excessive loudness of intense transients by the use of a fast-acting control system, in addition to the slow control system that normally determined the gain. The FULL-4 system achieved protection from intense transients by virtue of being a fast-acting system. The systems did differ somewhat in their responses to low-level background sounds. For example, the DUAL-HI and DUAL-LO

systems did not change their gain during brief pauses in ongoing speech, which meant that low-level background sounds were less obtrusive with these systems than with the FULL-4 system, for which the gain did change during brief pauses in speech. Differences in the intrusiveness of background noise may account for the slight preference for the DUAL-LO system that was found in the field trial conducted after the main experiment.

VI. SUMMARY AND CONCLUSIONS

Four different compression algorithms were implemented in wearable digital hearing aids:

- (1) The slow-acting dual front-end AGC system combined with appropriate frequency-response equalization, with a compression threshold of 63 dB SPL and with a compression ratio of 30 (DUAL-HI);
- (2) The dual front-end AGC combined with appropriate frequency-response equalization, with a compression threshold of 55 dB and with a compression ratio of 3. This was intended to give some impression of the levels of sounds in the environment (DUAL-LO);
- (3) Fast-acting full dynamic range compression in four channels. The compression was designed to minimize envelope distortion due to overshoots and undershoots (FULL-4);
- (4) A combination of (2) and (3) above, where each applied less compression than when used alone (DUAL-4).

In all of the systems, the compression was implemented “cleanly” so as to minimize distortion of various types. Subjects had at least 2 weeks experience with each system in everyday life before evaluation using the APHAB test and measures of speech intelligibility in quiet (AB word lists at 50 and 80 dB SPL) and noise (ASL sentence lists in speech-shaped noise, or that same noise amplitude modulated with the envelope of speech from a single talker). The APHAB scores did not indicate clear differences between the four systems. Scores for the AB words in quiet were high for all four systems at both 50 and 80 dB SPL and did not differ significantly across the systems. The speech-to-noise ratios required for 50% intelligibility were low (indicating good performance) in all conditions, although still somewhat higher than obtained using normally hearing subjects listening unaided. Performance was similar for the four systems, but there was a slight trend for better performance in modulated noise with the DUAL-4 system than with the other systems. A subsequent trial where three subjects directly compared each of the four systems in their everyday lives indicated a slight preference for the DUAL-LO system.

We conclude that markedly different compression systems can be almost equally effective in compensating for the effects of loudness recruitment and reduced dynamic range.

ACKNOWLEDGMENTS

This work was supported by the MRC (UK), the EU (TAP Program, Project SPACE), and by ReSound, AudioLogic, and Danavox. We thank Tom Baer for help in gener-

ating some of the stimuli used in this study. We also thank two anonymous reviewers for helpful comments.

APPENDIX: IMPLEMENTATION OF THE INITIAL FITTING PROCEDURE ON THE DUAL-4 SYSTEM

There were two constraints in determining the initial fitting for the DUAL-4 system:

- (1) The gains for a speech-shaped noise input with an overall level of 65 dB SPL should be as specified by the Cambridge formula. For this purpose, we assumed an eight-band system with nominal center frequencies of 250, 500, 1000, 1500, 2000, 3000, 4000, and 6000 Hz. Call the gains for these center frequencies $G(i)$, where $i = 1$ to 8.
- (2) Speech should be audible over its entire dynamic range in each frequency band when the overall speech level is 45 dB SPL. In each frequency band, the dynamic range of the speech was assumed to extend from 18 dB below the rms level to 12 dB above it.

We started by calculating the gains and compression ratios *as if* we had an eight-channel compression system. Then, these were used to calculate appropriate values for the four-channel system, while preserving the eight-band frequency-response shaping.

A. Parameter derivation for the eight-channel system

The steps for the hypothetical eight-channel system were as follows:

- (1) The absolute thresholds in dB HL, $ABSHL(i)$, were converted to equivalent free-field dB SPL by adding 13, 5, 4, 2, 0, -4, -5, and 4 dB, for the frequencies 0.25, 0.5, 1, 1.5, 2, 3, 4, and 6 kHz, respectively. Call these conversion factors $Conv(i)$.
- (2) For a speech-shaped noise input with an overall (free-field) level of 65 dB SPL, the input levels, $I8(i)$ for each of the eight frequency bands are

$$I8(i) = 61.3, 60.1, 57.6, 49.5, 47.0, 45.1, 39.5 \text{ dB.} \quad (A1)$$

These levels are based on the mean speech spectrum published by Byrne *et al.* (1994). The gains $G(1), \dots, G(8)$ should be achieved for input levels of $I8(1), \dots, I8(8)$. These input levels were changed by the high-frequency emphasis applied prior to the AGC. The amount of emphasis, $EMPH8(i)$, was 0, 0, 3.3, 5.3, 6.7, 8.6, 10, and 10 dB at 0.25, 0.5, 1, 1.5, 2, 3, 4, and 6 kHz, respectively. Since the dual front-end AGC of the DUAL-4 system applied a gain of 22.4 dB for a speech-shaped noise with an input level of 65 dB SPL, the output level from the dual front-end AGC in the i th band was

$$F_{65} = I8(i) + EMPH8(i) + 22.4. \quad (A2)$$

The level required at the output of the whole system was

$$O_{65} = I8(i) + G(i). \quad (A3)$$

- (3) Consider now speech with a mean input level of 45 dB, i.e., all input level $I8(i)$ are reduced by 20 dB. At this

level, which corresponded to the compression threshold for the dual front-end AGC system, the dual front-end AGC applied a gain of 30.6 dB (as the dual front-end AGC had a compression ratio of 1.7). Therefore, the mean level at the output of the dual front-end AGC system for the i th band was

$$F_{45} = (I_8(i) - 20) + \text{EMPH8}(i) + 30.6. \quad (\text{A4})$$

Consider now the case where, when the mean input level was 45 dB, the momentary level was 27 dB (i.e., 18 dB below 45 dB). It was required that the speech level in each band should just reach the absolute threshold. Therefore, the required level in this case, for the i th band, was

$$O_{27} = \text{ABSHL}(i) + \text{Conv}(i). \quad (\text{A5})$$

The dual front-end AGC system was linear for input levels below 45 dB SPL. Therefore, the output level for an input level of 45 dB SPL, O_{45} , was determined, relative to O_{27} , by the action of the appropriate channel compressor

$$O_{45} = O_{27} + (45 - 27)/\text{CR8}(i), \quad (\text{A6})$$

where $\text{CR8}(i)$ is the compression ratio in the i th band. Substituting for O_{27} , we get

$$O_{45} = \text{ABSHL}(i) + \text{Conv}(i) + 18/\text{CR8}(i). \quad (\text{A7})$$

- (4) The compression ratios for the fast compressor are given by

$$\text{CR8}(i) = (F_{65} - F_{45}) / (O_{65} - O_{45}). \quad (\text{A8})$$

Substituting from Eqs. (A2), (A3), (A4), and (A6) gives

$$\begin{aligned} \text{CR8}(i) &= \frac{I_8(i) + \text{EMPH8}(i) + 22.4 - I_8(i) + 20 - \text{EMPH8}(i) - 30.6}{(I_8(i) + G(i) - \text{ABSHL}(i) - \text{Conv}(i) - 18/\text{CR8}(i))} \\ &= 11.8 / (I_8(i) + G(i) - \text{ABSHL}(i) - \text{Conv}(i) - 18/\text{CR8}(i)). \end{aligned} \quad (\text{A9})$$

After rearrangement of terms, this becomes

$$\text{CR8}(i) = 29.7 / (I_8(i) + G(i) - \text{ABSHL}(i) - \text{Conv}(i)). \quad (\text{A10})$$

In practice, a limitation was placed on the amount of low-level gain. If the value of $\text{CR8}(i)$ called for by Eq. (A10) was greater than 2, then its value was set to 2. If Eq. (A10) required a value of $\text{CR8}(i)$ less than 1.001, the value was set to 1.001 [this avoided problems in the adaptive fitting procedure which arose if any value of $\text{CR8}(i)$ was equal to or less than 1].

B. Conversion from the eight-channel to the four-channel system

Conversion to the four-channel system was as follows:

- (1) The compression ratios in each channel, $\text{CR4}(i)$, were calculated from the values of $\text{CR8}(i)$ for the eight-channel system, as follows:

$$\begin{aligned} \text{CR4}(1) &= [\text{CR8}(1) + \text{CR8}(2)]/2, \\ \text{CR4}(2) &= [\text{CR8}(3) + \text{CR8}(4)]/2, \\ \text{CR4}(3) &= [\text{CR8}(5) + \text{CR8}(6)]/2, \\ \text{CR4}(4) &= [\text{CR8}(7) + \text{CR8}(8)]/2. \end{aligned}$$

- (2) The compression thresholds for the four channels were calculated as follows: It was desired to apply the appropriate amplification for input levels down to $I_4(i) - 38$ dB SPL, where $I_4(i)$ is the input level in the i th channel for a speech-shaped noise with input level 65 dB SPL, i.e., 64.0, 57.3, 48.8, and 47.8 dB, for $i = 1$ to 4. Therefore, the compression thresholds corresponded to input levels per band of 26, 19.3, 10.8, and 9.8 dB SPL, respectively.
- (3) The gains that needed to be applied subsequent to the compression to achieve the correct overall frequency-response shape were calculated such that, when the input to the dual front-end AGC was speech-shaped noise at 65 dB SPL, the overall gains at the center frequencies defined for the eight-channel system corresponded to $G(i)$. These gains were interpolated and extrapolated to 32 values, corresponding to the bins in the FFT. For frequencies below 0.25 kHz, the gains were set to the value for 0.25 kHz. For values above 6.0 kHz, the gains were set to the value for 6.0 kHz. Linear interpolation of dB on a log-frequency scale was used for intermediate frequencies.

- Boothroyd, A. (1968). "Developments in speech audiometry," *Sound* **2**, 3-10.
- Braida, L. D., Durlach, N. I., De Gennaro, S. V., Peterson, P. M., and Bustamante, D. K. (1982). "Review of recent research on multiband amplitude compression for the hearing impaired," in *The Vanderbilt Hearing-Aid Report*, edited by G. A. Studebaker and F. H. Bess (Monographs in Contemporary Audiology, Upper Darby, PA).
- Bustamante, D. K., and Braida, L. D. (1987). "Multiband compression limiting for hearing-impaired listeners," *J. Rehabil. Res. Dev.* **24**, 149-160.
- Byrne, D. et al. (1994). "An international comparison of long-term average speech spectra," *J. Acoust. Soc. Am.* **96**, 2108-2120.
- Carhart, R. C., and Tillman, T. W. (1970). "Interaction of competing speech signals with hearing losses," *Arch. Otolaryngol.* **91**, 273-279.
- Cox, R. M., and Alexander, G. C. (1995). "The abbreviated profile of hearing aid benefit," *Ear Hear.* **16**, 176-186.
- Dillon, H. (1996). "Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times?," *Ear Hear.* **17**, 287-307.
- Duquesnoy, A. J. (1983). "Effect of a single interfering noise or speech source on the binaural sentence intelligibility of aged persons," *J. Acoust. Soc. Am.* **74**, 739-743.
- Eisenberg, L. S., Dirks, D. D., and Bell, T. S. (1995). "Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing," *J. Speech Hear. Res.* **38**, 222-233.
- Finney, D. J. (1971). *Probit Analysis* (Cambridge University Press, Cambridge).
- Goedegebure, A., Hulshof, M., Mass, A. J. J., and Verschuure, J. (1996). "The effects of syllabic compression on speech intelligibility in hearing impaired," in *Psychoacoustics, Speech and Hearing Aids*, edited by B. Kollmeier (World Scientific, Singapore).
- Hickson, L. M. H. (1994). "Compression amplification in hearing aids," *Am. J. Audiol.* **3**, 51-65.
- Hohmann, V., and Kollmeier, B. (1995). "The effect of multichannel dynamic compression on speech intelligibility," *J. Acoust. Soc. Am.* **97**, 1191-1195.
- Hood, J. D. (1972). "Fundamentals of identification of sensorineural hearing loss," *Sound* **6**, 21-26.
- Houtgast, T., Steeneken, H. J. M., and Plomp, R. (1980). "Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acoustics," *Acustica* **46**, 60-72.

- Hygge, S., Rönnerberg, J., Larsby, B., and Arlinger, S. (1992). "Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds," *J. Speech Hear. Res.* **35**, 208–215.
- Kollmeier, B., and Hohmann, V. (1995). "Loudness estimation and compensation employing a categorical scale," in *Advances in Hearing Research*, edited by G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore).
- Lippmann, R. P., Braida, L. D., and Durlach, N. I. (1981). "Study of multi-channel amplitude compression and linear amplification for persons with sensorineural hearing loss," *J. Acoust. Soc. Am.* **69**, 524–534.
- MacLeod, A., and Summerfield, Q. (1990). "A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use," *Br. J. Audiol.* **24**, 29–43.
- Moore, B. C. J. (1993). "Signal processing to compensate for reduced dynamic range," in *Recent Developments in Hearing Instrument Technology*, edited by J. Beilin and G. R. Jensen (Stougaard Jensen, Copenhagen).
- Moore, B. C. J. (1998). *Cochlear Hearing Loss* (Whurr, London).
- Moore, B. C. J., Alcántara, J. I., and Glasberg, B. R. (1998). "Development and evaluation of a procedure for fitting multi-channel compression hearing aids," *Br. J. Audiol.* **32**, 177–195.
- Moore, B. C. J., Alcántara, J. I., Stone, M. A., and Glasberg, B. R. (1999a). "Use of a loudness model for hearing aid fitting II. Hearing aids with multi-channel compression," *Br. J. Audiol.* **33**, 99–113.
- Moore, B. C. J., and Glasberg, B. R. (1988). "A comparison of four methods of implementing automatic gain control (AGC) in hearing aids," *Br. J. Audiol.* **22**, 93–104.
- Moore, B. C. J., and Glasberg, B. R. (1997). "A model of loudness perception applied to cochlear hearing loss," *Aud. Neurosci.* **3**, 289–311.
- Moore, B. C. J., and Glasberg, B. R. (1998). "Use of a loudness model for hearing aid fitting. I. Linear hearing aids," *Br. J. Audiol.* **32**, 301–319.
- Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (1991). "Optimization of a slow-acting automatic gain control system for use in hearing aids," *Br. J. Audiol.* **25**, 171–182.
- Moore, B. C. J., Glasberg, B. R., and Vickers, D. A. (1995). "Simulation of the effects of loudness recruitment on the intelligibility of speech in noise," *Br. J. Audiol.* **29**, 131–143.
- Moore, B. C. J., Johnson, J. S., Clark, T. M., and Pluvinaige, V. (1992a). "Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss," *Ear Hear.* **13**, 349–370.
- Moore, B. C. J., Lynch, C., and Stone, M. A. (1992b). "Effects of the fitting parameters of a two-channel compression system on the intelligibility of speech in quiet and in noise," *Br. J. Audiol.* **26**, 369–379.
- Moore, B. C. J., Peters, R. W., and Stone, M. A. (1999b). "Benefits of linear amplification and multi-channel compression for speech comprehension in backgrounds with spectral and temporal dips," *J. Acoust. Soc. Am.* **105**, 400–411.
- Moore, B. C. J., Wojtczak, M., and Vickers, D. A. (1996). "Effect of loudness recruitment on the perception of amplitude modulation," *J. Acoust. Soc. Am.* **100**, 481–489.
- Pearsons, K. S., Bennett, R. L., and Fidell, S. (1976). *Speech Levels in Various Environments. Report No. 3281* (Bolt, Beranek and Newman, Cambridge, MA).
- Peters, R. W., Moore, B. C. J., and Baer, T. (1998). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," *J. Acoust. Soc. Am.* **103**, 577–587.
- Plomp, R. (1983). "The role of modulation in hearing," in *Hearing—Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartmann (Springer, Berlin).
- Plomp, R. (1988). "The negative effect of amplitude compression in multi-channel hearing aids in the light of the modulation-transfer function," *J. Acoust. Soc. Am.* **83**, 2322–2327.
- Plomp, R. (1994). "Noise, amplification, and compression: Considerations of three main issues in hearing aid design," *Ear Hear.* **15**, 2–12.
- Steinberg, J. C., and Gardner, M. B. (1937). "The dependency of hearing impairment on sound intensity," *J. Acoust. Soc. Am.* **9**, 11–23.
- Stone, M. A., and Moore, B. C. J. (1992). "Syllabic compression: Effective compression ratios for signals modulated at different rates," *Br. J. Audiol.* **26**, 351–361.
- Stone, M. A., Moore, B. C. J., Wojtczak, M., and Gudgin, E. (1997). "Effects of fast-acting high-frequency compression on the intelligibility of speech in steady and fluctuating background sounds," *Br. J. Audiol.* **31**, 257–273.
- Takahashi, G. A., and Bacon, S. P. (1992). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," *J. Speech Hear. Res.* **35**, 1410–1421.
- Verschuure, J., and Dreschler, W. A. (1996). "Dynamic compression in hearing aids," in *Psychoacoustics, Speech and Hearing Aids*, edited by B. Kollmeier (World Scientific, Singapore).
- Verschuure, J., Maas, A. J. J., Stikvoort, E., de Jong, R. M., Goedegebure, A., and Dreschler, W. A. (1996). "Compression and its effect on the speech signal," *Ear Hear.* **17**, 162–175.
- Villchur, E. (1973). "Signal processing to improve speech intelligibility in perceptive deafness," *J. Acoust. Soc. Am.* **53**, 1646–1657.
- Yund, E. W., and Buckles, K. M. (1995). "Enhanced speech perception at low signal-to-noise ratios with multichannel compression hearing aids," *J. Acoust. Soc. Am.* **97**, 1224–1240.

Determination of parameters for lumped parameter models of the vocal folds using a finite-element method approach

M. P. de Vries, H. K. Schutte, and G. J. Verkerke

Division of Artificial Organs, University of Groningen, Bloemensingel 10, 9712 KZ Groningen, The Netherlands

(Received 15 April 1999; revised 29 July 1999; accepted 19 August 1999)

To study the mechanical behavior of the vocal folds, lumped parameter models of the vocal folds have been developed in the past. Coupling with a model of the aerodynamics in the glottis provides the possibility of simulating glottal waves. A new method is presented to obtain values for the masses and springs of the lumped parameter models by using a finite-element method model of the vocal folds. This finite-element method model is based on geometry and material data from the literature, resulting in a model that describes the vocal fold in a realistic way. Requiring the dynamic behavior of the lumped parameter model of the vocal fold to be equal to the dynamic behavior of the finite-element method model of the vocal fold, parameter values are obtained that are purely based on dynamic considerations. Therefore, the behavior of the vocal folds is described in a realistic way by these parameters. These values are compared with the values used by previous authors and are implemented in lumped parameter models. Self-sustained oscillation is achieved with the new values for masses and springs. © 1999 Acoustical Society of America.

[S0001-4966(99)03412-8]

PACS numbers: 43.70.Bk [AL]

LIST OF SYMBOLS

m_1	lower mass	G_{cover}	longitudinal shear modulus of the cover
m_2	upper mass	G_{body}	longitudinal shear modulus of the body
k_1	lower spring stiffness	ν	Poisson's ratio
k_2	upper spring stiffness	M	mass matrix of two-mass model
k_c	coupling spring stiffness	R	damping matrix of two-mass model
k_{col1}	lower collision spring	K	stiffness matrix of two-mass model
k_{col2}	upper collision spring	k_{11}	element 1,1 of stiffness matrix
r_1	lower damper	k_{12}	element 1,2 and 2,1 of stiffness matrix
r_2	upper damper	k_{22}	element 2,2 of stiffness matrix
ζ_1	lower damping ratio	f_1	frequency of normal mode 1
ζ_2	upper damping ratio	f_2	frequency of normal mode 2
x_1	deflection of m_1	F_0	fundamental frequency
x_2	deflection of m_2	FEM_{mech}	set of lumped parameters determined using the FEM
F_1	force on m_1	$\text{I\&F}_{\text{mech}}$	set of lumped parameters in the Ishizaka and Flanagan model
F_2	force on m_2	Pel_{mech}	set of lumped parameters in the Pelorson model
d_1	thickness lower part	CQ	closed quotient
d_2	thickness upper part	U_g	glottal peak flow
L_1	location on the FEM model representing m_1	u_1/u_2	ratio of deflections of m_1 and m_2 for f_1
L_2	location on the FEM model representing m_2		
E_{cover}	Young's modulus of cover tissue of vocal fold		
E_{body}	Young's modulus of body tissue of vocal fold (in longitudinal direction)		

INTRODUCTION

Numerical modeling of the vocal folds appeared to be very useful to study the processes involved in normal voice production. Over the years, several investigators have developed a number of different models. To study the dynamic interaction between vocal folds and aerodynamics, the so-called lumped parameter type of numerical models of the vocal folds are suitable. These models consist of a model of the aerodynamics in the glottis combined with a parametric description of the vocal folds. They combine a high degree

of reality in their results with a considerable simplicity. Several lumped parameter models have been developed, e.g., the 2-mass models of Ishizaka and Flanagan (1972), Koizumi *et al.* (1987), Pelorson *et al.* (1994), Lous *et al.* (1998); the 3-mass model of Story and Titze (1995); the 16-mass model of Titze (1973, 1974). These models differ from each other by the number of parameters describing the vocal folds, by the value of the parameters, or by the way aerodynamic influences are implemented. To avoid confusion, we make a distinction between the mechanical part of the lumped pa-

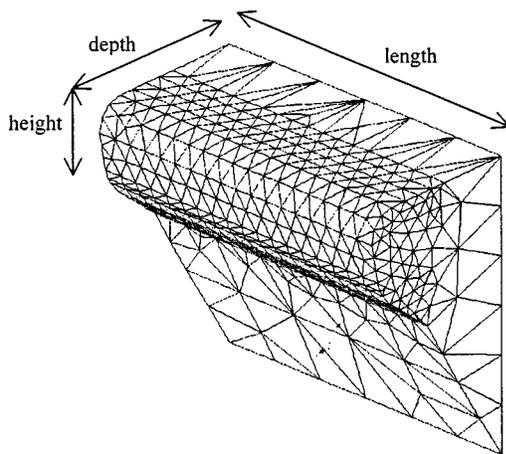


FIG. 1. Geometry of the vocal fold in the FEM model.

parameter models and the aerodynamic part that described the airflow through the glottis. In this study, we focus on the mechanical part of the lumped parameter models.

When considering the two-mass models that are used frequently, it is difficult to determine realistic values for the lumped parameters. For example, most of the authors presenting two-mass models adopt parameter values of the mechanical part that are identical to the values proposed by Ishizaka and Flanagan (Herzel *et al.*, 1995; Steinecke and Herzel, 1995). These values are partially chosen on mechanical considerations, but also result from fitting glottal waves produced by the two-mass model including aerodynamics with estimations of glottal waves produced in normal phonation.

The aim of this paper is to present realistic values for the mass and spring parameters of lumped parameter models of the vocal folds. By using the finite-element method (FEM), a detailed description of the geometry and the material properties of the vocal folds is possible. By requiring equal mechanical behavior of the FEM model and of a lumped parameter model, realistic mass and spring parameter values of the lumped parameter model can be determined.

MATERIALS

Finite-element method model of the vocal folds

The finite-element method (FEM) gives the possibility of implementing detailed geometric and material information in a model. In this way, a FEM model can be used to describe the geometric and material properties of the vocal folds in a realistic way. Other authors also used the FEM to study the physiology of phonation (e.g., Titze *et al.*, 1992; Berry *et al.*, 1994).

In our study, anatomical data from Myerson (1964) and Baer (1981) have been used to implement the geometry of a vocal fold in the FEM model (Fig. 1). Myerson cited Schaeffer, who determined the length of the membranous portion of the vocal fold to be 11.5 for female and 15.5 mm for male. Baer considered the membranous part of the vocal fold to be 13 mm. Combining the data of Myerson and Baer, we assume the average length of the membranous portion of the vocal fold to be 13 mm. This is the length of the vocal fold

that we implemented in the FEM model. After Hirano (1974), an improved understanding of the behavior of the vocal folds can be reached when the vocal folds are considered to be built up by a body and a cover. The differences in the behavior of the body and cover are expressed in different material properties. For our FEM approach we used the following geometric and material properties. The total thickness (vertical plane) of the vocal fold is 3 mm. The cover has a maximum thickness of 0.5 mm. The depth of the vocal fold is the part that plays a role in vibration and is assumed to be 4 mm. The length (horizontal plane, from front to back) is assumed to be 13 mm. The material properties implemented in the FEM model are density and elasticity of the cover and of the body of the vocal fold. The densities of the body and the cover have been mentioned in the literature by several authors (e.g., Titze and Strong, 1975; Story and Titze, 1995; Berry and Titze, 1996). The differences between the values they use are small. We follow Berry using a value of 1.03 g/cm^3 for the density of the whole vocal fold. The elasticity of the vocal fold is defined by the Young's modulus and shear modulus of vocal fold tissue. In the present study, Berry's data for the transversely isotropic elasticity modulus of the body and the cover of the vocal fold are used. Berry used a transverse Young's modulus of the body E_{body} of 4 kPa, a transverse Young's modulus of the cover E_{cover} of 2 kPa, a longitudinal shear modulus of the body G_{body} of 12 kPa, and a longitudinal shear modulus of the cover G_{cover} of 10 kPa. To complete the set of material properties, Poisson's ratio ν is set to 0.4 according to Berry and Titze (1996). No longitudinal tension is applied, in contrast with Titze and Strong (1975), but in accordance with Berry and Titze (1994, 1996).

Both geometric and material properties of the vocal fold are implemented in the FEM model. Assuming left-right symmetry, one vocal fold is considered. The vocal fold is divided into approximately 3000 elements. Figure 1 shows the resulting mesh. The anterior, lateral, and posterior sides are constrained (immobile), assuming the cartilage to be very stiff compared to the vocal fold tissue.

Lumped parameter model

For the lumped parameter model, the two-mass model of the vocal folds (Fig. 2) was chosen. The concept of dividing the vocal fold in an upper part and a lower part has proven to give insight into the physiology of phonation (Ishizaka and Flanagan, 1972; Titze, 1974; Herzel *et al.*, 1994; Stevens, 1988; Pelorson *et al.*, 1994).

The choice to use the two-mass model in this study is based on the wide use of it, combined with its simplicity. The numerical method presented in this study can be applied to all existing lumped parameter models of the vocal folds.

The two-mass model describes one vocal fold by two coupled oscillators [Fig. 2(a)]. Each oscillator consists of a mass, a spring stiffness, and a damper. Mass m_1 , spring stiffness k_1 , and damper r_1 represent the lower part with thickness d_1 of the vocal fold. Mass m_2 , spring stiffness k_2 , and damper r_2 represent the upper part with thickness d_2 . The two masses are coupled by a spring stiffness k_c . The two masses, m_1 and m_2 , are permitted to move in a lateral direction. The deflections of m_1 and m_2 are x_1 and x_2 , re-

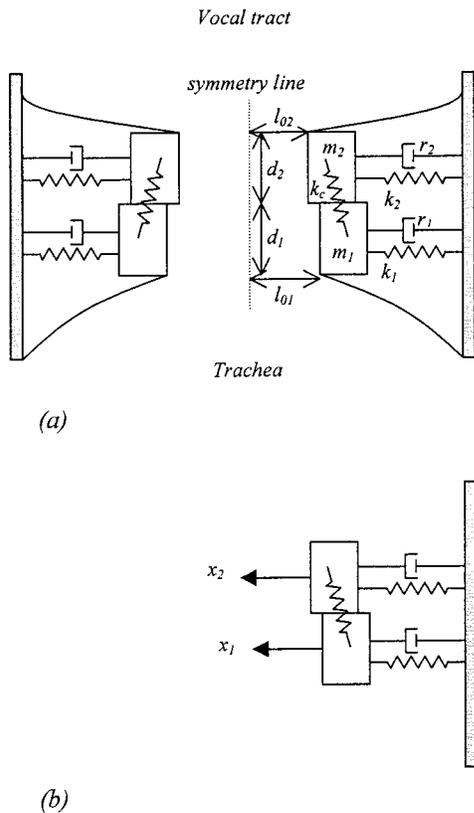


FIG. 2. (a) Two-mass model; (b) x_1 and x_2 are the deflections of the masses m_1 and m_2 .

spectively [Fig. 2(b)]. In the two-mass model, symmetry along the length of the glottis is assumed, therefore only one vocal fold is considered. When the vocal fold approaches the symmetry line within a very short distance, collision springs with stiffness k_{col1} and k_{col2} will be activated and have an influence on m_1 and m_2 , respectively, in the contralateral direction.

Implementation of new parameter values in existing lumped parameter models

This study results in a new set of mass and spring parameters. For the lumped parameter models, the Ishizaka and Flanagan model and the Pelorson model have been chosen. The new set of mass and spring parameters is implemented in these two lumped parameter models. The values of the damping parameters follow from the mass and spring parameters according to $r_i = 2\zeta\sqrt{m_i k_i}$, $i = 1, 2$. Two different lumped parameter models have been chosen because we want to demonstrate that the new values lead to self-sustained oscillation with the use of different descriptions of the aerodynamics (the Ishizaka and Flanagan model and the Pelorson model) in the glottis. In this way, the behavior of the vocal fold described by the new set of mass and spring parameters can be examined under different aerodynamic influences. Both models are shown in Fig. 3.

The relevant differences between the models are as follows:

- (1) Separation of the flow from the vocal fold occurs at the

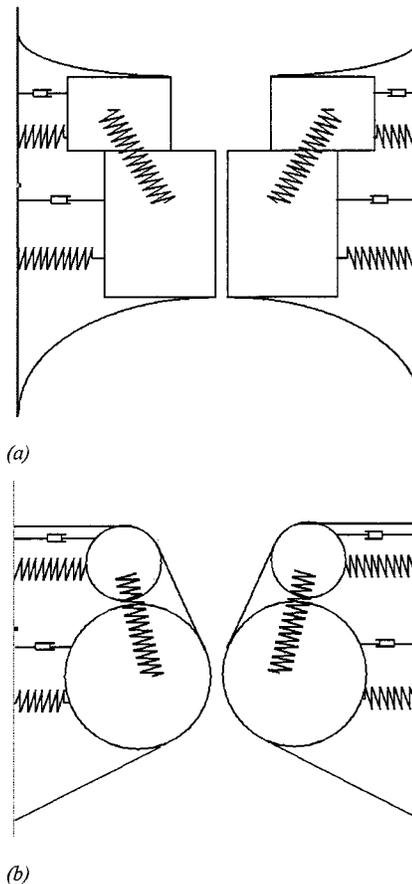


FIG. 3. Comparison of two lumped parameter models: (a) Ishizaka and Flanagan; (b) Pelorson.

entrance of the glottis (Ishizaka and Flanagan) or at a location depending on convergence or divergence of the glottis (Pelorson).

- (2) Pressure recovery is present in the Ishizaka and Flanagan model and absent in the Pelorson model.
- (3) Springs have a nonlinear behavior in the Ishizaka and Flanagan model and are assumed to behave linearly in the Pelorson model.
- (4) Closure of the glottis occurs abruptly in the Ishizaka and Flanagan model and gradually in the Pelorson model.

METHODS

First, a general description will be given of the steps necessary to determine the values of the mass and spring parameters for the mechanical part of the two-mass model. The steps will be explained in more detail in following sections.

General description

The basis of the numerical tool is the requirement that the dynamic response of the lumped parameter model (in our study the two-mass model) to aerodynamic forces must be equal to the dynamic response of the FEM model to equivalent aerodynamic forces. The lateral deflections x_1 and x_2 of the two masses determine the dynamic response of the mechanical part of the two-mass model. In the FEM model it is necessary to choose two locations which represent the loca-

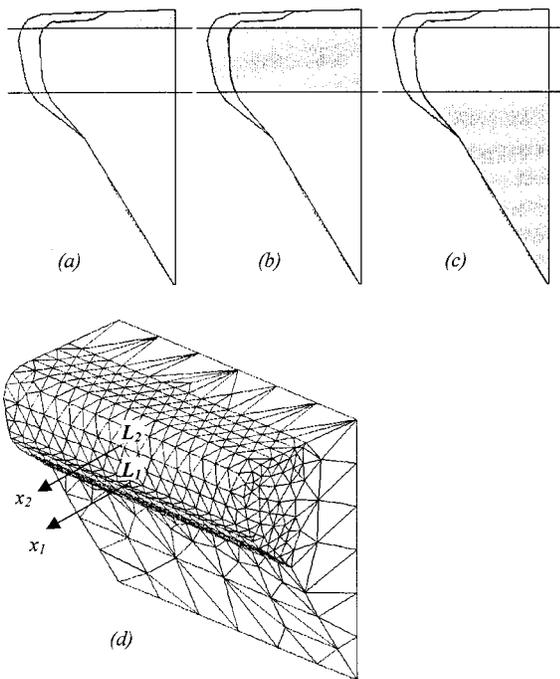


FIG. 4. Cross section of FEM model; (a) division in upper part; (b) mid part; (c) lower part; (d) locations L_1 and L_2 .

tion of the masses of the lumped parameter model. The lateral deflections of these locations on the FEM model must correspond to the lateral deflection of the lumped parameter model. Because in the lumped parameter models the locations are situated on the glottal surface, the two locations on the FEM model are chosen at the glottal surface too. Therefore, the two locations L_1 and L_2 are defined in the mid-coronal plane of the FEM model (center of the glottis) at 0.5 mm below the superior surface (L_1) and at 2.5 mm below the superior surface [L_2 , see Fig. 4(d)].

By equaling the dynamic response of the mechanical part of the two-mass model (determined by x_1 and x_2) to the dynamic response of the FEM model (determined by the deflections of L_1 and L_2), the numerical values of the lumped parameters can be derived. Two steps are used to determine these values:

- (1) First, the stiffness of the FEM model is examined. The deflections of L_1 and L_2 as a result of a pressure distribution in the glottis are calculated. By requiring the two-mass model to behave as stiffly as the FEM model, the numerical values of the spring parameters of the two-mass model result. Because the two-mass model has three spring parameters (k_1 , k_2 , and k_c), three different situations have to be analyzed to determine the three unknowns.
- (2) The normal modes of the FEM model are calculated and the normal modes of the two-mass model are expressed in terms of masses and springs. In the lowest eigenmode of the two-mass model, all tissue moves in phase. The eigenmode of the FEM model in which all tissue moves in phase is determined. Now the masses are determined by equaling the normal mode of the two-mass model with the corresponding normal mode in the FEM model.

Equations of motion

The dynamic response of the mechanical part of the two-mass model can be described by the equations of motion of the two masses. These equations are

$$m_1 \ddot{x}_1 + r_1 \dot{x}_1 + (k_1 + k_c)x_1 - k_c x_2 = F_1, \quad (1)$$

$$m_2 \ddot{x}_2 + r_2 \dot{x}_2 + (k_2 + k_c)x_2 - k_c x_1 = F_2, \quad (2)$$

where F_1 and F_2 are the forces exerted on the masses. A dot above a variable means a time derivative of the variable, a double dot above a variable means the second time derivative of the variable. These equations are used in the steps to determine the values for mass and spring parameters.

Determination of spring parameters of the two-mass model

The spring parameters of the two-mass model are determined by equaling the stiffness of the two-mass model with the stiffness of the FEM model. The spring stiffness k_1 is equaled with the stiffness of the tissue between L_1 and the cartilage. The stiffness of spring stiffness k_2 is equaled with the stiffness of the tissue between L_2 and the cartilage. The stiffness of spring stiffness k_c is equaled with the stiffness of the tissue between L_1 and L_2 . As mentioned before, three situations have to be analyzed. Three situations are created by assuming constraints on three different parts of the vocal fold. The three parts into which the FEM model is divided are indicated in Fig. 4: an upper part consisting of all elements situated above L_2 [Fig. 4(a)], a lower part consisting of all elements located below L_1 [Fig. 4(c)], and a mid part consisting of all elements located between the upper and lower part [Fig. 4(b)].

On the mechanical part of the two-mass model, a load F_1 is applied on m_1 and a load F_2 is applied on m_2 . In the FEM model, a pressure is applied at the glottal surface. The pressure is chosen in such a way that the resulting lateral forces in the two-mass model and the resulting lateral force in the FEM model are equal. For the determination of the spring parameters, static analyses are performed. Therefore, the time derivatives in Eqs. (1) and (2) vanish. The reduced equations of motion are

$$(k_1 + k_c)x_1 - k_c x_2 = F_1, \quad (3)$$

$$(k_2 + k_c)x_2 - k_c x_1 = F_2. \quad (4)$$

Three situations are considered:

- (1) In the two-mass model, m_2 is constrained. By this constraint, Eq. (4) is canceled and Eq. (3) reduces to

$$(k_1 + k_c) \cdot x_1 = F_1. \quad (5)$$

In the FEM model, the upper part [Fig. 4(a)] is constrained. In this way, the influence of the tissue between L_2 and the cartilage is eliminated. A static calculation with the FEM model results in lateral deflections of L_1 . This deflection is entered in Eq. (5) as x_1 .

- (2) In the two-mass model, m_1 is constrained. By this constraint, Eq. (3) is canceled and Eq. (4) reduces to

$$(k_2+k_c)\cdot x_2=F_2. \quad (6)$$

In the FEM model, the lower part [Fig. 4(c)] is constrained. In this way, the influence of the tissue between L_1 and the cartilage is eliminated. A static calculation with the FEM model results in lateral deflections of L_2 . This deflection is entered in Eq. (6) as x_2 .

- (3) To obtain the third equation, no constraints are applied. Equations (3) and (4) are summed, resulting in

$$k_1x_1+k_2x_2=F_1+F_2. \quad (7)$$

Equations (5)–(7) are combined, resulting in numerical values for k_1 , k_2 , and k_c .

Determination of mass parameters

The magnitude of the masses can be extracted when the normal modes are calculated using the equations of motion of the two-mass model. To obtain the normal modes, the equations of motion are written in matrix notation,

$$[M]\{\ddot{x}\}+[R]\{\dot{x}\}+[K]\{x\}=\{F\}, \quad (8)$$

where M is the mass matrix

$$M=\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix},$$

R is the damping matrix,

$$R=\begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix},$$

and K is the stiffness matrix

$$K=\begin{bmatrix} k_1+k_c & -k_c \\ -k_c & k_2+k_c \end{bmatrix}.$$

With these equations of motion the frequencies of the normal modes can be determined by

$$\det[K-(2\pi f)^2M], \quad (9)$$

where f_1 and f_2 represent the two eigenvalues of the system. Extracting the determinant of the matrix, the two frequencies f_1 and f_2 result from the eigenvalues

$$\begin{aligned} \frac{(f_1 \cdot 2\pi)^2}{(f_2 \cdot 2\pi)^2} &= \frac{1}{2} \frac{m_1k_{22}+m_2k_{11}}{m_1m_2} \\ &\times \mu \frac{1}{2} \sqrt{\left(\frac{m_1k_{22}+m_2k_{11}}{m_1m_2}\right)^2 - 4 \frac{k_{11}k_{22}-k_{12}^2}{m_1m_2}}, \end{aligned} \quad (10)$$

where the numbers k_{ij} are the i,j th element of the stiffness matrix,

$$k_{11}=k_1+k_c, \quad k_{22}=k_2+k_c, \quad k_{12}=k_{21}=-k_c.$$

To be able to determine the two values of the masses, two equations are needed. From the normal modes, four possible equations are available: two frequencies available, each resulting in one equation, and both frequencies have a corre-

sponding mode shape, each resulting in an equation also. So a choice has to be made. The criterion we choose is: the best fit of the normal mode that corresponds to f_1 is the dominant normal mode. This corresponds to the finding of Berry *et al.* (1994) that 72.5% of the vocal fold vibration consists of this dominant normal mode. We want to be able to describe this normal mode in the best possible way.

When f_1 is known, the mode shape of the vocal fold corresponding to f_1 is used to determine the masses,

$$\frac{u_2}{u_1} = -\frac{k_{11}-(2\pi f_1)^2 m_1}{k_{12}} = -\frac{k_{12}}{k_{22}-(2\pi f_1)^2 m_2},$$

where u_2/u_1 is the ratio of the deflections of L_1 and L_2 for the mode shape of f_1 . From these equations, m_1 and m_2 result.

Comparison between results of the numerical method and results of existing two-mass models

The newly obtained values for mass and spring parameters are compared with the values of the mass and spring parameters of the Ishizaka and Flanagan model and of the Pelorson model. To investigate the possibility of self-sustained oscillation of the vocal fold represented by the newly obtained values, the values will be implemented in the Ishizaka and Flanagan and the Pelorson models of aerodynamics. For comparison, the original mass and spring parameters of Ishizaka and Flanagan and of Pelorson are implemented in both aerodynamic models also.

So three sets of lumped parameters have been used: the new set of mass and spring parameter values determined using the FEM (FEM_{mech}), the set of parameters from the Ishizaka and Flanagan model (I&F_{mech}), and the set of parameters from the Pelorson model (Pel_{mech}). When the three sets of lumped parameters (the FEM_{mech} set, the I&F_{mech} set, and the Pel_{mech} set) are implemented in the two two-mass models (the Ishizaka and Flanagan model and the Pelorson model), six series of glottal waves result. To obtain glottal waves, a subglottal pressure was applied. Schutte (1980) determined the mean pressure for male and female for normal phonation to be 0.44 kPa, which is 4.3 cm H₂O. Holmberg *et al.* (1989) determined the mean pressure for female and male for normal phonation to be 6.0 cm H₂O. No acceptable glottal waves could be obtained using the mean subglottal pressure of Schutte of 4.3 cm H₂O with the I&F_{mech} set in the Pelorson model and also not with the Pel_{mech} set in the Ishizaka and Flanagan model; therefore the mean pressure of Holmberg of 6.0 cm H₂O has been used. This pressure is applied in the Ishizaka and Flanagan model and in the Pelorson model, causing the two-mass system to vibrate. The glottal waves produced with the different sets of lumped parameters are analyzed. The glottal waves produced will be compared by the fundamental frequency F_0 , the glottal wave amplitude U_g , and the closed quotient CQ .

Holmberg determined mean values of F_0 , U_g , and CQ for 25 male and 20 female subjects. We will use these mean values for comparison of our numerical results with normal phonation.

TABLE I. Parameter values of the I&F72 model (I&F_{mech}), the Pelorson model (Pel_{mech}), and the parameter values determined using the FEM (FEM_{mech}).

	I&F _{mech}	Pel _{mech}	FEM _{mech}
m_1 lower mass (g)	0.125	0.170	0.024
m_2 upper mass (g)	0.025	0.030	0.020
k_1 lower spring stiffness (N/m)	80	45	22
k_2 upper spring stiffness (N/m)	8	8	14
$\eta_{1,2}$ nonlinear coefficient	100	N.A.	10 in I&F _{mech} N.A. in Pel _{mech}
k_c coupling spring stiffness (N/m)	25	25	10
ζ_1 damping ratio (g/s)		0.1	
ζ_2 damping ratio (g/s)		0.6	
d_1 lower thickness (cm)	0.25	N.A.	0.15 in I&F _{mech} N.A. in Pel _{mech}
d_2 upper thickness (cm)	0.05	N.A.	0.15 in I&F _{mech} N.A. in Pel _{mech}
l_g glottal length (cm)	1	1.4	1.3
l_{01} initial position (cm)		0.018	
l_{02} initial position (cm)		0.0	
k_{col1} collision spring stiffness (N/m)		$3k_1$	
k_{col2} collision spring stiffness (N/m)		$3k_2$	

RESULTS

The values for the lumped parameters of the FEM_{mech} set are listed in Table I, together with the values of the lumped parameters of the I&F_{mech} set and of the Pel_{mech} set. Note the difference between the value of m_1 in the FEM_{mech} set on the one hand and the value of m_1 in the I&F_{mech} set and the Pel_{mech} set on the other hand: the newly obtained value for m_1 is five to seven times smaller than m_1 in I&F_{mech} or Pel_{mech}. Another difference between the FEM_{mech} on the one hand and I&F_{mech} and Pel_{mech} on the other hand is the fact that the two masses in the FEM_{mech} set are more equal to each other than the two masses in the other two lumped parameter sets. A third difference concerning the masses is the fact that the sum of m_1 and m_2 in the FEM_{mech} set is 0.044 instead of 0.15 g in the I&F_{mech} set and 0.20 g in the Pel_{mech} set. This means that in the FEM_{mech} set, the sum of the two masses is 18% of the total mass of the tissue modeled in the FEM model (which is 0.250 g); in the

I&F_{mech} set, the sum of the two masses is 60% of the total mass of the tissue in the FEM model; in the Pel_{mech} set, the sum of the two masses is 80% of the total mass of the tissue modeled in the FEM model. In the FEM_{mech} set, not only different values for the masses are proposed, but also different values for the springs. An interesting difference concerning the values of the springs between the FEM_{mech} set and the two other sets is the spring stiffness k_2 , which is closer to k_1 in the FEM_{mech} set than in the other parameter sets.

The sets of lumped parameters FEM_{mech}, I&F_{mech}, and Pel_{mech} have been implemented in the Ishizaka and Flanagan model and in the Pelorson model. Self-sustained oscillation was obtained in all six situations. The resulting glottal waves have been analyzed. The fundamental frequency of the glottal waves F_0 for the different configurations is shown in Fig. 5. Also, the average F_0 for females and males according to Holmberg is indicated. Both the FEM_{mech} set and the I&F_{mech} set result in glottal waves with an F_0 that is in between the female and male average. The Pel_{mech} set results in an F_0 that falls slightly below normal ranges.

U_g is shown in Fig. 6. Large differences between the results using the Ishizaka and Flanagan model and the Pelorson model have been established, especially for the Pel_{mech} set. The U_g produced with the FEM_{mech} set are in between the results of the I&F_{mech} set and the Pel_{mech} set. The U_g is outside normal ranges using the FEM_{mech} set and the Pel_{mech} set in the Ishizaka and Flanagan model; the Pelorson model gives a U_g that is outside the physiologic ranges for all the three mechanical sets.

Figure 7 shows the CQ of the glottal waves. With the FEM_{mech} set, a CQ of an average male is reached with the Pelorson model, and a CQ between male and female is reached with the Ishizaka and Flanagan model. I&F_{mech} results in an average CQ that is in the average female range. Using Pel_{mech}, a CQ slightly below the female range has been reached.

DISCUSSION

The values of the newly determined lumped parameters differ from the values of the two two-mass models considered in this study. The sum of the two masses in the FEM_{mech}

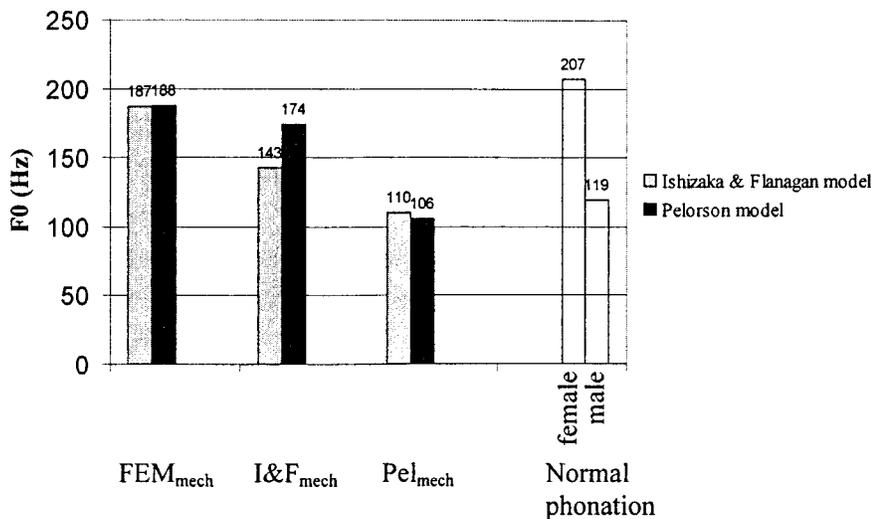


FIG. 5. Fundamental frequency of glottal waves with different aerodynamics and mechanics.

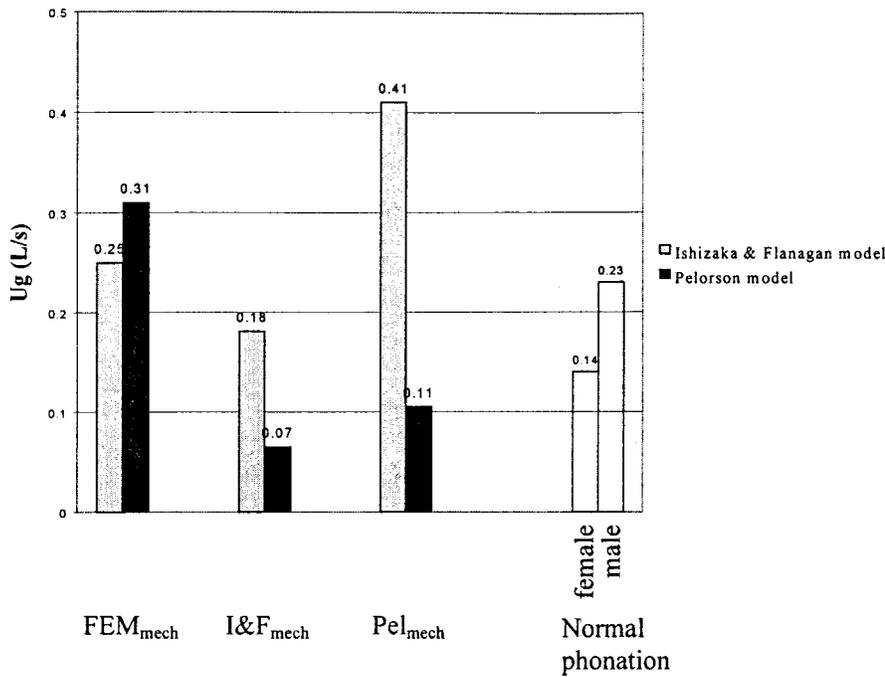


FIG. 6. Glottal peak flow with different aerodynamics and mechanics.

set is a factor of four to five smaller than the sum of the two masses in the I&F_{mech} set and in the Pel_{mech} set. This difference could be explained by the fact that in the I&F and Pelorson model the total mass of the vocal fold (volume \times density) is considered. We postulate that in lumped parameter models, the total mass of the vocal fold is not a relevant measure, but the effective mass that plays a role in vibration. Considering our results, we can conclude that in the I&F_{mech} set and the Pel_{mech} set the values of masses and springs might have been chosen too large. This is in correspondence with Lous *et al.* (1998), where he stated that in the lumped parameter models, the aerodynamic forces are probably overestimated. This overestimation of the aerodynamic forces needs to be compensated by an overestimation of the masses and springs of the vocal fold to prevent the glottal waves from being unrealistic. This can also be confirmed by the fact that for

the I&F_{mech} set and the Pel_{mech} set, no glottal waves could be obtained in the Pelorson model and the Ishizaka and Flanagan model, respectively, when the average subglottal pressure of 4.4 cm H₂O of Schutte (1980) has been applied, while the FEM_{mech} set gives glottal waves with this pressure in both the Ishizaka and Flanagan model and the Pelorson model. Considering our FEM_{mech} set with smaller values for masses and springs, reflecting the use of the parameter values of the I&F_{mech} set and the Pel_{mech} set is recommended.

We know that the choice of the locations of the masses influences the values of the mass and spring parameters. In the I&F_{mech} set and the Pel_{mech} set, one mass represents the body and the other represents the cover. This is in contrast with the fact that both masses are placed at the glottal surface and in this way interact with the aerodynamic forces: one

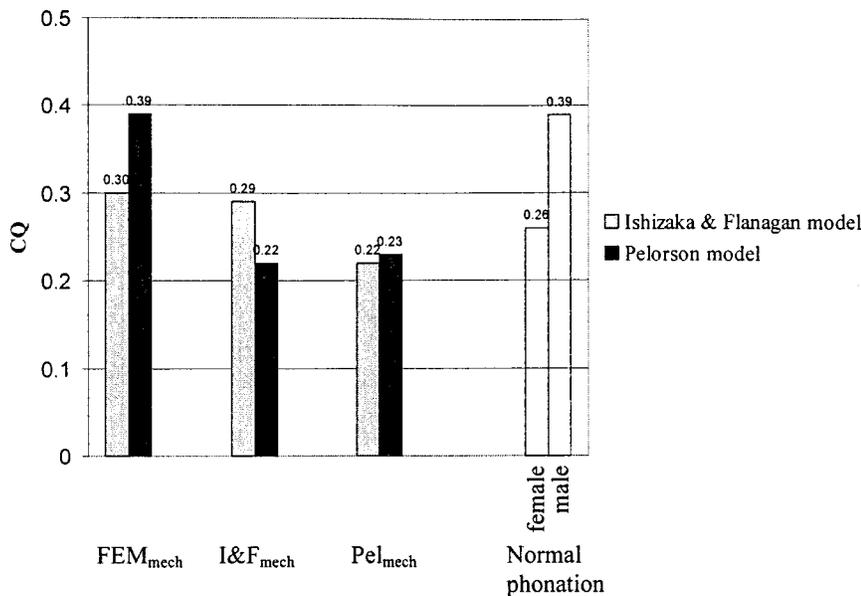


FIG. 7. Closed quotient with different aerodynamics and mechanics.

mass represents the surface (the cover mass) and one mass represents the deeper parts of the vocal fold (the body mass), whereas both masses interact with the airflow; the fact that an airflow interacts with the body of the vocal fold is physically unrealistic. In the FEM_{mech} set, this unrealistic situation is prevented by choosing L_1 and L_2 at the glottal surface. Not only the absolute values of m_1 and m_2 have been changed, also the ratio of m_1/m_2 has been changed. This ratio tends to approach one in the FEM_{mech} set, instead of five in the typical Ishizaka and Flanagan values. Lous *et al.* (1998) also proposed a smaller ratio of m_1/m_2 , resulting in a more realistic behavior of the vocal folds. The new spring stiffness values k_1 and k_2 also tend to differ less from each other than in the typical Ishizaka and Flanagan model. The smaller ratio of m_1/m_2 in combination with a smaller ratio k_1/k_2 results in the fact that the two oscillators are less different. In this way, the lower oscillator (consisting of m_1 and k_1) is less dominant over the upper oscillator (consisting of m_2 and k_2) as it is in I&F_{mech} and Pel_{mech}. In this way, the idea of having one oscillator that represents the body (the dominant one) and one oscillator that represents the cover (the less dominant one) is left. A distinction between body and cover is best described by Story and Titze (1995): two cover masses that describe the glottal surface, and a body mass deeper down the tissue of the vocal fold; values for these masses can also be determined by the method proposed in this study by setting up a new system of equations and using the same FEM model.

The values of the dampers were not part of the study. When the values of the masses and spring parameters change, the values of the dampers change according to $r_i = 2\zeta_i\sqrt{m_i \cdot k_i}$, $i=1,2$. The damping ratio ζ_1 is 0.1 and ζ_2 is 0.6 in both the Ishizaka and Flanagan model and in the Pelorson model. The reason for this choice is not given. The choice of a large difference between the two damping ratios is probably made because of the large differences in the values of the masses. If this were the reason, we would propose two damping ratios with values closer to each other, because we established values for the masses that are more equal. Story and Titze (1995) also chose two masses at the glottal surface with equal numerical values, in combination with damping ratios both equal to 0.2.

Implementation of the FEM_{mech} set in the Ishizaka and Flanagan model as well as in the Pelorson model results in self-sustained oscillation. Using the FEM_{mech} set, combined with both aerodynamic models, F_0 falls between the normal ranges of female and male voice, like the I&F_{mech} set does. The Pel_{mech} set results in an F_0 that falls more in the range of normal male voice. A wide range of values of U_g and CQ has been reached using the three mechanical parameter sets in combination with the two two-mass models. This wide variety shows the sensitivity of the system of the values of the masses and springs. Viewing the results of the glottal waves, the lumped parameters models with the FEM_{mech} set produce glottal waves that are within physiologic ranges considering F_0 and CQ . For U_g , the glottal waves produced with the FEM_{mech} set deviate relatively less from the physiologic values than the I&F_{mech} set and the Pel_{mech} set. This means, by using an accurate FEM description of the vocal

fold, values for the lumped parameters can be used which result in realistic glottal waves. Despite the fact that the values of the I&F_{mech} set and the Pel_{mech} set are chosen by the authors to give an optimal output (by means of the glottal waves), the lumped parameter values resulting from the presented study give glottal waves with at least the same degree of reality.

Introducing more detailed information can improve the numerical method for the determination of the lumped parameters. For example, nonlinear material behavior could be implemented in the FEM model. To determine the mass parameters, modal analyses cannot be performed on a nonlinear model. The harmonic response of the system can be used instead.

The numerical method presented can be used to determine values for mass and spring parameters for different laryngeal states. Applying different longitudinal tensions, the laryngeal state is changed (Titze and Strong, 1975; Stevens, 1981). In this way different registers can be investigated using the lumped parameter model. Ishizaka and Flanagan simulated different registers by applying a tension factor Q that reduces the values of the mass parameters and increases the values of the spring parameters when the register was changed from modal to falsetto. This resulted also in an increase of F_0 .

CONCLUSIONS

It can be concluded that, using our approach with the FEM, it is possible to achieve parameter values for lumped parameter models of the vocal folds, which are based on assumptions that are more realistic than the assumptions used before. The requirement of an equal dynamic response of a two-mass model of the vocal fold and of a FEM model of the vocal fold results in a set of lumped parameters. Self-sustained oscillation occurs when the lumped parameters are combined with aerodynamic forces that occur in the glottis. The physiologic parameters resulting from the glottal waves produced with the new set of lumped parameters show that the glottal waves are within normal ranges of phonation.

ACKNOWLEDGMENT

This research was supported by the Technology Foundation.

- Baer, T. (1981). "Observation on vocal fold vibration: Measurement of excised larynges," in *Vocal Fold Physiology*, edited by K. Stevens and M. Hirano (University of Tokyo, Japan), pp. 119–133.
- Berry, D. A., Herzel, H., Titze, I. R., and Krischer, K. (1994). "Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions," *J. Acoust. Soc. Am.* **95**, 3595–3604.
- Berry, D. A., and Titze, I. R. (1996). "Normal modes in a continuum model of vocal fold tissues," *J. Acoust. Soc. Am.* **100**, 3345–3354.
- Herzel, H., Berry, D., Titze, I. R., and Steinecke, I. (1995). "Nonlinear dynamics of the voice: Signal analysis and biomechanical modelling," *Chaos* **5**(1), 30–34.
- Herzel, H., Berry, D., Titze, I. R., and Saleh, M. (1994). "Analysis of vocal disorders with methods from nonlinear dynamics," *J. Speech Hear. Res.* **37**, 1008–1019.
- Hirano, M. (1974). "Morphological structure of the vocal cord as a vibrator and its variations," *Folia Phoniatri.* **26**, 89–94.

- Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (1989). "Glottal airflow and translottal air pressure measurements for male and female speakers in low, normal, and high pitch," *J. Voice* **3**(4), 294–305.
- Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell Syst. Tech. J.* **51**, 1233–1267.
- Koizumi, T., Taniguchi, S., and Hiromitsu, S. (1987). "Two-mass models of the vocal cords for natural sounding voice synthesis," *J. Acoust. Soc. Am.* **82**, 1179–1192.
- Lous, N. J. C., Hofmans, G. C. J., Veldhuis, R. N. J., and Hirschberg, A. (1998). "A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design," *Acta Acustica* **84**, 1135–1150.
- Myerson M. C. (1964). *The Human Larynx* (Charles C. Thomas, Springfield).
- Pelorson, X., Hirschberg, A., van Hassel, R. R., Wijnands, A. P. J., and Auregan, Y. (1994). "Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model," *J. Acoust. Soc. Am.* **96**, 3416–3431.
- Schutte, H. K. (1980). "The efficiency of voice production," Ph.D. thesis, University of Groningen.
- Steinecke, I., and Herzel, H. (1995). "Bifurcations in an asymmetric vocal-fold model," *J. Acoust. Soc. Am.* **97**, 1874–1884.
- Stevens, K. N. (1981). "Vibration modes in relation to model parameters," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo, Tokyo).
- Stevens, K. N. (1988). "Modes of vocal fold vibration based on a two-section model," in *Vocal Fold Physiology: Voice Production, Mechanisms and Functions*, edited by O. Fujimura (Raven, New York).
- Story, B. H., and Titze, I. R. (1995). "Voice simulation with a body-cover model of the vocal folds," *J. Acoust. Soc. Am.* **97**, 1249–1260.
- Titze, I. R. (1973). "The human vocal cords: A mathematical model. I," *Phonetica* **28**, 129–170.
- Titze, I. R. (1974). "The human vocal cords: A mathematical model. II," *Phonetica* **29**, 1–21.
- Titze, I. R., and Strong, W. J. (1975). "Normal modes in vocal cord tissues," *J. Acoust. Soc. Am.* **57**, 736–749.
- Titze, I. R., Baken, R. J., and Herzel, H. (1992). "Evidence of chaos in vocal fold vibration," *NCVS Status Progress Rep.* **3**, 39–63.

Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants

Stuart Rosen, Andrew Faulkner, and Lucy Wilkinson

Department of Phonetics and Linguistics, University College London, 4 Stephenson Way, London NW1 2HE, England

(Received 29 April 1999; revised 11 August 1999; accepted 16 August 1999)

Multi-channel cochlear implants typically present spectral information to the wrong “place” in the auditory nerve array, because electrodes can only be inserted partway into the cochlea. Although such spectral shifts are known to cause large immediate decrements in performance in simulations, the extent to which listeners can adapt to such shifts has yet to be investigated. Here, the effects of a four-channel implant in normal listeners have been simulated, and performance tested with unshifted spectral information and with the equivalent of a 6.5-mm basalward shift on the basilar membrane (1.3–2.9 octaves, depending on frequency). As expected, the unshifted simulation led to relatively high levels of mean performance (e.g., 64% of words in sentences correctly identified) whereas the shifted simulation led to very poor results (e.g., 1% of words). However, after just nine 20-min sessions of connected discourse tracking with the shifted simulation, performance improved significantly for the identification of intervocalic consonants, medial vowels in monosyllables, and words in sentences (30% of words). Also, listeners were able to track connected discourse of shifted signals without lipreading at rates up to 40 words per minute. Although we do not know if complete adaptation to the shifted signals is possible, it is clear that short-term experiments seriously exaggerate the long-term consequences of such spectral shifts. © 1999 Acoustical Society of America. [S0001-4966(99)02012-3]

PACS numbers: 43.71.Es, 43.71.Ky [JMH]

INTRODUCTION

Although multi-channel cochlear implants have proven to be a great boon for profoundly and totally deaf people, there is still much to be done in improving patient performance. One barrier to better results may be the fact that spectral information is typically presented in the wrong “place” of the auditory nerve array, due to the fact that electrodes can only be inserted partway into the cochlea. In a recent study which used x-ray computed tomography to measure electrode position in 20 patients with the Nucleus implant fully inserted, the most apical electrode was estimated to be at a cochlear place tuned to a mean frequency of about 1 kHz (Ketten *et al.*, 1998). Four of the patients (20%) had their most apical electrode at a location tuned to a frequency greater than 1400 Hz.

All multi-channel implants make use of a tonotopic presentation of acoustic information, using a bottom channel that is typically at a frequency lower than is reached by the most apical electrode. As clinical implant speech processors use channels tuned to as low as 200 Hz, it is clear that the place/frequency mismatch can be substantial. The net effect of such misplacement is a shift of spectral information to nerves that typically carry higher-frequency information.

Recent studies by Dorman *et al.* (1997) and Shannon *et al.* (1998) lend support to the notion that such a shift in spectral envelope can be devastating for speech perceptual performance. Shannon and his colleagues implemented a simulation of a four-channel cochlear implant, and used that to process signals for presentation to normal listeners. In their reference condition, channels were unshifted and spaced equally by purported distance along the basilar mem-

brane. Performance in this condition was worse than that obtained with natural speech, but still relatively high (about 80% of words in sentences). However, when the spectral information was shifted so as to simulate an 8-mm shift on the basilar membrane basalward, performance dropped precipitously (<5% of words in sentences). Dorman *et al.* (1997) also found significant decrements in performance for basalward shifts of 4–5 mm in a five-channel simulation. In both these studies, however, listeners were given little or no opportunity to adapt to such signals, so it is impossible to say of what importance such a mislocation of spectral shape is for cochlear implant users, who will be gaining experience with their implant typically for more than 10 h per day.

In fact, there is much evidence to support the notion that listeners can learn to adapt to such changes, and even more extreme ones. Blesser (1972, 1969) instructed pairs of listeners to learn to communicate in whatever way they could over a two-way audio communication channel that low-pass filtered speech at 3.2 kHz, and then inverted its spectrum around the frequency of 1.6 kHz. Although intelligibility over this channel was extremely low initially (in fact, virtually nil), listeners did learn to converse through it over a period of time. They also showed improved ability at perceiving processed unknown sentences, although even after about 10 h of experience, performance was still relatively low (a mean of 35% of syllables identified correctly).

There is evidence also from normal speech perception to suggest that an extraordinary degree of tolerance to acoustic variability must be operating. In vowel perception, for example, it is clear that the spectral information that distinguishes vowel qualities can only be assessed in a relative

manner, as different speakers use different absolute frequencies for the formants which determine spectral envelope structure. It might even be said that the most salient aspect of speech perception is the ability to extract invariant linguistic units from acoustic signals that vary widely in rate, intensity, and spectral shape.

In an attempt to address this issue, we implemented the type of signal processing used by Shannon *et al.* (1998), and tested our subjects on a similar range of speech materials with both spectrally shifted and unshifted speech. What makes this study very different is that our subjects were given an explicit opportunity to learn about the shifted signals, both by repeating the speech tests over a period of time, but more importantly, by letting them experience the frequency-shifted signals as receivers in Connected Discourse Tracking (De Filippo and Scott, 1978). The advantages of Connected Discourse Tracking for this purpose are manifold, insofar as it is a quantifiable, highly interactive task using genuine connected speech, and thus has high face validity. Using it, we are not only able to give our subjects extensive experience with constant feedback, but also to monitor their progress.

I. METHOD

A. Subjects

Four normally hearing adults, aged 18–22, participated in the tests. Two were male and two were female. All were native speakers of British English.

B. Test material

Three tests of speech perception were used. All were presented over Sennheiser HD 475 headphones without visual cues and without feedback. Two of these were computer-based segmental tests, with a closed set of responses. The intervocalic consonant, or VCV test (vowel–consonant–vowel) consisted of 18 consonants between the vowel /a/, hence /ama/, /aba/, etc, uttered with stress on the second syllable by a female speaker of Southern Standard British English. Every VCV was represented by at least 5 distinct tokens, with most having 12–15 tokens. Each of the consonants (/b tʃ d f g k l m n p r s ʃ t v w j z/) occurred three times in a random order in each test session, with the particular token chosen randomly without replacement on each trial. Listeners responded by using a mouse to select 1 of the 18 possibilities, displayed orthographically on the computer screen in alphabetical order (b c h d f g k l m n p r s sh t v w y z). Results were analyzed not only in terms of overall percent correct, but also for percent correct, and information transmitted with respect to the features of *voicing* (**voiced**/m n w r l j b d g z v/ vs **voiceless** /p t k tʃ ʃ s f/), *manner of articulation* (**nasal** /m n/ vs **glide** /w r l j/ vs **plosive** /b p d t g k/ vs **affricate** /tʃ/ vs **fricative** /ʃ s f z v/) and *place of articulation* (**bilabial** /m w b p/ vs **labiodental** /f v/ vs **alveolar** /n l j d t s z/ vs **palatal** /r tʃ ʃ/ vs **velar** /g k/). Note that studies like this often use an information transfer measure to analyze performance by feature, rather than percent correct. Although percent correct suffers from the drawback that different levels of chance performance are not compensated for

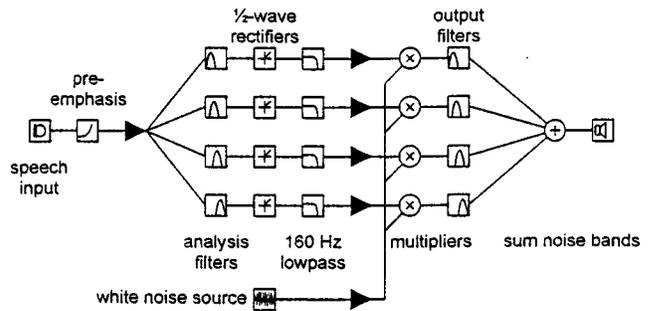


FIG. 1. Block diagram of the processing used for transforming the speech signal. Note that the filled triangles represent places where a gain adjustment can be made, but these were all fixed prior to the experiment.

in the calculation (e.g., that voicing judgements will be approximately 50% correct by chance alone whereas place judgements will be about 20% correct by chance), it is a more readily understood metric whose statistical properties are better characterized. Therefore, all statistical claims are made on the basis of percent correct only, although some summary statistics using information transfer are also presented.

The vowel test consisted of 17 different vowels or diphthongs in a /b/-vowel/-d/ context, in which all the utterances were real words or a common proper name—bad, bard, bared, bayed, bead, beard, bed, bid, bide, bird, bod, bode, booed, board, boughed, Boyd, or bud. The speaker was a (different) female speaker of Southern Standard British English. In each session, each vowel occurred three times in a random order, with each stimulus chosen randomly without replacement from a set of six distinct tokens. Again, listeners responded with a mouse to the possibilities displayed on the computer screen.

The third test consisted of the BKB sentence lists (Bench and Bamford, 1979). These are a set of 21 lists, each consisting of 16 sentences containing 50 key words, which are the only words scored. The particular recording (described by Foster *et al.*, 1993) used the same female speaker as the consonant test. Listeners wrote their responses down on a sheet of paper, and key words were scored using the so-called *loose* method (in which a response is scored as correct if its root matches the root of the presented word).

C. Signal processing

All signal processing was in real-time, implemented in the software system *Aladdin* (Nyvalla DSP AB, Sweden) and executed on a digital-signal-processing PC card (Loughborough Sound Images TMS320C31) running at a sampling rate of 22.05 kHz. The technique was essentially that described by Shannon *et al.* (1995) as shown in the block diagram in Fig. 1. The input speech was low-pass filtered, sampled, and pre-emphasized to whiten the spectrum for more accurate computations in mid- to high-frequency regions (first-order with a cutoff of 1 kHz). The signal was then passed through a bank of four analysis filters (sixth-order elliptical IIR) with frequency responses that crossed 15 dB down from the pass-band peak. Envelope detection occurred at the output of each analysis filter by half-wave rectification and first-order low-

TABLE I. Frequencies of the band edges used for the four output filters in the two main conditions of the experiment, specified in Hz. The analysis filters always used the *unshifted* frequencies.

	Band				
	1	2	3	4	
Unshifted	50	286	782	1821	4000
Shifted	360	937	2147	4684	10 000

pass filtering at 160 Hz. These envelopes were then multiplied by a white noise, and each filtered by a sixth-order elliptical IIR output filter, before being summed together for final digital-to-analog conversion. The gain of the four channels was adjusted so that a flat-spectrum input signal resulted in an output spectrum with each noise band having the same level (measured at the center frequency of each output filter).

Cross-over frequencies for both the analysis and output filters (Table I) were calculated using an equation and its inverse, relating position on the basilar membrane to its best frequency (Greenwood, 1990):

$$\text{frequency} = 165.4(10^{0.06x} - 1),$$

$$x = \frac{1}{0.06} \log\left(\frac{\text{frequency}}{165.4} + 1\right),$$

where x is position on the basilar membrane (in mm) from the apex, and *frequency* is given in Hz.

The *unshifted* condition, in which analysis and output filters had the same center frequencies, was obtained by dividing the frequency range from 50 to 4000 Hz equally using the equations above. This is similar to the LOG condition used by Shannon *et al.* (1998). In the *shifted* condition, output filters had their band edges increased upward in frequency by an amount equal to 6.46 mm on the basilar membrane (e.g., shifting 4 kHz to 10 kHz).

D. Procedure

In the first testing session, listeners were administered the three speech tests in each of three signal processing conditions: (1) normal speech (primarily to familiarize listeners with the test procedures, and not used with the BKB sentences); (2) unshifted four-channel; (3) frequency-shifted four-channel. One run of each of the vowel and consonant tests was performed with normal speech, and two runs of all three tests were presented for the two four-channel conditions.

Each subsequent testing session began with four 5-min blocks of audio-visual connected discourse tracking (CDT—De Filippo and Scott, 1978) with a short break between blocks. The talker in CDT was always the same (the third author). Talker and receiver faced each other through a double-pane glass partition in two adjacent sound-proofed rooms. The receiver wore Sennheiser HD475 headphones through which the audio signal was presented. Near the receiver was a stand-mounted microphone to transmit the receiver's comments undistorted to the talker. All CDT was done with the audio channel to the receiver undergoing the frequency-shifted four-channel processing. A low-level

masking noise was introduced into the receiver's room so as to ensure the inaudibility of any of the talker's speech not sufficiently attenuated by the intervening wall. Talker and receiver worked together to maximize the rate at which verbatim repetition by the receiver could be maintained. The materials used for CDT were of a controlled level of grammar and vocabulary, being drawn from the Heinemann Guided Readers series aimed at learners of English. The initial stages of CDT were performed audio-visually because it seemed highly unlikely that any subject would be able to track connected speech at all on the basis of the *shifted* sound alone at the beginning of the training.

In the sixth to tenth testing session, the first 5-min block of CDT was audio-visual, as in the previous sessions. Then visual cues were removed by covering the glass partition, and the second block of CDT was attempted in an audio alone condition. If the receiver scored more than ten words per minute (wpm), the remaining two blocks of CDT were conducted in the audio alone condition. If, however, the receiver scored less than 10 wpm, visual cues were restored for the remaining two 5-min blocks of CDT.

After each CDT training session, subjects were required to repeat the three speech perception tests given on the initial session (again for two runs of each test), but only in the *shifted* condition. After ten sessions of training (each consisting of four 5-min blocks of CDT) and testing, a final set of tests in the *unshifted* condition was also performed.

The 21 BKB sentence lists were presented in numerical order starting from list 1 in the *unshifted* condition, thus list 3 for *shifted*. As our experimental design required 24 lists, three lists only were repeated at the very end of testing (list 3 for the last test in the *shifted* condition, and lists 4 and 5 for the last tests in the *unshifted*). As scores were near zero for these lists when first presented; more than a week passed between presentations, and no feedback was ever given, it seems highly unlikely that subjects would have improved their performance on these lists simply through having experienced them before. In any case, we would expect any improvement to figure larger for the *unshifted* than the *shifted* condition.

E. Analysis

A common set of statistical analyses was performed for results from the three speech tests. As all these scores represent binomial variables, a statistical modeling approach based on logistic regression was taken, using the GLIM 4 system (Francis *et al.*, 1993). Unless otherwise stated, all statistical claims are based on changes in deviance at a 0.05 significance level.

One analysis concerned performance in the first and last sessions, the only two sessions in which results from both test conditions (*shifted* and *unshifted*) were obtained. Session and condition were treated as two-factor categorical variables, and listener as a four-factor variable. Of particular interest was the significance of a session by condition interaction, which would indicate the extent to which performance improves more across sessions for one condition than the other. In addition, a number of subsidiary analyses were performed, focusing on performance in either one session, or

TABLE II. Percent correct scores obtained in the recorded speech tests for the *unshifted* (**un**) and *shifted* (**shft**) conditions in the first testing session. Scores for each subject represent a mean of two tests.

Subject	BKB		bVd		VCV		Place		Voicing		Manner	
	un	shft	un	shft	un	shft	un	shft	un	shft	un	shft
CP	69	1	39	5	52	37	59	44	98	97	81	78
NW	64	0	43	5	57	32	61	38	94	85	76	80
SM	62	0	41	4	52	30	65	40	97	92	82	81
YW	61	2	45	5	55	33	74	42	94	82	90	74
mean	64	1	42	5	54	33	65	41	96	89	82	78

one condition. The aim of these analyses was to determine the extent to which performance improved in either condition, and whether performance was better for one condition or the other at the two points in time.

The other main analysis concerned trends across sessions for the *shifted* condition only, focusing on the extent to which increases in performance were significant, and the extent to which they appeared to be slowing over sessions. Again, a logistic regression was used, here to look for significant linear and quadratic trends across session. Session number was treated as a continuous variable, and listener as a four-category factor. A significant positive linear trend indicates performance is improving, while an additional significant quadratic trend always indicated a deceleration in the increase of performance. Although it is typical to use a logit link in such analyses, here an identity link was used. In this way, a linear trend in the statistical model corresponds exactly with a linear trend in proportion correct as a function of session number. In fact, the analyses were done with both link functions. Although differences arose in the details of the statistical models resulting, a change of link never resulted in a different substantive conclusion.

The analysis of the results from CDT also explored the existence of linear and quadratic trends across session, although under the assumption that the rate obtained, in words per minute, could be modeled as a Gaussian random variable. These analyses thus took the form of a general linear model (analyses of variance and covariance).

II. RESULTS

A. Initial test session

As expected, performance was high when the subjects were presented with natural speech. The mean score was 98.6% correct (range: 96.3–100.0) for the VCVs, and a little lower for the vowels (mean of 91.6% and a range of 86.0–96.1).

In the unshifted condition, performance was worse than with natural speech (as would be expected from Shannon *et al.*, 1995), but still quite high, as seen in Table II. The shift in spectrum, however, had a devastating effect on speech scores, especially for those tests that require the perception of spectral detail for good performance.

For the understanding of BKB sentences, mean performance dropped from 64% of key words correct to just under 1%. Vowel perception, too, was severely affected. Performance on VCVs was least affected, primarily because manner and voicing were relatively well received. These features

are known to be well signalled by temporal (Rosen, 1992) and gross spectral cues—cues which are apparently not disrupted by the spectral shift. Place of articulation, depending as it does upon fine spectral cues, was the most perceptually degraded phonetic feature.

B. Connected discourse tracking (CDT)

Although the main purpose of CDT was to provide a highly interactive training method, it is of interest to examine the trends found (Fig. 2). Only one subject (CP) failed to meet the criterion of 10 wpm in the auditory alone condition consistently for sessions 6–10, and even he met it on two of the sessions.

As would be expected, audio-visual performance was always considerably better than that obtained from auditory cues alone. There was also a clear improvement in the audio-visual condition, especially in the initial sessions. Three of the four subjects (excepting YW) exhibited a statistically significant linear increase in performance across session ($p < 0.01$). CP and YW both showed a significant quadratic trend across sessions, consistent with a deceleration in im-

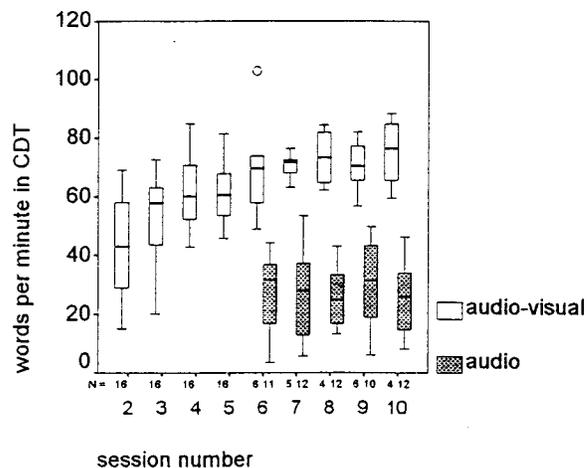


FIG. 2. Box plots of rates obtained from each 5-min run of Connected Discourse Tracking (across subjects) as a function of session. The box indicates the inter-quartile range of values obtained, with the median indicated by the solid horizontal line. The range of measurements is shown by the whiskers except for points more than 1.5 (indicated by “○”) or 3 box lengths (“*”) from the upper or lower edge of the box. Although no “*” appears on this plot, box plots are also used for Figs. 3–9, where these symbols do sometimes occur. The small numbers below the abscissa under each box indicate the total number of measurements used in the construction of each “box.” Two errors in the number of blocks run led to an extra CDT score in each of session numbers 6 and 7.

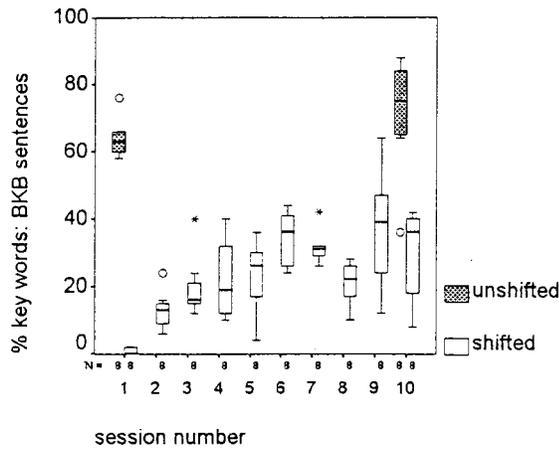


FIG. 3. Box plots of performance with BKB sentences, as a function of session and condition, across subjects.

provements in performance. Similarly, an ANOVA of the rates obtained without regard to subject revealed linear and quadratic trends over session.¹

The same ANOVA over subjects revealed no trends in the auditory alone condition, indicative also of a greater heterogeneity of trends across listeners. CP and SM showed no linear trend, NW showed a significant increase in CDT rate, whereas YW showed a significant linear *decrease* in CDT rate across session. Given the relatively high variability across sessions of CDT, and the small number of measurements, this heterogeneity should not be too surprising.

In short, whereas performance in the audio-alone condition appears to be stable, performance improves across sessions in the audio-visual condition, but to a diminishing degree toward the later sessions. Note too that audio-visual tracking rates become quite high in the later sessions (maximum rates of CDT under ideal conditions are about 110 wpm—De Filippo and Scott, 1978), and this also may be limiting the rate of increase that is possible.

C. Sentences (BKB)

Figure 3 shows the results obtained in the BKB sentence test. As noted above, performance is far superior for *unshifted* speech in session 1. However, performance improved significantly across sessions in the *shifted* condition, even though it did not reach the level obtained for *unshifted* speech. A statistical model of the results obtained only in sessions 1 and 10 showed a strong *session by condition* interaction ($p < 0.0001$), indicating that performance increased more in the *shifted* condition than in the *unshifted* condition. Other analyses showed significant improvements across sessions in both conditions, and that performance remained superior for the *unshifted* speech even in session 10.²

Statistical trends across sessions were generally similar to those found for audio-visual CDT. A model describing performance in the *shifted* condition showed a quadratic dependence of words correctly identified on session, indicating performance to be increasing over sessions, with the greatest increases in the early sessions. Although the complexity of the statistical model (and the paucity of data) makes rigorous investigation of individual differences difficult, inspection of

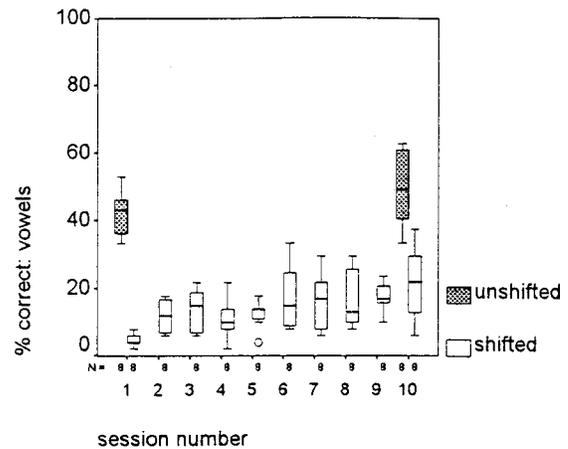


FIG. 4. Box plots of performance on the vowel test, as a function of session and condition, across subjects.

the model fits showed three of the listeners to be very similar in overall levels of performance and trends across sessions. CP was unusual in showing much less pronounced improvements over time. Although a linear fit to CP's data showed a significant increase in performance across sessions, the slope obtained was about half of that calculated from the other three listeners.

D. Vowels

Results for the vowel test are displayed in Fig. 4. Looking first only at results obtained in sessions 1 and 10, the pattern is as found for BKB sentences. Performance was always worse in the *shifted* condition, even though it improved significantly over the course of training. The increase in performance in the *unshifted* condition is barely significant ($p \approx 0.03$) whereas it is highly significant in the *shifted* condition ($p < 0.0001$). This is also reflected in a strong *session by condition* interaction.

Trends across sessions were somewhat different than those found for sentences. Here, there was only evidence for a linear improvement in performance, with no significant quadratic term. It therefore appears that performance increased linearly over session, with no evidence of a deceleration. This pattern held for all four listeners, although the slopes varied significantly between them. CP again exhibited a shallower slope than the others, although it did differ significantly from zero even when tested in a separate analysis of that data alone.

E. Intervocalic consonants (VCVs)

Figure 5 shows performance on the VCV test pooled across listeners. Analysis of the *shifted* results shows a significant linear effect of session, with no quadratic trend, just as found with the vowels. The statistical effects were smaller though, and the measured slopes considerably less shallow. In addition, there is no statistical evidence in the complete model of any differences in the slopes among listeners, although they did differ in overall level of performance.

Analyses of sessions 1 and 10 again exhibited a strong *session by condition* interaction ($p < 0.01$), showing that

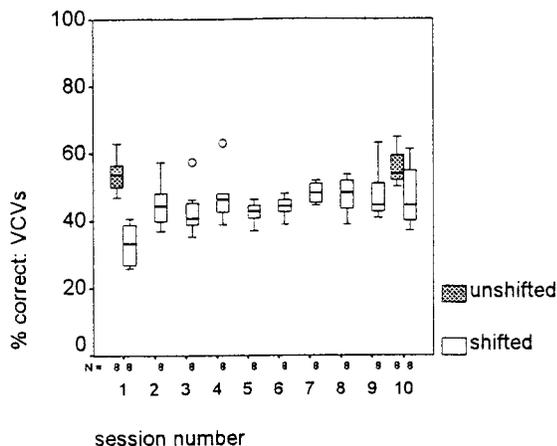


FIG. 5. Box plots of percent correct in the VCV test as a function of session number for both *shifted* and *unshifted* conditions, across subjects.

performance increased more in the *shifted* than in the *unshifted* condition. In fact, the results from sessions 1 and 10 in the *unshifted* condition are not statistically different. Moreover, performance with *unshifted* speech was only better than that for *shifted* speech in the first session. At session 10, results from the two conditions are not statistically different. This outcome is quite different to those from the other speech tests, in which performance in the *shifted* condition never reached that attained in the *unshifted* condition.

A slightly different outcome arose for the perception of place of articulation (Fig. 6). As for percent correct, performance in the *unshifted* condition did not change across sessions, and *shifted* performance in session 1 was poorer than for *unshifted* speech. Here, however, *shifted* performance at session 10 still did not reach the level of the *unshifted* condition, even though it was significantly better than at session 1. But, just as with percent correct, the *shifted* results show a significant linear (but no quadratic) trend across the ten sessions.

Changes in the accuracy of voicing and manner perception were smaller through training, as would be expected from the greater role temporal and gross spectral aspects play

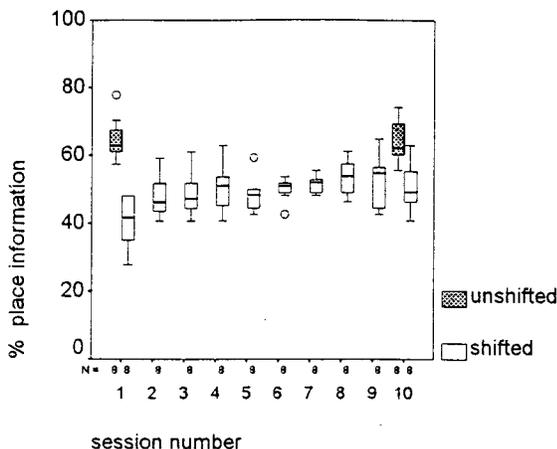


FIG. 6. Percent correct for place of articulation in the VCV test as a function of session number for both *shifted* and *unshifted* conditions.

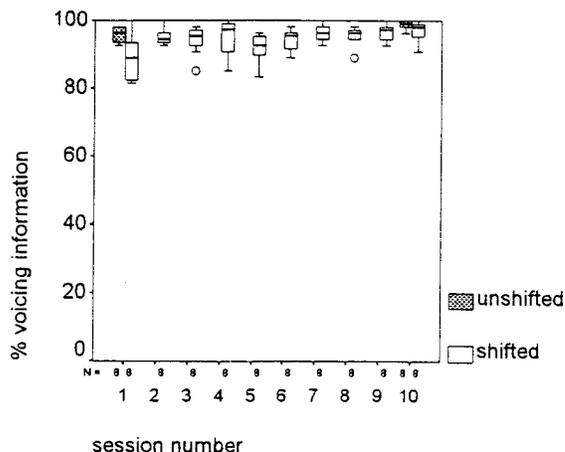


FIG. 7. Percent correct for voicing in the VCV test as a function of session number for both *shifted* and *unshifted* conditions.

in signaling these features and the higher initial performance levels (Figs. 7 and 8). Results for voicing were different to those found previously in that there was no significant *condition* by *session* interaction—rather, there were significant main effects of both factors. This indicates that performance increased to the same degree for *shifted* and *unshifted* speech, and that performance with *shifted* speech was inferior at both sessions. These changes in performance are small, however, and may also be constrained by ceiling effects.

For manner, there was again a significant *condition* by *session* interaction, but here performance in the *shifted* condition was actually better on average than in the *unshifted* condition. Subsidiary analyses show that performance increased significantly across sessions only for *shifted* speech.

Both voicing and manner perception showed significant linear trends across the ten testing sessions with *shifted* speech. Manner perception also exhibited a significant quadratic term. The form of the predictions for manner were quite varied across subject, with overall changes small in any case. The most important outcome for both these phonetic features was a significant improvement over time (albeit small), but the degree of improvement may have been limited by the relatively high performance overall.

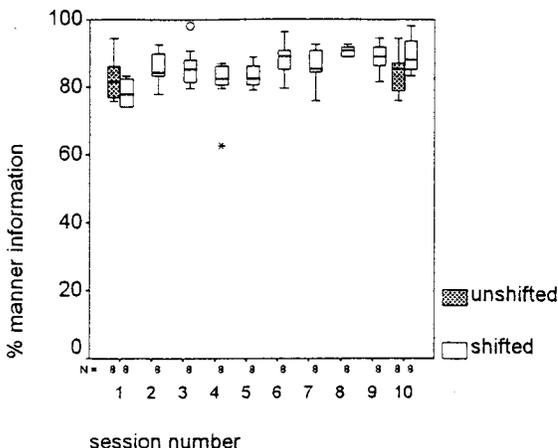


FIG. 8. Percent correct for manner of articulation in the VCV test as a function of session number for both *shifted* and *unshifted* conditions.

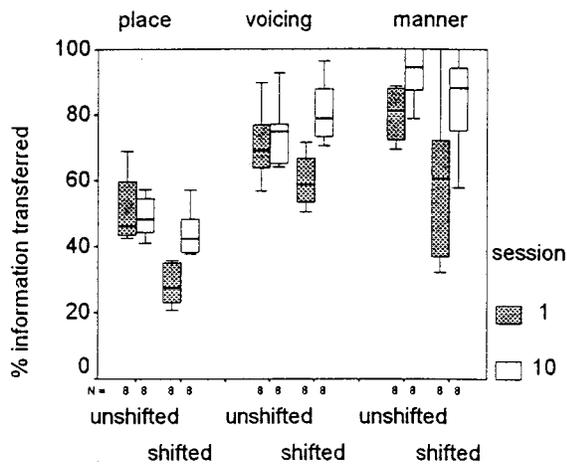


FIG. 9. Information transfer statistics calculated for each of the three phonetic features (place, manner, and voicing) as a function of condition and session.

In order to make comparisons of performance across phonetic features that are uncontaminated by different levels of chance performance, Fig. 9 shows information transfer measures as a function of session number and condition. Clearly, overall performance was poorest for place of articulation, the phonetic feature which depends most strongly on fine spectral detail. Differences between the *shifted* and *unshifted* conditions were larger in the first session, and tended to diminish over the course of training.

To summarize, performance in the VCV task for *shifted* speech improved over the course of training, with overall accuracy and perception of manner statistically indistinguishable from the *unshifted* condition. However, the results from the perception of place of articulation, expected to be most affected by frequency shifts, suggest that subjects had not quite reached the level of performance they were able to obtain with *unshifted* speech.

III. DISCUSSION

Two aspects of the current study seem especially striking. First, there is the enormous decrement in performance in understanding speech when processed to contain envelope information in four spectral channels when these are shifted sufficiently in frequency (a fact already known from the earlier study of Shannon *et al.*, 1998). That different tests suffer different degrees of degradation is easily understood, as it would be expected that speech materials that require effective transmission of detailed spectral information for good performance (e.g., vowels and sentences) would be more affected by a spectral shift than those in which much can be apprehended through temporal cues or gross spectral contrasts (e.g., consonants).

Second, there is the remarkable speed at which listeners learn to compensate for the spectral shift. After just 3 h of experience (not counting the tests themselves, which consist of quite short periods of speech without feedback), performance in the most severely affected tasks (vowels and sentences) increased from near zero to about one-half the performance in the *unshifted* condition. Also, all listeners

exhibited at least some improvement in all three recorded speech tests, even though the degree of improvement appeared to vary across listeners.

It might be argued that an important part of this improvement with the *shifted* signals reflects adaptation not only to the spectral shift, but to other aspects of the stimulus transformation and/or testing procedures. We would, however, expect learning of these latter aspects to be reflected in changes in performance for the *unshifted* signals as well. Although there was learning of this kind, the improvements tended to be small. All three recorded speech tests showed strong statistical evidence that performance increased more for the *shifted* signals than for the *unshifted* ones. It is therefore clear that listeners are learning *something* about the spectral shift, although it is impossible for us to say exactly what that is.

In light of this evidence, it might seem odd that the audio-alone condition of CDT showed no improvements over sessions. For one thing, it is clear that some very large learning effects at the start of training have been missed (where tracking rates would have been near zero), as subjects did not attempt this condition until session 6, where they already were tracking at a median rate of 30+ wpm. There was also a greater degree of variability among the subjects. It may well be that the relatively large variability shown in the CDT task has masked any trends, that training needs to be done over significantly longer periods of time, or even that the level of performance reached here represents the maximum that will ever be achieved. Given the consistent evidence of improvements in all the other tasks, this last possibility seems unlikely, but a clarification of this issue requires further study.

We cannot, of course, say anything about whether compensation to *shifted* signals would be complete after some further degree of training, how long it would take were it to be possible, nor the extent to which performance might improve with training for *unshifted* speech. Nor do we know the extent to which CDT is effective as a training procedure, whether other procedures would be better, nor indeed whether the progress the subjects made can be attributed primarily to the use of CDT. These, though, are secondary questions. What is clear is that subjects were able to improve their performance considerably over short periods of time, periods that are inconsequential from the point of view of an implant patient.

That implant patients do, in fact, adapt to an altered spectral representation is seen most clearly in a recent study by Fu and Shannon (1999). They manipulated the spectral representation of vowels in normal listeners (using techniques similar to the ones we used) and in users of cochlear implants. The normal listeners, who had little or no opportunity to adapt to the altered stimulation, always did best when the analysis and carrier bands matched in frequency (an *unshifted* condition). On the other hand, the implant users, with at least six months of experience, always showed best performance for the frequency allocations used in their every day speech processor (typically a *shifted* condition), in spite of the fact that electrode insertion depths varied widely over the group.

IV. SUMMARY AND FINAL REMARKS

Spectral distortions of the kind that are likely to be present in multi-channel cochlear implants can pose significant limitations on the performance of the listener, at least initially. With practice, a substantial part of these decrements can be erased. Although we cannot say on the basis of this study whether place/frequency mismatches can ever be completely adapted to, it is clear that short-term experiments seriously exaggerate the long-term consequences of such spectral shifts. If we were to argue that matching frequency and place is essential, then listeners with shallow electrode penetrations should not receive speech information below, say, 1–2 kHz. That such an approach would be preferable to one in which the lowest-frequency band of speech is assigned to the most apical electrode seems highly unlikely to us. For one thing, it is clear that the lower-frequency regions of speech are the best for transmitting the temporal information that can most suitably complement the information available through lipreading. Can we possibly imagine that the shallower an electrode array is implanted, the higher should be the band of frequencies we present to the patient? It may well be that patients with shallower electrode penetrations will perform more poorly on average than those with deeper penetrations. But this probably results more from the loss of access to the better-surviving apical neural population (Johnsson, 1985), or from the fact that the speech frequency range must be delivered to a shorter section of the nerve fibre array, than from the place/frequency mismatch per se. It seems entirely possible that the speech perceptual difficulties which implant users experience as a result of a place/frequency mismatch may be a short-term limitation overcome with experience.

ACKNOWLEDGMENTS

This work was supported by Defeating Deafness (The Hearing Research Trust) and a Wellcome Trust Vacation

Scholarship to LCW (Grant reference No. VS/97/UCL/016). Our gratitude goes to the reviewers Bob Shannon, Ken Grant, and James Hillenbrand for their useful comments on an earlier version.

¹The outlier in the audio-visual condition at session number 6 was excluded from the statistical analyses.

²These analyses excluded an outlier for one listener in the *unshifted* condition in session 10 where performance is about half the level expected, and in line with results from the *shifted* condition.

- Bench J., and Bamford, J. (Eds.) (1979). *Speech-hearing Tests and the Spoken Language of Hearing-impaired Children* (Academic, London).
- Blessner, B. (1972). "Speech perception under conditions of spectral transformation: I. Phonetic characteristics," *J. Speech Hear. Res.* **15**, 5–41.
- Blessner, B. A. (1969). "Perception of spectrally rotated speech," Unpublished Ph. D., MIT, Cambridge, MA.
- De Filippo, C. L., and Scott, B. L. (1978). "A method for training and evaluating the reception of ongoing speech," *J. Acoust. Soc. Am.* **63**, 1186–1192.
- Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). "Simulating the effect of cochlear-implant electrode insertion depth on speech understanding," *J. Acoust. Soc. Am.* **102**, 2993–2996.
- Foster, J. R., Summerfield, A. Q., Marshall, D. H., Palmer, L., Ball, V., and Rosen, S. (1993). "Lip-reading the BKB sentence lists: corrections for list and practice effects," *Br. J. Audiol.* **27**, 233–246.
- Francis, B., Green, M., and Payne, C. (Eds.) (1993). *The GLIM System: Release 4 Manual* (Clarendon, Oxford).
- Fu, Q. J., and Shannon, R. V. (1999). "Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing," *J. Acoust. Soc. Am.* **105**, 1889–1900.
- Greenwood, D. D. (1990). "A cochlear frequency-position function for several species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.
- Johnsson, L.-G. (1985). "Cochlear anatomy and histopathology," in *Cochlear Implants*, edited by R. F. Gray (Croom Helm, London).
- Ketten, D. R., Vannier, M. W., Skinner, M. W., Gates, G. A., Wang, G., and Neely, J. G. (1998). "In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays," *Ann. Otol. Rhinol. Laryngol.* **107**, 1–16.
- Rosen, S. (1992). "Temporal information in speech: Acoustic, auditory, and linguistic aspects," *Philos. Trans. R. Soc. London, Ser. B* **336**, 367–373.
- Shannon, R. V., Zeng, F. G., and Wygonski, J. (1998). "Speech recognition with altered spectral distribution of envelope cues," *J. Acoust. Soc. Am.* **104**, 2467–2476.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.

A method to determine the speech transmission index from speech waveforms

Karen L. Payton

*Electrical and Computer Engineering Department, University of Massachusetts Dartmouth,
North Dartmouth, Massachusetts 02747*

Louis D. Braida^{a)}

Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 12 November 1997; revised 7 June 1999; accepted 19 July 1999)

A method for computing the speech transmission index (STI) using real speech stimuli is presented and evaluated. The method reduces the effects of some of the artifacts that can be encountered when speech waveforms are used as probe stimuli. Speech-based STIs are computed for conversational and clearly articulated speech in several noisy, reverberant, and noisy-reverberant environments and compared with speech intelligibility scores. The results indicate that, for each speaking style, the speech-based STI values are monotonically related to intelligibility scores for the degraded speech conditions tested. Therefore, the STI can be computed using speech probe waveforms and the values of the resulting indices are as good predictors of intelligibility scores as those derived from MTFs by theoretical methods. © 1999 Acoustical Society of America. [S0001-4966(99)03411-6]

PACS numbers: 43.71.Gv, 43.72.Dv, 43.72.Ew [JH]

INTRODUCTION

The speech transmission index (STI) measures the extent to which speech envelope modulations are preserved in degraded listening environments. The STI has been shown to be highly correlated with speech intelligibility in a wide range of listening conditions (Houtgast and Steeneken, 1985; Humes *et al.*, 1986; Payton *et al.*, 1994). These conditions include additive noise, reverberation, and their combination. In each study mentioned, the STI was derived from measured changes in the modulation depth of modulated noise presented in the acoustic environment (referred to herein as traditional STI) or from acoustic theory, using signal-to-noise ratios (SNR), room reverberation times, and/or room impulse responses (referred to herein as theoretical STI).

There are situations though, when the traditional and theoretical techniques to compute the STI will not work. For example, speech must be used as the probe stimulus if the STI is to attempt to predict the effect of speaking style on speech intelligibility. Clearly articulated speech is more intelligible than conversational speech in noisy and reverberant environments (Payton *et al.*, 1994) even though the long term spectra of clear and conversational speech are very similar (Picheny *et al.*, 1986). The relatively high intelligibility of clear speech in the presence of reverberation may be related to differences in envelope modulation spectra associated with differences in speaking rate. To determine whether such differences are responsible for differences in intelligibility, speech waveforms characteristic of different speaking styles must be used to derive STI values.

Also, there has been considerable interest in using the STI to estimate the effects of some types of nonlinear signal processing, such as amplitude compression, on intelligibility (Plomp, 1988; Villchur, 1989; Ludvigsen *et al.*, 1990; Lud-

vigsen *et al.*, 1993; Drullman *et al.*, 1994; Hohmann and Kollmeier, 1995). Since such effects are not predicted by acoustic theory, empirical approaches have been used. Using speech rather than modulated noise as the probe stimulus would seem to offer considerable advantages when determining changes in modulation characteristics for compression schemes that primarily affect the speech envelopes rather than the fine structures (e.g., syllabic compression). Some of the aforementioned studies (Ludvigsen *et al.*, 1990, 1993; Drullman *et al.*, 1994; Hohmann and Kollmeier, 1995) have attempted to do just that with varying success.

First, though, it is important to demonstrate that a speech-based STI is as good as the traditional and theoretical STI methods in circumstances where the latter two are used successfully. When speech is used as a probe stimulus, artifacts are often observed in the measured envelope spectra of speech in degraded listening environments (Houtgast and Steeneken, 1972; Payton *et al.*, 1993; Ludvigsen, 1993; Hohmann and Kollmeier, 1995). The observed artifacts consist of increases in intensity envelope spectra when theory predicts decreases. These artifacts can seriously interfere with the computation of modulation transfer functions (MTFs), an intermediate step in the traditional and theoretical derivations of the STI. This paper reports our efforts to develop a method to compute the STI from MTFs using speech waveform probes that is relatively insensitive to the effects of such artifacts. The relationship of this speech-based STI to previously reported speech intelligibility measurements (Payton *et al.*, 1994) will be compared to that of the theoretically computed STI.

I. BACKGROUND AND METHODS

Although several approaches have been suggested to derive the STI from speech signals, we used digital signal processing techniques to implement a method based on the one

^{a)}Electronic mail: BRAIDA@mit.edu

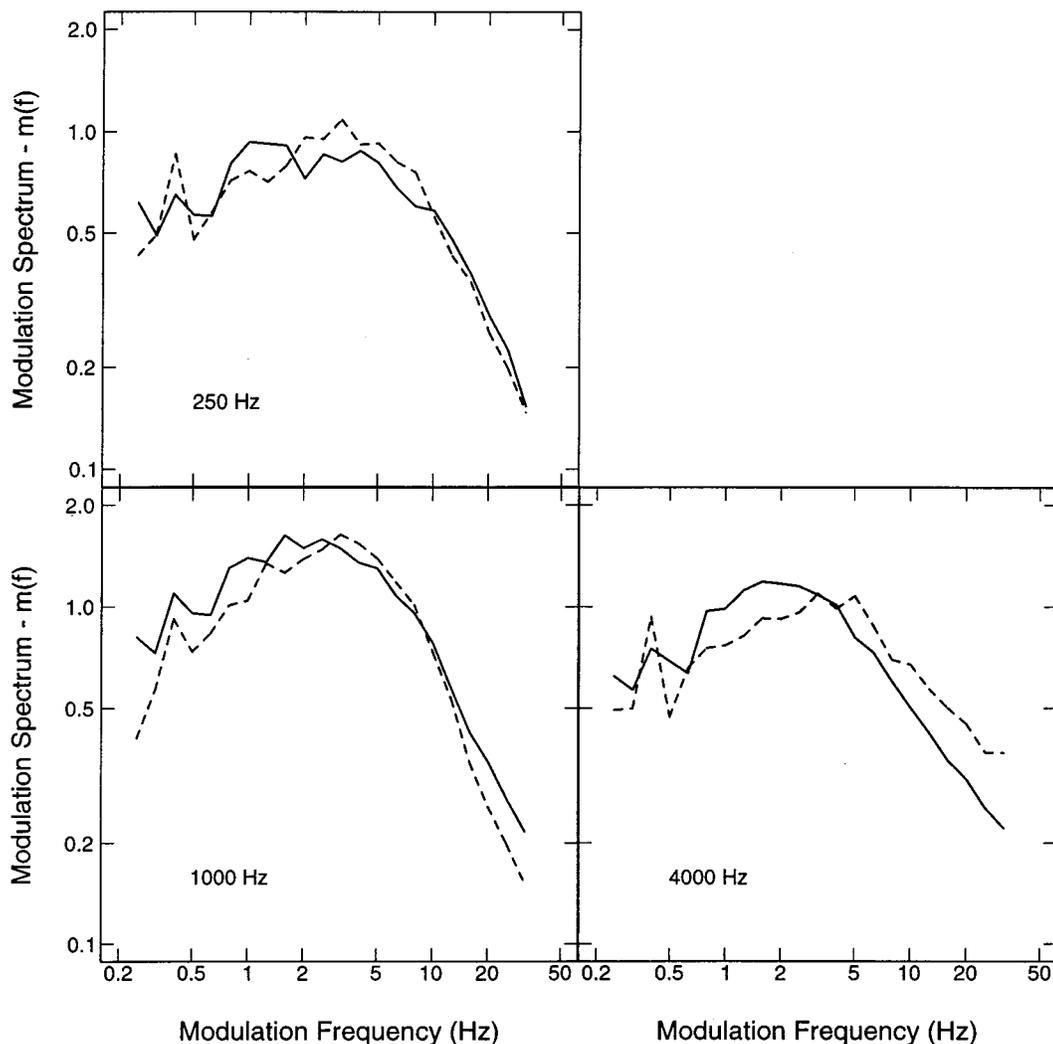


FIG. 1. One-third octave envelope spectra of filtered octave bands of conversational (dashed lines) and clear speech (solid lines).

originally proposed by Houtgast and Steeneken (1983, 1985). This method is the most similar to the traditional technique for computing the STI from modulated noise. The method also includes computation of the MTF as an intermediate result which can be compared directly with theoretically computed MTFs that are not subject to artifacts. Others have attempted to compute the STI from speech using correlation techniques that result directly in apparent signal-to-noise ratios rather than MTFs (Ludvigsen *et al.*, 1990; Holube and Kollmeier, 1996) or in modified MTFs (Drullman *et al.*, 1994).

Envelope spectra, from which MTFs and STIs were calculated, were computed from a corpus of 48 nonsense sentences¹ that were spoken both conversationally and clearly. The sentences were digitized at 20 000-Hz sampling rate with a 9500-Hz antialiasing filter. Sentences were concatenated into strings of four with a silent interval inserted before each sentence to mimic pauses in running speech.²

The environmental degradations were the same as those reported in Payton *et al.* (1994). Speech-shaped Gaussian noise was added to the sentence strings at specified signal-to-noise ratios for the noise conditions. The noise spectrum approximated the combined average long-term spectra of

clear and conversational speech. The approximation resulted in slight SNR variations from octave band to octave band when compared with either individual speaking style. Reverberant speech was created by convolving the sentence strings with simulated room impulse responses (Peterson, 1986). For conditions with both noise and reverberation, the noise was added first and then the noisy strings were reverberated.

Twelve speech strings, each of duration 8.2 s (for conversational speech) or 16.4 s (for clear speech) were extracted from the concatenated sentences. The clear speech strings are twice as long as the conversational strings because clear speech is spoken at roughly half the speed as conversational speech. A 250-ms cosine ramp was applied to the beginning and end of each string.

Seven octave-band envelope strings were derived from each concatenated sentence string as follows. First, each sentence string was bandpass filtered by a bank of eight-pole octave-bandwidth digital Butterworth filters, with center frequencies in the range 125–8000 Hz (the 8000-Hz filter was actually a high-pass filter). Each filtered string was squared, low-pass filtered by convolution with a 16.6-ms Hamming window, and then down-sampled by a factor of 10. A Hamming window was used as the low-pass filter impulse re-

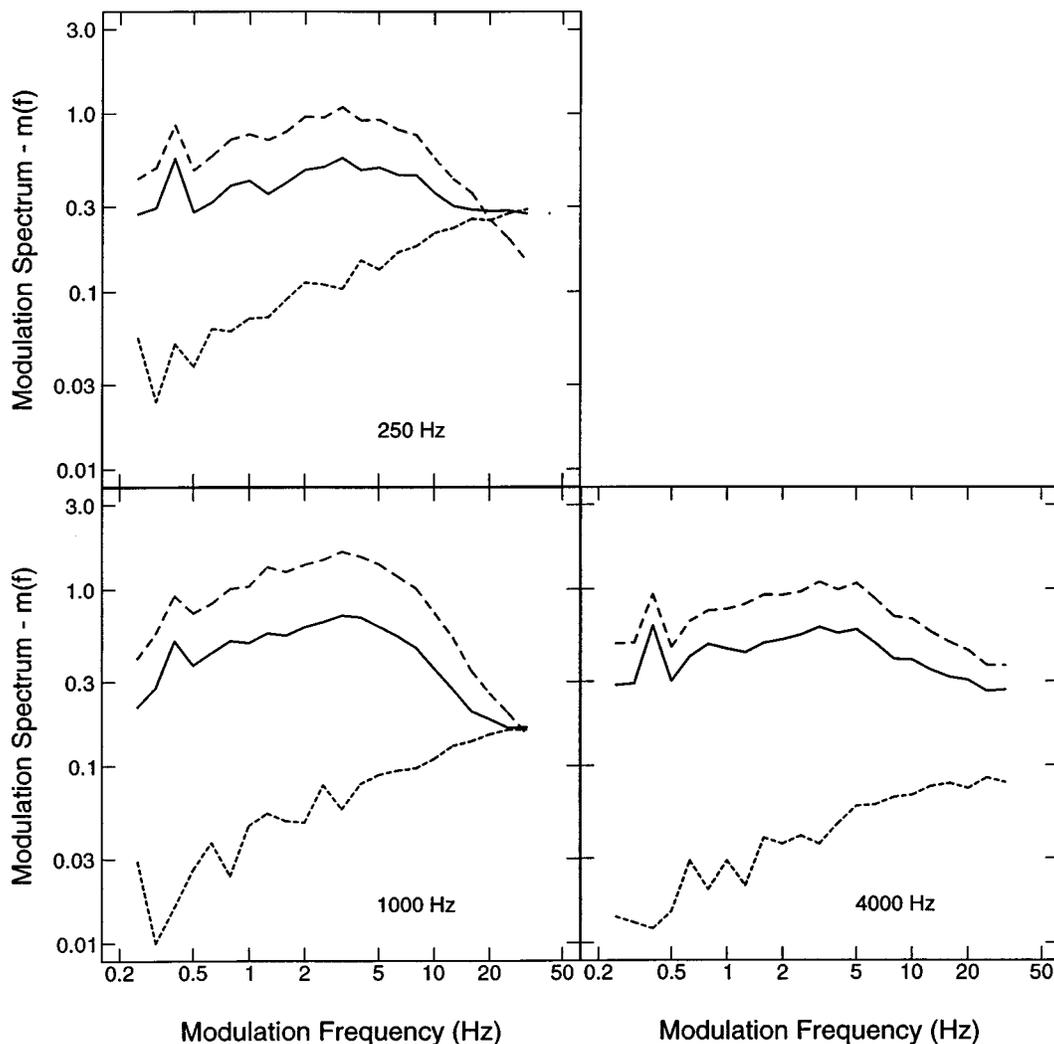


FIG. 2. One-third octave envelope spectra of filtered bands of conversational speech under conditions of no degradation (dashed lines) and additive speech spectrum noise at 0 dB SNR (solid lines). In each panel the dotted line denotes the envelope spectrum of the speech-shaped noise.

sponse since it is always greater than zero, to avoid producing negative values in the filtered envelope.

Power spectra of the envelope strings were computed from sequences of length 32 768 samples (required to provide adequate frequency resolution at the 2000-Hz sampling rate). For conversational speech, pairs of envelope strings were concatenated to provide the required number of samples. For clear speech no such concatenation was needed. The power spectra were then normalized so that 100% intensity modulation at a single frequency would correspond to unity. An ensemble average envelope spectrum was then computed for each octave band from 6 conversational or 12 clear sequences. The average spectra were reduced to one-third octave representations by summing components over one-third octave intervals with center frequencies ranging from 0.25 to 25 Hz. The square root of the resulting sum was taken as the one-third octave modulation spectrum. This wide modulation frequency range is examined for two reasons. First, the modulation range cited in the literature has varied over the years (Houtgast and Steeneken, 1973, 1985; Steeneken and Houtgast, 1980). Second, in the case of mild reverberation, an overly conservative (i.e., low) upper bound

on the MTF could result in an overprediction of intelligibility.

II. MEASUREMENTS

Three one-third octave envelope modulation spectra of octave bands of undistorted speech, computed as described in Sec. I, are shown in Fig. 1 for each speaking style. These spectra are generally consistent with those reported by Houtgast and Steeneken (1985). Note that the clear speech envelope spectra peak at lower modulation frequencies than the conversational envelope spectra but, otherwise, are very similar. In the remainder of this paper, conversational envelope spectra and MTFs of these three bands (centered at 250, 1000, and 4000 Hz) are used to illustrate effects seen in the low-, mid-, and high-frequency parts of the speech spectrum for both speaking styles. Clear envelope spectra and MTFs are not included because they are affected by degradations in a manner very similar to conversational envelope spectra. Although the 8000-Hz band was not a full octave wide, the envelope spectra of that band behave according to the theoretical predictions discussed below.

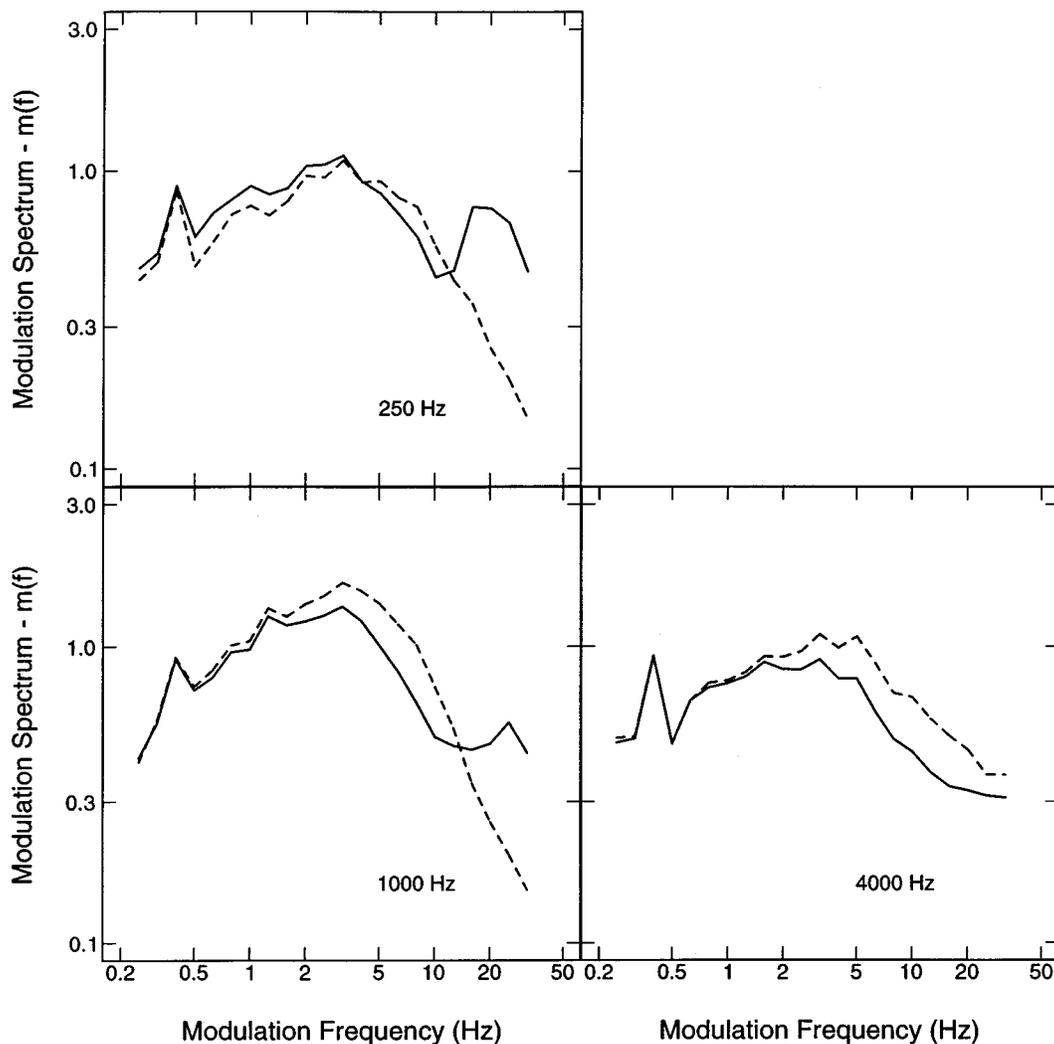


FIG. 3. One-third octave envelope spectra of filtered bands of conversational speech under conditions of no degradation (dashed line) and reverberation with RT=0.6 s (solid line).

The envelope spectra of noisy and/or reverberant speech often differ from theoretical predictions in important aspects. For example, when noise is added to speech, the valleys in the envelope are filled in, resulting in an attenuation of the envelope spectrum that is expected, on theoretical grounds, to be relatively equal at all modulation frequencies. However, as can be seen in Fig. 2, the modulation spectra are not always attenuated as much at high modulation frequencies as at low frequencies and, in some instances, can be larger than for undegraded speech (Payton *et al.*, 1993; Hohmann and Kollmeier, 1995). This situation is usually observed only in the octave bands with low center frequencies (Fig. 2) and is rarely seen in high-frequency octave bands. Some computations of the STI from speech waveforms (e.g., Hohmann and Kollmeier, 1995) appear to have been based on MTF values derived from modulation spectra that were not attenuated by the interfering noise as much as expected, due in part to these artifacts.

In computing the STI for reverberant speech it is usually assumed that reverberation has little effect on envelope components with low modulation frequencies, but attenuates components at high modulation frequencies. This effect is similar to low-pass filtering of the speech envelope with a

cutoff frequency that depends on the reverberation time (RT). Measured speech-based envelope spectra are typically consistent with this analysis, but, as can be seen in Fig. 3, discrepancies are sometimes observed at high modulation frequencies. Although the modulation spectra of reverberant speech tend to roll off as expected, the rolloff does not always extend to the highest modulation frequencies, as is evident in the modulation spectrum of the octave band at 1000 Hz. Occasionally, as seen in the envelope spectra for the octave band of speech at 250 Hz, reverberation can even increase the depth of modulation at low modulation frequencies. Despite differences in the envelope spectra of the two speaking styles, there were no qualitative differences in the reverberant MTFs for the two speaking styles.

When both noise and reverberation degrade speech, both types of artifacts are present (Fig. 4). The envelope spectra of noisy-reverberant speech are attenuated across most modulation frequencies due to the effects of the additive noise and further attenuated at high modulation frequencies due to the effects of reverberation. However, artifacts can introduce anomalous increases in depth of modulation for low-frequency octave bands of speech (Fig. 4). For these bands the depth of modulation for the noisy-reverberant speech can

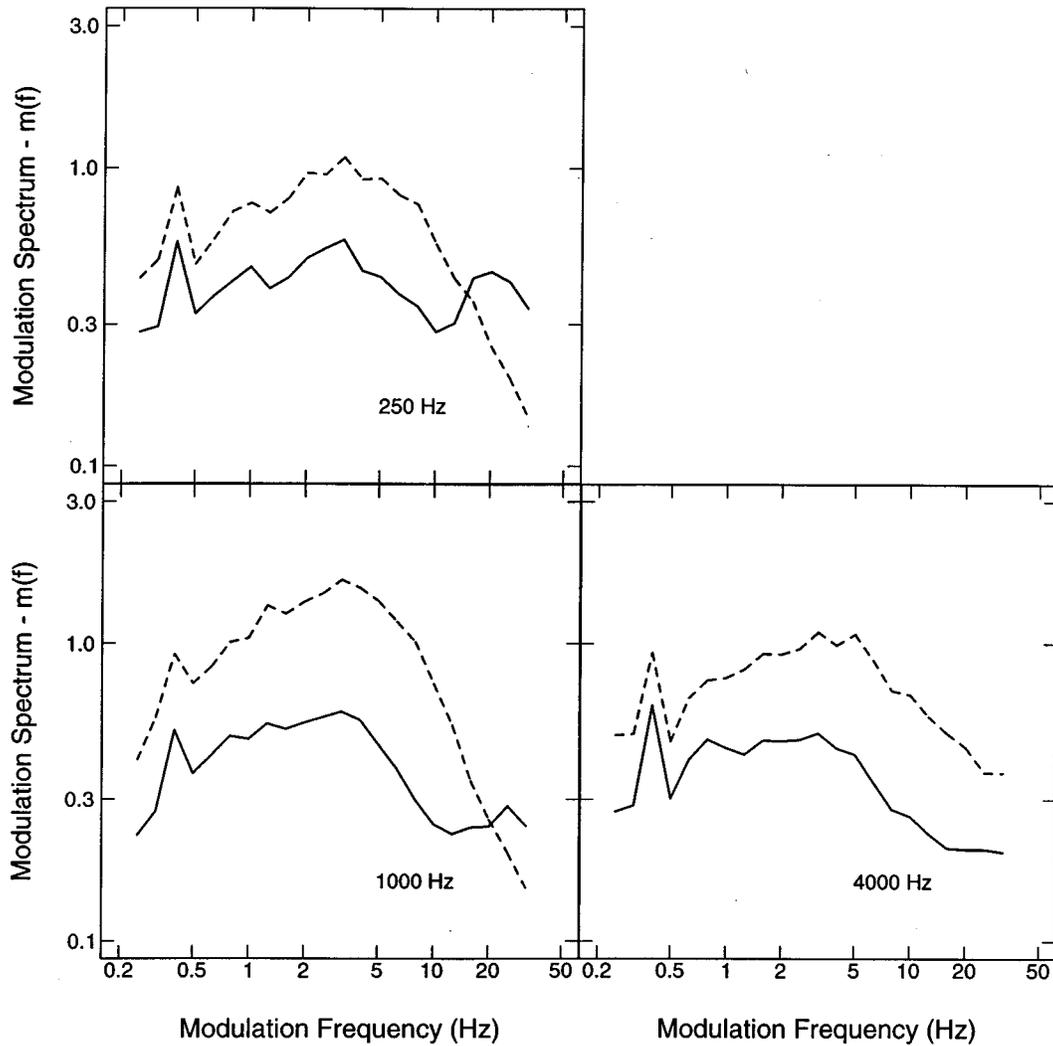


FIG. 4. One-third octave envelope spectra of filtered bands of conversational speech under conditions of no degradation (dashed line) and degradation by additive speech spectrum noise (SNR=0 dB) and reverberation with RT=0.6 s (solid line).

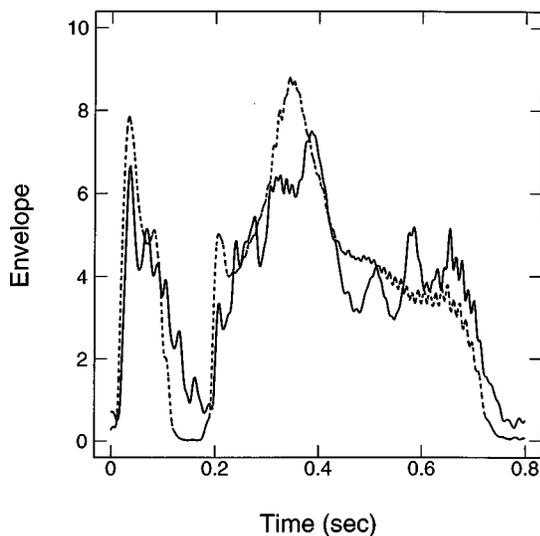


FIG. 5. Segments of the envelope of the octave band centered at 250 Hz of conversational speech before processing (dotted line) and after reverberation with RT=0.6 s (solid line).

be greater than that for noisy speech (Fig. 2).

Detailed examination of the waveforms of degraded speech reveals at least some of the causes for the observed deviations from expectations that are based on acoustic theory and experience with envelope spectra measured using modulated-noise probes. In the case of noisy speech, the added noise contributes its own modulations to the envelope of the summed waveform. The noise modulation is constant on a per Hertz basis, resulting in one-third octave noise envelope spectra that increase with modulation frequency. However, the mean amplitude of the noise envelope fluctuations is inversely proportional to the bandwidth of the octave filter applied to the noisy speech signal, so the slopes of the noise envelope spectra also vary as a function of octave bandwidth.³ The one-third octave envelope spectra of octave bands of the added speech-shaped noise (Fig. 2) are close to the envelope spectra of the noisy speech at the highest modulation frequencies in the 250- and 1000-Hz octave bands.

A different phenomenon appears to be operative in the case of reverberant speech. At low frequencies, the speech waveform is quasiperiodic rather than noiselike, reflecting

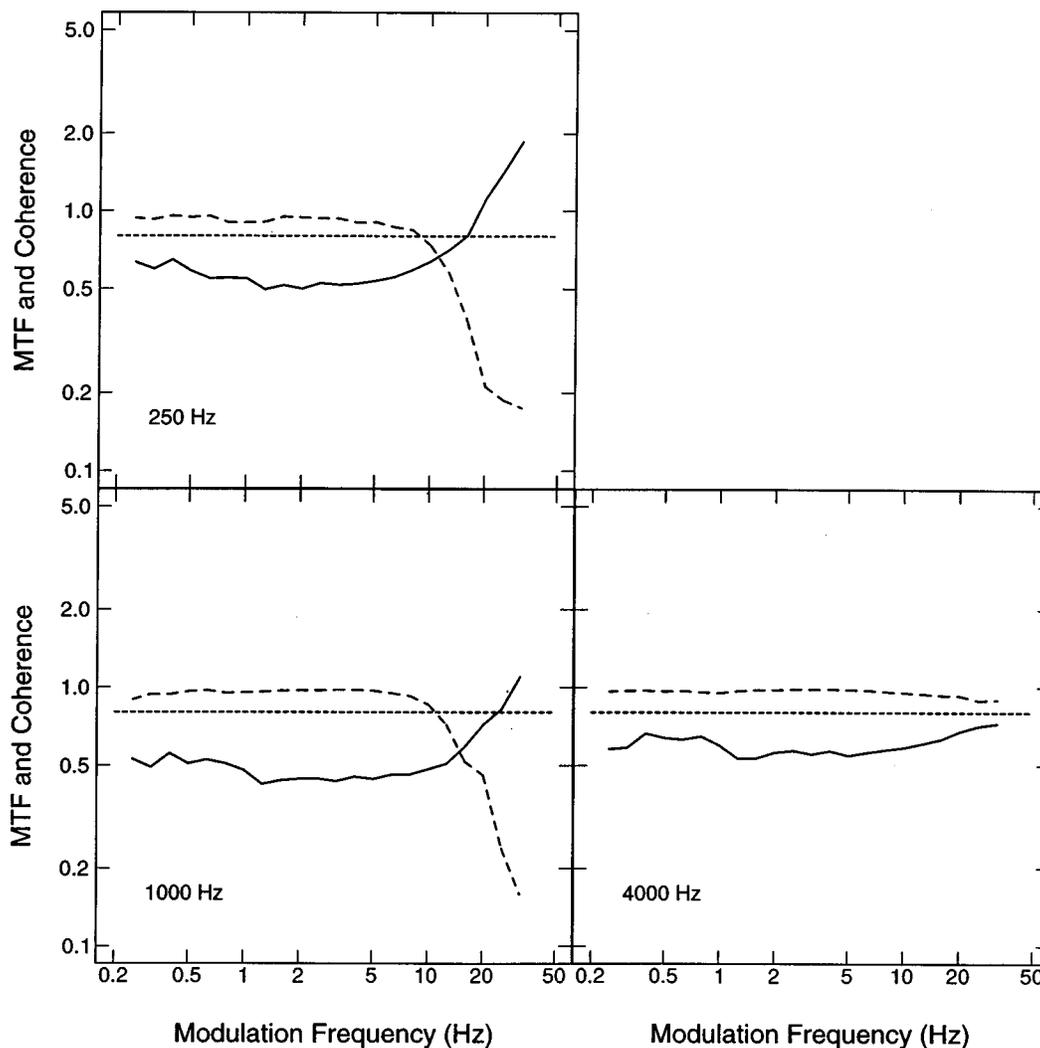


FIG. 6. The MTF (solid lines) and coherence (dashed lines) computed on a one-third octave basis for undegraded conversational speech and the same speech degraded by additive speech-spectrum noise (0 dB SNR). In each panel the dotted line denotes a coherence value of 0.8. The modulation frequency at which the dotted and dashed lines intersect is the frequency at which the MTF is truncated.

the presence of vowel sounds. Reverberation causes both constructive and destructive interference, depending on the relation between the arrival times of reflections and the fundamental period of the voice. As the voice pitch changes, this interference can modulate the originally smooth envelopes of octave bands of speech (Fig. 5).

III. APPROACH FOR SPEECH-BASED STI

Straightforward use of the modulation spectra of speech that has been degraded by noise and/or reverberation cannot form a reliable basis for estimating the STI because spurious modulation components are introduced which disrupt the perception of relevant speech modulations. To be qualitatively consistent with the theoretical computations that have been shown to correlate well with measures of speech intelligibility, the effects of artifacts such as those in Figs. 2–4 must be suppressed when computing the STI from speech waveforms.

One way to eliminate the deviations of speech-based MTFs from theoretical MTFs at the higher modulation frequencies is to set an upper limit on the maximum modulation

frequency included in the speech-based STI computation for both noise and reverberant degradations. Although a conservatively low bound such as the 12.5 Hz used by Houtgast and Steeneken (1985) could be used in many conditions, this might underestimate the reduction in intelligibility associated with mild reverberation: truncation might occur at a lower frequency than the low-pass cutoff frequency associated with the reverberation.

We chose to derive a measure from the envelope signals themselves that would indicate automatically which portions of the envelope spectra should be included in the STI computation. The coherence between envelopes of clean and degraded speech is typically close to unity for modulation frequencies unaffected by artifacts, but decreases rapidly when artifacts are evident. The coherence computes the ratio of the squared magnitude of the cross-power spectral density to the product of the auto-power spectral densities for the clean and degraded speech envelopes, respectively, based on concatenated envelopes with a 32 768-sample window and with a 75% overlap between windows. The power spectra of 21 conversational or 45 clear windows were averaged for each

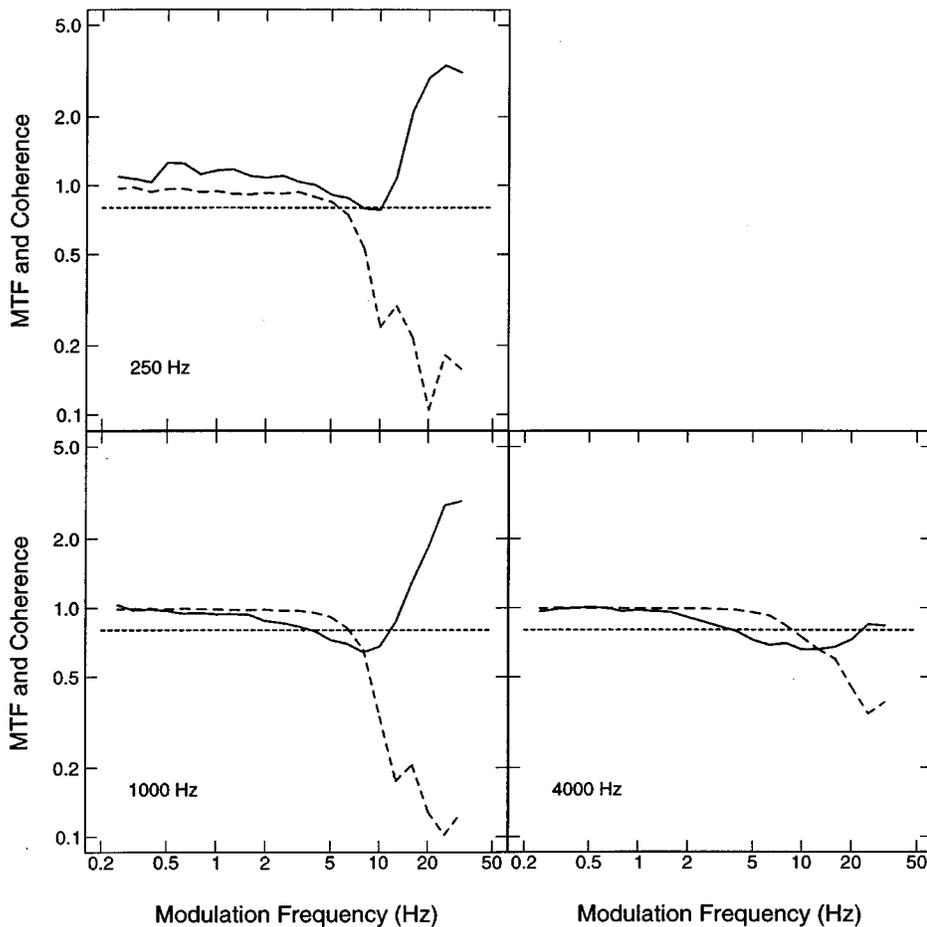


FIG. 7. The MTF (solid lines) and coherence (dashed lines) computed on a one-third octave basis for undegraded conversational speech and the same speech degraded by reverberation (RT=06 s). In each panel the dotted line denotes a coherence value of 0.8.

condition. The dependence of conversational speech coherence and MTF on modulation frequency is shown in Fig. 6 for the additive noise condition of Fig. 2, in Fig. 7 for the reverberant condition of Fig. 3, and in Fig. 8 for the noise-reverberation condition of Fig. 4. The coherence plots for clear speech were very similar to those for conversational speech.

We established an upper bound on the modulation frequencies that would be included in the STI computation: that modulation frequency at which the coherence falls below 0.8. In computing the STI, the speech-based MTF is truncated to include only frequencies below this bound. Visual examination of plots like Figs. 6–8 indicated that 0.8 was a reasonable compromise to try to maximize the modulation frequency range while minimizing the contribution of artifacts. Truncated speech-based MTFs are compared to MTFs computed on the basis of acoustic theory in Fig. 9. For noisy speech, the theoretical MTF is depicted for perfectly speech-shaped noise at 0 dB SNR. The slight shifts of the reported curves above and below the theoretical prediction of 0.5 are due to SNR differences in individual octave bands since the noise spectrum was fit to the combined spectrum of clear and conversational speech. For reverberant speech, the theoret-

ical MTF was derived from the room impulse response (RT=0.6 s) using the method of Schroeder (1981) wherein the room impulse response is squared and Fourier transformed to give the theoretical reverberant MTF. The theoretical MTF for noise plus reverberant speech was obtained by multiplying the noise MTF by the reverberant MTF (Houtgast and Steeneken, 1985). As indicated previously, the theoretical and speech-based MTFs agree relatively well for highest octave bands for each degradation, but relatively poorly for the lowest octave bands. However, in all cases, the truncated MTF agrees fairly well with the theoretical MTF, an indication that the bound frequency selection process successfully eliminates most artifacts.

The STIs were computed from truncated MTFs for the two speaking styles and eight distortion conditions studied by Payton *et al.* (1994).⁴ In this study a continuous flat-spectrum noise was added to the processed speech materials (at -4 dB SNR) to simulate the effects of a flat hearing loss for the normal-hearing subjects. Since samples of noise used to simulate the hearing loss were not recorded, its effects on envelope spectra could not be computed directly. Instead, the SNR in each octave band was converted to an MTF value

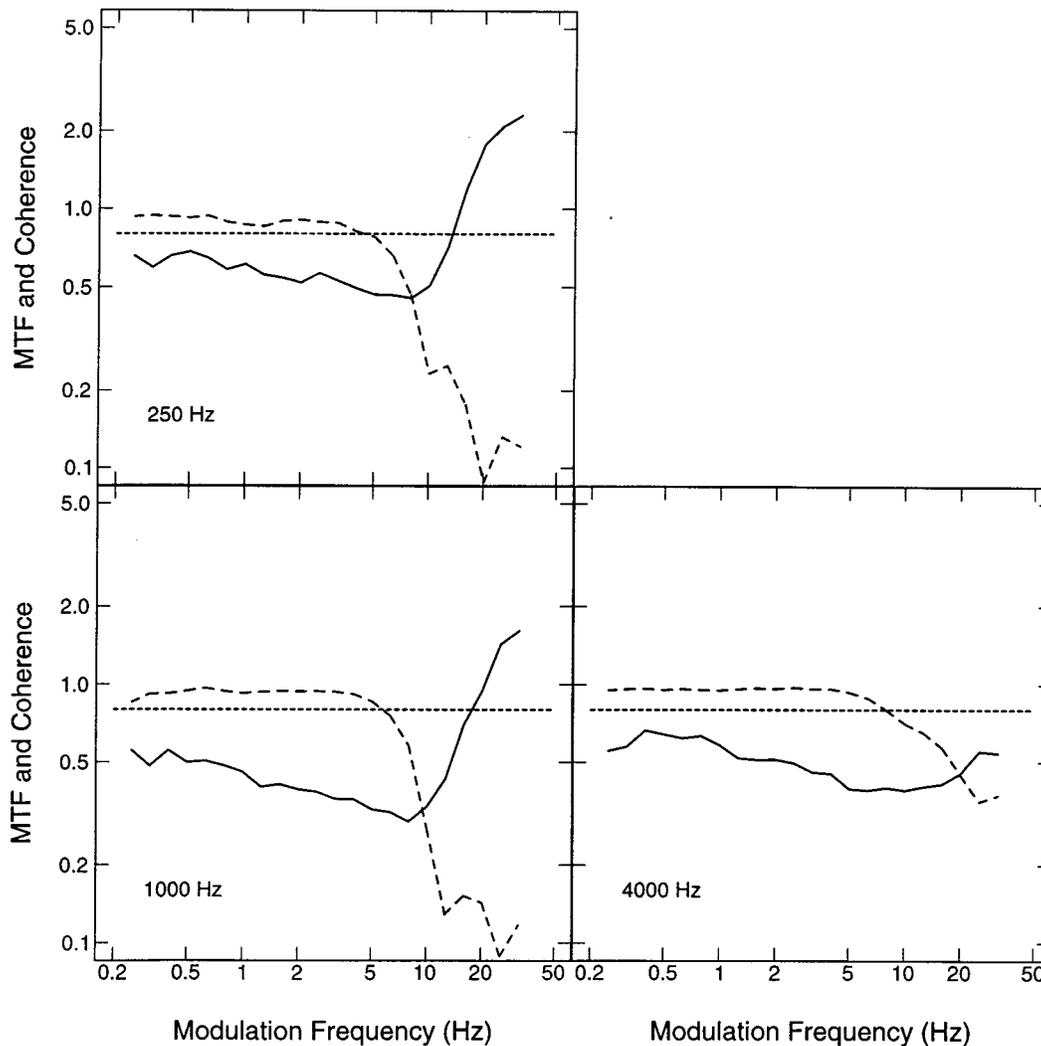


FIG. 8. The MTF (solid lines) and coherence (dashed lines) computed on a one-third octave basis for undegraded conversational speech and the same speech degraded by additive speech spectrum noise at 0 dB SNR and reverberation with RT=06 s (dotted line). In each panel the dotted line denotes a coherence value of 0.8.

(constant as a function of modulation frequency) according to Houtgast and Steeneken's (1985) formula:

$$m_S(i) = [1 + 10^{-\text{SNR}(i)/10}]^{-1}, \quad (1)$$

where i is the octave band index. The truncated speech-based MTFs were then multiplied by $m_S(i)$, clipped to range between 0 and 1, and converted to an apparent SNR value:

$$\text{SNR}(i,j) = 10 \log \frac{m(i,j)}{1 - m(i,j)}, \quad (2)$$

where j is the modulation frequency index and $m(i,j)$ is the clipped MTF value for band j of octave i . The apparent SNR values were clipped to range between +15 and -15 dB and averaged over j to give a single apparent SNR per octave band:

$$\text{SNR}(i) = \frac{1}{N_i} \sum_{j=1}^{N_i} C[\text{SNR}(i,j)], \quad (3)$$

where $C[\text{SNR}(i,j)]$ is the clipped apparent signal-to-noise ratio for the j th one-third octave band of the envelope signal for the i th octave band, and N_i is the number of one-third

octave modulation indices included in the average. A weighted average of the seven octave-band SNRs is computed

$$\text{SNR}' = \sum_{i=1}^7 \alpha_i w_i \text{SNR}(i), \quad (4)$$

where w_i is the importance weighting factor for the i th octave band suggested by Houtgast and Steeneken (1985). The term

$$\alpha_i = 7 \frac{N_i}{\sum_{k=1}^7 N_k} \quad (5)$$

is a new factor that reduces the relative contributions of octave bands in which modulation spectra are dominated by artifacts and increases the relative contributions of octave bands without artifacts. The α 's always sum to 7. The net influence of α is to weight the higher octave bands more heavily than originally proposed by Houtgast and Steeneken since they are affected least by artifacts. Finally, the STI is computed as

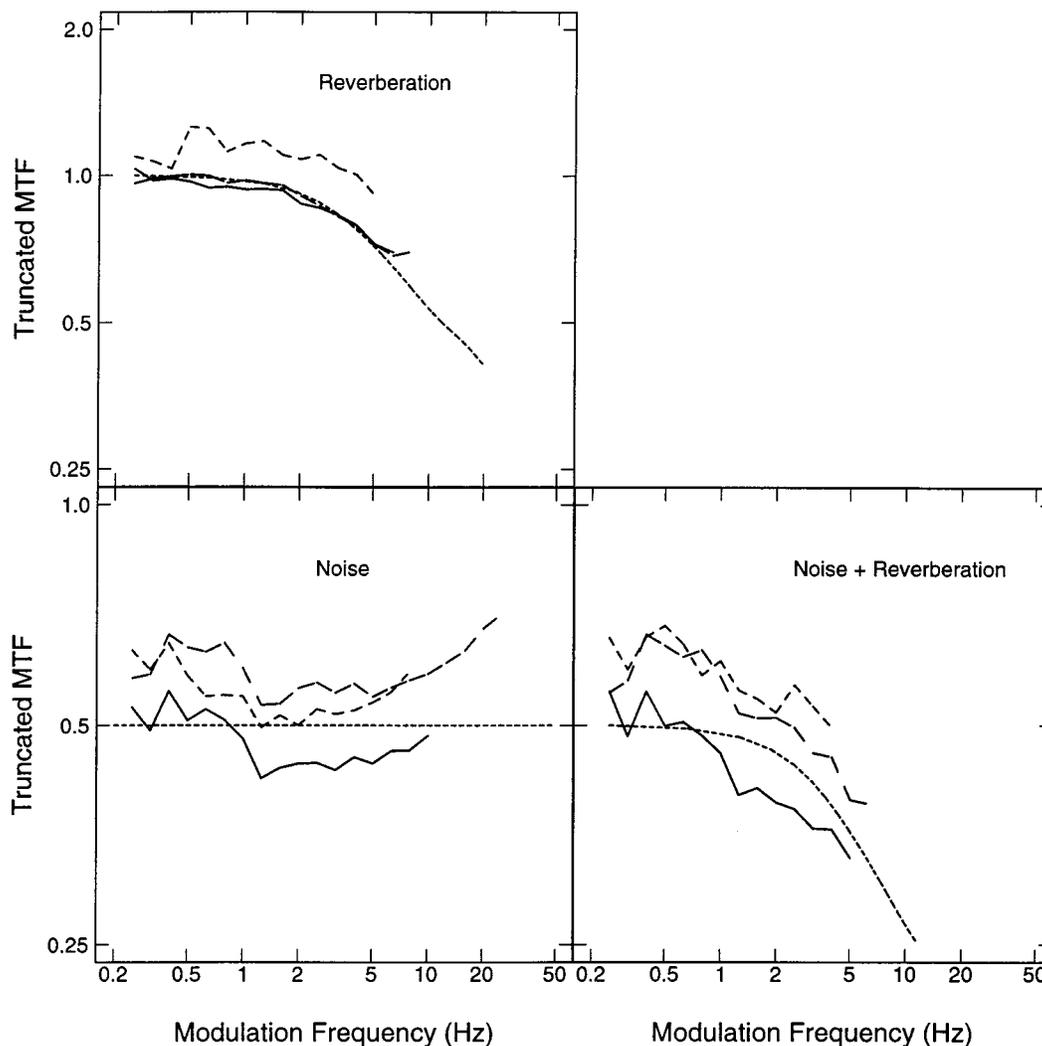


FIG. 9. Truncated modulation transfer functions determined from speech waveforms for the octave bands of speech centered at 250 Hz (short dashed lines), 1000 Hz (solid lines), and 4000 Hz (long dashed lines). Functions for conversational speech degraded by reverberation with RT=0.6 s are shown in the top panel, by additive speech-spectrum noise at 0 dB SNR are shown in the bottom-left panel, and for the noisy speech further degraded by reverberation in the bottom-right panel. In each panel the dotted line shows the modulation transfer function determined from acoustic theory.

$$STI = \frac{SNR' + 15}{30}. \quad (6)$$

IV. RESULTS

To evaluate the effect of modifying the STI computation on predictions of speech intelligibility, we examined the relationship between the theoretical and speech-based STI values and previously obtained measurements of speech intelligibility in the degraded conditions.⁵

Table I and Figs. 10 and 11 compare the theoretical and speech-based STI values for each of the presentation conditions, together with the average intelligibility scores for ten normal-hearing listeners (theoretical STI values and intelligibility scores are from Payton *et al.*, 1994). Although speech-based STI values generally are in close agreement with theoretical STI values for clear speech, they tend to be somewhat smaller for conversational speech (a larger difference than can be accounted for by our method of simulated impairment noise inclusion). For both methods of computing STI values, and both speaking styles, the milder acoustic

degradations produce a wider range of STI values and smaller range of intelligibility scores than the more severe degradations. Figure 12 indicates that, for each speaking style, the dependence of intelligibility on STI is essentially the same for the speech-based and theoretical STI values.

The theoretical method of computing the STI does not take into account differences in envelope modulation spectra associated with differences in speaking style. As a result, theoretical STI values are relatively independent of speaking-style, differing by at most 0.06 for each degradation.⁶ By contrast, the speech-based STI values for the two speaking styles typically differ by larger amounts (up to 0.11), particularly in the most degraded listening conditions. Moreover, the speech-based STI values are lower than the theoretical STI values in these conditions for conversational speech, but not for clear speech. As a result, the speech-based STI accounts for a larger part of the intelligibility difference associated with speaking style than the theoretical STI. However, this improvement is still small compared to the large differences in intelligibility scores associated with speaking style. Moreover, since the speech-

TABLE I. Environment, speaking style, theoretical and speech-based STIs and intelligibility scores.

Environment	Speaking style	Theoretical STI	Speech-based STI	Word % score
ANEC	CL	0.84	0.84	97
(Anechoic)	CO	0.78	0.78	81
LIVR	CL	0.79	0.78	93
(RT=0.18 s)	CO	0.73	0.69	78
SNR1	CL	0.75	0.73	93
(SNR=9.4 dB)	CO	0.70	0.67	73
NILV	CL	0.71	0.70	90
(RT=0.18 s and SNR=9.4 dB)	CO	0.66	0.62	69
CONF	CL	0.69	0.76	91
(RT=0.6 s)	CO	0.65	0.67	71
SNR2	CL	0.65	0.63	89
(SNR=5.3 dB)	CO	0.61	0.58	71
SNR3	CL	0.51	0.50	79
(SNR=0 dB)	CO	0.47	0.43	51
N3LV	CL	0.50	0.50	73
(RT=0.18 s and SNR=0 dB)	CO	0.46	0.40	43
N3CF	CL	0.46	0.48	61
(RT=0.6 s and SNR=0 dB)	CO	0.42	0.39	27

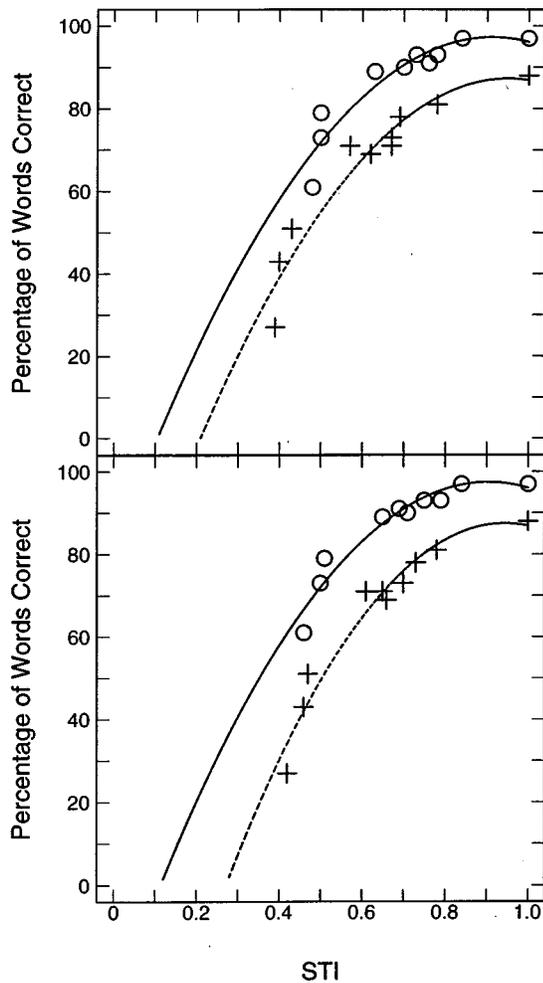


FIG. 10. Relation between speech intelligibility (percentage of key words in nonsense sentences recognized correctly) and STI values computed by speech-based (top panel) and theoretical methods (bottom panel) for clear (circles) and conversational speech (crosses). The solid and broken lines are best second-order fits to the data.

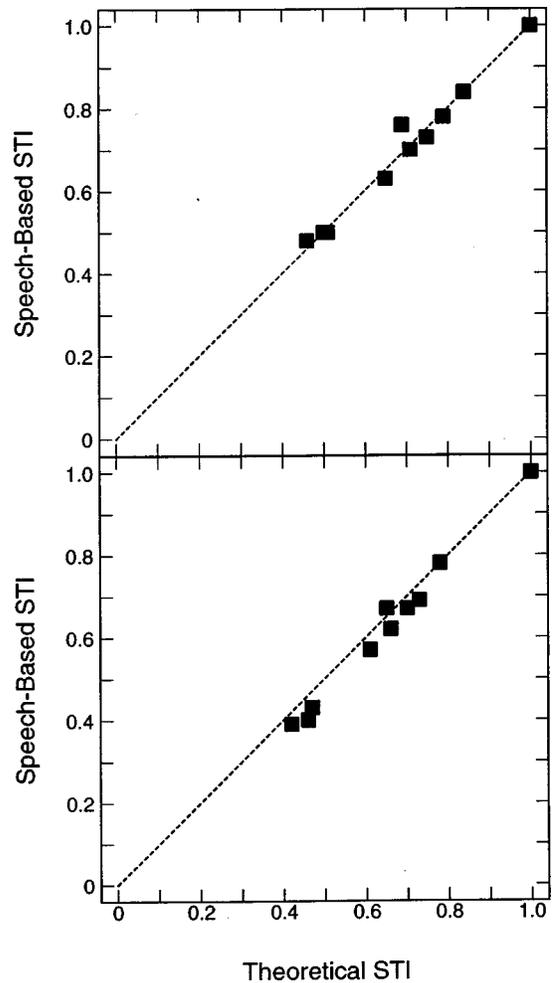


FIG. 11. Relation between STI values computed by speech-based and theoretical methods for clear (top panel) and conversational speech (bottom panel).

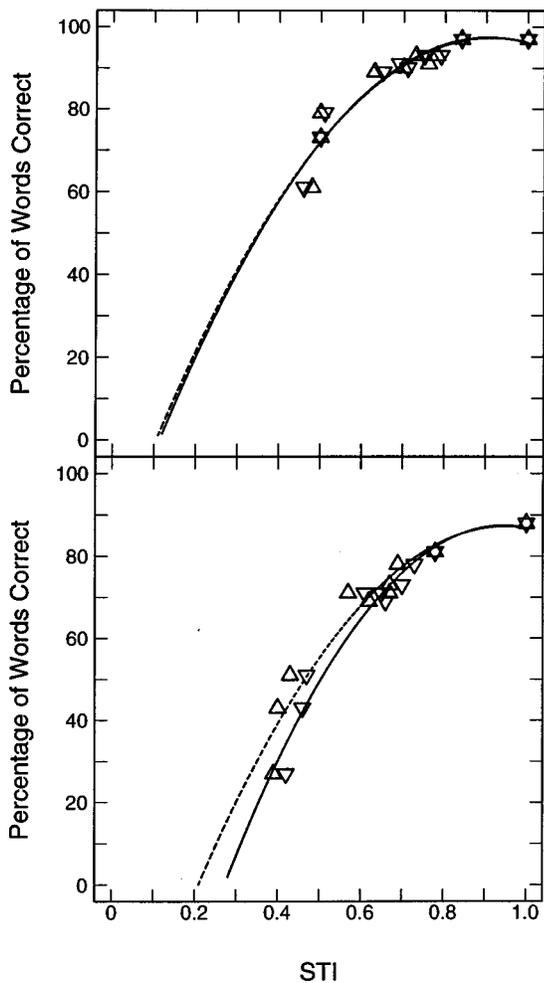


FIG. 12. Relation between speech intelligibility (percentage of key words in nonsense sentences recognized correctly) and STI values computed by speech-based (Δ) and theoretical (∇) methods for clear (top panel) and conversational speech (bottom panel). The solid and broken lines are best second-order fits to the data.

based approach simply averages all estimated MTF values within the coherence bounds, differences in the derived STI values are unlikely to reflect differences between the modulation spectra of clear and conversational speech.

V. SUMMARY AND CONCLUSION

When speech is used as a probe stimulus, substantial artifacts can be observed in envelope modulation spectra that can affect the computation of the STI. In low-frequency octave bands of noisy and/or reverberant speech, these spectra are distorted at the higher modulation frequencies in ways that increase the apparent modulation values. This increase is contrary to theoretical expectations and can increase the apparent STI, thus underestimating the effect of the degradation on speech intelligibility.

There are two substantive differences between our speech-based STI and the traditional STI. First, the traditional MTF is obtained from modulated noise and always spans the same range of modulation frequencies. The proposed method of computing the STI uses speech as a probe stimulus and puts a variable upper bound on the modulation frequencies included in the STI computation based on the

coherence between the envelopes of degraded and undegraded speech. A coherence value of 0.8 has been found to establish bounds that allow much of the envelope spectra to be used in the computation while eliminating many of the effects of artifacts. Second, the proposed method introduces a new weighting term α_i that weights the octave bands with highest modulation cutoff most heavily.

Speech-based STI values have been shown to be reliable predictors of intelligibility when computed using the proposed truncation scheme, at least for the range of acoustic degradations considered by Payton *et al.* (1994). In addition, the STI values computed using the proposed method seem to capture some of the physical differences in speech waveforms associated with differences in speaking style that produce differences in speech intelligibility. Further research is needed to determine to what extent these differences are related to differences in modulation spectra. A valid speech-based STI may also be applicable to signal distortions that are not amenable to traditional techniques for estimating the STI, as in nonlinear signal processing conditions such as amplitude compression.

ACKNOWLEDGMENTS

The authors wish to thank B. Kollmeier, an anonymous reviewer, and J. Hieronymus for their assistance in improving this manuscript. Support for the research was provided in part by Research Grant No. 5 RO1 DC 00117 from the National Institute on Deafness and Other Communication Disorders.

¹Nonsense sentences are grammatically correct while being semantically meaningless, e.g., "A right cane could guard an edge."

²The durations of these intervals was uniformly and randomly distributed over the range 0.2–0.9 s for conversational speech and 0.3–1.5 s for clear speech.

³Similar fluctuations are present in traditional STI measurements when modulated noise is used as a probe, but they have only minor effects on envelope spectra because the explicit modulation of the probe noise is at least an order of magnitude larger than any inherent modulation.

⁴STI values for the ANEC (anechoic) condition are identical to those computed by Payton *et al.* because no degradations were introduced.

⁵Payton *et al.* (1994) accounted for the effects of the flat-spectrum noise used to simulate hearing loss by using the greater of the flat-spectrum and speech-spectrum noise levels in each octave band when computing $SNR(i)$ rather than multiplying modulation transfer functions. This difference results in theoretical STI values that are slightly larger than speech-based STI values for both speaking styles.

⁶This difference results from small differences in the long-term speech spectra of clear and conversational speech.

Drullman, R., Festen, J. M., and Plomp, R. (1994). "Effect of reducing slow temporal modulations on speech reception," *J. Acoust. Soc. Am.* **95**, 2670–2680.

Hohmann, V., and Kollmeier, B. (1995). "The effect of multichannel dynamic compression on speech intelligibility," *J. Acoust. Soc. Am.* **97**, 1191–1195.

Holube, I., and Kollmeier, B. (1996). "Speech intelligibility prediction in hearing-impaired listeners based on a psychoacoustically motivated perception model," *J. Acoust. Soc. Am.* **100**, 1703–1716.

Houtgast, T., and Steeneken, H. (1972). "Envelope spectrum and intelligibility of speech in enclosures," in *IEEE-AFCRL Conference on Speech Communications and Processing* (IEEE, Newton, MA), pp. 392–395.

Houtgast, T., and Steeneken, H. (1973). "The modulation transfer function in room acoustics as a predictor of speech intelligibility," *Acustica* **28**, 66–73.

- Houtgast, T., and Steeneken, H. (1985). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," *J. Acoust. Soc. Am.* **77**, 1069–1077.
- Houtgast, T., Steeneken, H., and Plomp, R. (1980). "Predicting speech intelligibility in rooms from the modulation transfer function. I. general room acoustics," *Acustica* **46**(1), 60–72.
- Humes, L. E., Dirks, D. D., Bell, T. S., Ahlstrom, C., and Kincaid, G. E. (1986). "Application of the Articulation Index and the Speech Transmission Index to the recognition of speech by normal-hearing and hearing-impaired listeners," *J. Speech Hear. Res.* **29**, 447–462.
- Ludvigsen, C. (1993). "The use of objective measures to predict the intelligibility of hearing aid processed speech," in *Recent Developments in Hearing Instrument Technology*, edited by J. Beilin and G. R. Jensen (Scanticon, Kolding, Denmark), pp. 81–94.
- Ludvigsen, C., Elberling, C., and Keidser, G. (1993). "Evaluation of a noise reduction method-comparison between observed scores and scores predicted from STI," *Scand. Audiol. Suppl.* **38**, 50–55.
- Ludvigsen, C., Elberling, C., Keidser, G., and Poulsen, T. (1990). "Prediction of intelligibility of non-linearly processed speech," *Acta Oto-Laryngol. Suppl.* **469**, 190–195.
- Payton, K. L., Uchanski, R. M., and Braida, L. D. (1993). "Computation of modulation spectra for the Speech Transmission Index using real speech," in *Proceedings of the 1993 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (IEEE Signal Processing Society), Paper 5.4.
- Payton, K. L., Uchanski, R. M., and Braida, L. D. (1994). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," *J. Acoust. Soc. Am.* **95**, 1581–1592.
- Peterson, P. M. (1986). "Simulating the response of multiple microphones to a single acoustic source in a reverberant room," *J. Acoust. Soc. Am.* **80**, 1527–1529.
- Picheny, M., Durlach, N., and Braida, L. (1986). "Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech," *J. Speech Hear. Res.* **29**(4), 434–446.
- Plomp, R. (1988). "The negative effect of amplitude compression in multichannel hearing aids in light of the modulation transfer function," *J. Acoust. Soc. Am.* **83**, 2322–2327.
- Schroeder, M. (1981). "Modulation transfer functions: Definition and measurement," *Acustica* **49**, 179–182.
- Steeneken, H., and Houtgast, T. (1980). "A physical method for measuring speech-transmission quality," *J. Acoust. Soc. Am.* **67**, 318–326.
- Steeneken, H., and Houtgast, T. (1983). "The temporal envelope spectrum of speech and its significance in room acoustics," in *Proceedings of the Eleventh International Congress on Acoustics* (GALF, Paris), Vol. 7, pp. 85–88.
- Villchur, E. (1989). "Comments on 'The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation transfer function,'" *J. Acoust. Soc. Am.* **86**, 425–427.

Training American listeners to perceive Mandarin tones

Yue Wang^{a)} and Michelle M. Spence

Cornell Phonetics Laboratory, Cornell University, Ithaca, New York 14853

Allard Jongman and Joan A. Sereno

Linguistics Department, University of Kansas, Lawrence, Kansas 66045

(Received 20 April 1999; revised 19 July 1999; accepted 20 July 1999)

Auditory training has been shown to be effective in the identification of non-native segmental distinctions. In this study, it was investigated whether such training is applicable to the acquisition of non-native suprasegmental contrasts, i.e., Mandarin tones. Using the high-variability paradigm, eight American learners of Mandarin were trained in eight sessions during the course of two weeks to identify the four tones in natural words produced by native Mandarin talkers. The trainees' identification accuracy revealed an average 21% increase from the pretest to the post-test, and the improvement gained in training was generalized to new stimuli (18% increase) and to new talkers and stimuli (25% increase). Moreover, the six-month retention test showed that the improvement was retained long after training by an average 21% increase from the pretest. The results are discussed in terms of non-native suprasegmental perceptual modification, and the analogies between L2 acquisition processes at the segmental and suprasegmental levels. © 1999 Acoustical Society of America. [S0001-4966(99)04811-0]

PACS numbers: 43.71.Hw, 43.71.Es [JMH]

INTRODUCTION

It is commonly stated that Mandarin tones are difficult for American learners to acquire (e.g., Kiriloff, 1969; Bluhme and Burr, 1971; Shen, 1989), since English and Mandarin differ in their pitch patterns, distributions, and functions (Chen, 1974; White, 1981). In the present study, American listeners were trained to identify the four Mandarin tones, using an auditory training procedure which has been shown to be effective in helping learners acquire non-native segmental contrasts in a comparatively short period of time.

A. Auditory training

Research in the domain of second language (L2) acquisition has generally found that adults are inferior to children in the ability to perceive and produce foreign speech sounds, manifested by the commonly known "adult foreign accent." The belief in the possibility that children enjoy an innate ability to acquire languages more easily and accurately than adults leads to the Critical Period Hypothesis (CPH), stating that cerebral lateralization occurs after puberty, accompanied by the loss of neurological plasticity of the brain, resulting in a reduction in language learning ability (Lenneberg, 1967).

An alternative account of foreign accent is the phonologically based argument that foreign accent is not caused by the completion of cerebral lateralization, but is rather the result of the interaction between L2 learners' two phonetic systems (e.g., Flege, 1995; Best, 1995). In this view, adult L2 learners differ from children acquiring their first language (L1) in that the former perceive and produce L2 sounds with reference to the linguistic categories of their existing native language system. Thus the influence of the adults' firmly established L1 phonetic system is believed to be responsible

for "foreign accent." However, unlike the CPH statement of a complete diminution of speech learning ability at puberty, the phonologically based argument is that the decline in human vocal learning ability with age does not apply to all L2 sounds. It is assumed that the degree of approximation to L2 sounds depends on learners' "perceived phonetic similarity" of L2 sounds to their L1 phonetic categories. Empirical research has revealed that, with sufficient experience and exposure, adult L2 learners can authentically perceive or produce novel L2 phones which are judged to have no L1 phonemic counterparts, although it is still difficult for them to form separate phonetic categories for those L2 sounds that are similar to L1 counterparts but realized in a phonetically different manner (Flege, 1987; Best *et al.*, 1988).

The evidence that learners can improve their L2 pronunciation at least for some target language sounds suggests adult perceptual mechanisms have more plasticity than was previously recognized. Therefore, researchers have attempted to train listeners to perceive non-native sounds in a linguistically meaningful manner, based on the assumption that the perceptual system of mature adults can be modified. The goal of these auditory training studies is, by using relatively simple laboratory procedures, to help listeners create a new phonetic category that is usable in various phonetic contexts and can be retained in long-term memory.

An early attempt of this approach was to train American listeners to perceive three-way (i.e., voiced, voiceless unaspirated, voiceless aspirated) voice onset time (VOT) distinctions (e.g., Pisoni *et al.*, 1982; McClaskey *et al.*, 1983), since English does not phonemically distinguish voiced and voiceless unaspirated stops. There were also experiments that trained French listeners to identify the English / θ - δ / contrast, which is absent in French (e.g., Jamieson and Morosan, 1986, 1989). Most recent training studies have concentrated on training Japanese listeners to identify English / r / and / l /

^{a)}Electronic mail: yw36@cornell.edu

(e.g., Strange and Dittmann, 1984; Logan *et al.*, 1991; Lively *et al.*, 1993; Lively *et al.*, 1994; Bradlow *et al.*, 1997).

Summing up the results of these training studies, first and most importantly, the identification of non-native speech contrasts generally improved after training. For instance, Jamieson and Morosan (1986) reported that the French trainees' average percentage of correct identification for natural stimuli (containing /ə/ or /ɔ̃/) improved from the pretest (68% correct responses) to the post-test (79% correct responses) by 11%. Logan *et al.*'s (1991) study on training Japanese listeners to perceive English /r/ and /l/ also showed a significant increase of 8% from pretest (78%) to post-test (86%). Similarly, there was a 16% increase (from 65% to 81%) in the Japanese trainees' /r-l/ identification accuracy in Bradlow *et al.* (1997).

In addition, researchers have also found an effect of training with regard to generalization and long-term retention. First, experience gained from training on one phonetic category (e.g., VOT contrast for labial stops) can be transferred to another phonetic category (e.g., VOT for alveolar stops) without additional training (McClaskey *et al.*, 1983). Second, generalization can extend to novel words and talkers that are not used in training (Lively *et al.*, 1993). Third, contrasts learned can be maintained long (i.e., three to six months) after training (Lively *et al.*, 1994). And finally, contrasts gained perceptually can be transferred to production without additional training (Rochet, 1995; Bradlow *et al.*, 1997).

Concerning methodological issues, the previous studies have agreed that training should be designed to ensure the formation of a robust phonetic category, since the ultimate goal is to facilitate the development of a new phonemic category that is usable among a variety of sources of variability (Logan and Pruitt, 1995). For example, Jamieson and Morosan (1986, 1989) designed the fading technique (i.e., training is not only on the prototypical stimuli, but also on a variety of exemplars within the category) in an attempt to extend generalization from synthetic to natural stimuli. While Strange and Dittmann (1984) report no significant effect of discrimination training using synthetic stimuli in only one phonetic environment, Logan *et al.* (1991) demonstrated that a high-variability training paradigm (i.e., identification of natural stimuli in various phonetic contexts and spoken by various talkers) encouraged long-term modification of listeners' phonetic perception.

B. Mandarin tones

Mandarin phonemically distinguishes four tones, with Tone 1 having high-level pitch, Tone 2 high-rising pitch, Tone 3 low-dipping pitch, and Tone 4 high-falling pitch (Chao, 1948). The prosodic features of tones are manifested physically by different fundamental frequency (F_0) values, as shown in Fig. 1. Moreover, the F_0 pattern for particular tones varies as a function of vowel (Howie, 1976). In addition, the intrinsic duration differs for the four tones, the longest being Tone 3, and the shortest being Tone 4 (Lin, 1965). Intrinsic amplitude has been found to vary among the four tones as well, with Tone 3 having the lowest, and Tone 4 the highest amplitude (Chuang *et al.*, 1972).

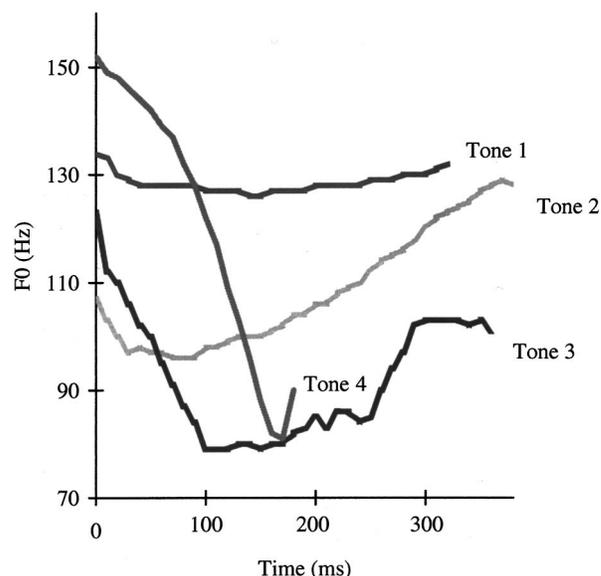


FIG. 1. F_0 contours for the four Mandarin tones, each combined with the syllable *fa*, produced by a male native speaker of Mandarin.

Studies in the perceptual domain have shown that the above acoustic cues are functionally integrated in the identification of Mandarin tones by native listeners. For example, perception tests using synthetic F_0 contours and multidimensional scaling studies have demonstrated the two dimensions of F_0 height and contour as fundamental perceptual cues of Mandarin tones, of which listeners seem to attach more importance to the "contour" than "height" dimensions (Gandour, 1984; Massaro *et al.*, 1985). F_0 contour as a perceptual cue has been further investigated in terms of F_0 turning point, i.e., the point at which the direction of the F_0 contour changes from falling to rising, the results of which showed that the timing of F_0 turning point constitutes a salient perceptual cue for differentiating Tone 2 from Tone 3 (Shen and Lin, 1991; Moore and Jongman, 1997), and Tone 3 from Tone 4 (Gårding *et al.*, 1986). In addition, duration has also been shown to affect tonal perception. For instance, Blicher *et al.* (1990) reported that systematic lengthening of the vowel shifted the labeling boundary in the direction of the Tone 2 exemplar, thus producing more Tone 3 responses. Moreover, native Mandarin listeners have been found to refer to extrinsic F_0 (corresponding to speaker identity) as a frame of reference for tone perception; that is, they perceive tones by normalizing for speaker F_0 range (Moore and Jongman, 1997).

Perception studies on Mandarin tones have also been conducted cross-linguistically to examine if and how non-tonal listeners distinguish themselves from the Mandarin listeners by their patterns of perceptual processing of the dimensions of F_0 . For example, by comparing tone perceptual patterns of native English and Mandarin (as well as Cantonese, Taiwanese, and Thai) listeners, Gandour (1983) found that native English listeners attached more importance to the height, and less to the direction dimension, than did listeners from most of the tone languages. He argued that since English has no contrastive tones, contour or otherwise, English listeners directed their attention almost exclusively

to the F_0 height of the stimuli. Addressing the same question, Leather (1987) examined the identification of Mandarin Tone 1 and Tone 2 (in a synthetic Tone 1-2 continuum) by native listeners of English and Dutch (both nontonal), as compared to that by Mandarin listeners. The result of a greater spread in location of the category crossover among the Dutch and English, as opposed to the Chinese, reflects linguistically inappropriate perceptual weighting of the parameters of F_0 contour by the phonetically unskilled non-natives. Stagra and Downs (1993) examined the differential sensitivity for frequency among Mandarin and English listeners from a psychoacoustic perspective. They found that Mandarin listeners had poorer differential sensitivity than English listeners because the former had learned to categorize sounds of similar frequency together to facilitate their perception of tones. Taken together, these cross-linguistic studies suggest that linguistic experience plays an important role in tone perception.

For adult nontonal speakers learning Mandarin as an L2, tones have presented great difficulty (e.g., Kiriloff, 1969; Bluhme and Burr, 1971; Shen, 1989). For native speakers acquiring Mandarin as L1, tonal pattern is an integral part of each word they learn, but such functional association between segmental structure and F_0 contour is nonexistent, for example, in American learners' linguistic behavior. Therefore, the source of difficulty in learning tones has generally been attributed to interference from English suprasegmental features. Knowledge of the function of pitch in the English stress and intonation systems was found to highly influence American listeners' perception of Mandarin tones (White, 1981; Broselow *et al.*, 1987; Chen, 1997). For example, White (1981) claimed that English listeners will perceive the Mandarin high tones as stressed and the low Tone 3 as unstressed, despite the fact that in Mandarin, the stress on a syllable is mainly realized by duration and amplitude rather than F_0 . Given her observations that Tones 1 and 4 are more difficult to acquire, Shen (1989) argued that these two tones are more likely to be receptive to L1 interference since they are prosodically less marked than Tones 2 and 3. It should be noted that, although Tones 2 and 3 have been observed to be easier to learn than Tones 1 and 4, this tone pair is still the most confusing pair for English learners of Mandarin (Kiriloff, 1969).

C. The present study

As reviewed previously, research has shown substantial improvements (8%–16%), after simple phonetic laboratory training procedures, in the identification of segmental distinctions which are absent in the listeners' native language. However, little research has reported the application of such training procedures to the acquisition of non-native speech contrasts at the suprasegmental level. Since the acquisition of Mandarin tones has been found to be difficult for native nontonal learners, it provides an ideal case for the study of suprasegmental training. By training American listeners to perceive Mandarin tones, the goal of the present study was to examine whether auditory training, which has been shown to be effective at the segmental level, can be applied to the acquisition of non-native suprasegmental contrasts.

TABLE I. Characteristics of the trainees and the controls in terms of language background.

	Gender	Age	Mode of learning	Length of learning ^a	Class when training ^b	L2 experience
Trainee						
1	F	20	class ^c	7 months	yes	none
2	M	25	class	4 months	no	Spanish
3	F	19	class	7 months	yes	French
4	F	29	class	4 months	no	Cantonese
5	F	20	class	7 months	yes	French
6	M	24	intensive ^d	7 months	yes	none
7	F	19	class	7 months	yes	Cantonese
8	F	24	intensive	7 months	yes	none
Control						
1	M	21	class	7 months	yes	none
2	F	20	class	4 months	no	Cantonese
3	M	25	class	7 months	yes	Japanese
4	M	22	class	10 months	yes	Cantonese
5	M	23	class	7 months	yes	Spanish
6	M	21	class	7 months	yes	French
7	F	20	class	7 months	yes	Spanish
8	M	22	class	10 months	yes	none

^aLength of learning Mandarin as a foreign language.

^bWhether taking Mandarin course during the training period.

^cA first-year Chinese course (5 hours/week).

^dAn intensive Mandarin program (20 hours/week).

I. METHOD

The perceptual training program followed the high-variability procedure developed by Logan *et al.* (1991). That is, American listeners were trained to identify the four Mandarin tones appearing in a variety of phonetic contexts in natural words, produced by a variety of talkers. In order to assess the trainees' improvements, the program included a pretest before training, a post-test, two generalization tests, and a long-term retention test. Listeners' performance in the pretest and the post-test was compared to determine to what extent tone identification could be improved due to training. The two generalization tests were designed to examine if any improvement gained in training could be extended to novel stimuli (Generalization Test 1), and to novel talkers and stimuli (Generalization Test 2). The retention test was conducted six months after the training program to determine the long-term training effects.

A. Participants

Sixteen native speakers of American English without speech and hearing impairments participated in the study, with eight as trainees and eight as controls. All were paid for their participation. The trainees and controls are all students at Cornell University who have taken one or two semesters of Mandarin Chinese language courses. None of the trainees or controls has ever lived in a Mandarin-speaking environment, and most of them (except for the four who speak limited Cantonese) have no experience with a tone language prior to learning Mandarin. The characteristics of the trainees and controls are described in Table I.

Six native speakers of Mandarin Chinese participated voluntarily as talkers. One male speaker read the pretest and post-test stimuli, while four others (two males and two females) served as talkers during training. One of these male

speakers also read the novel stimuli for Generalization Test 1 (henceforth Gen 1). The sixth speaker was a female who provided the novel stimuli for Generalization Test 2 (henceforth Gen 2).

B. Stimuli

The stimuli are real monosyllabic Mandarin words presented in isolation. In order to ensure context variability, the stimuli were chosen to have combinations of various initial consonants and final vowels, and different syllabic structures (i.e., V, CV, CVNasal, VN, CGLideV, CGVN). A total of 400 different stimuli were selected: 100 items (25 for each tone) were used in the pre/post-test, 180 (45 for each tone) in training, 60 (15 for each tone) in Gen 1, and an additional 60 (15 for each tone) in Gen 2. The stimuli used in the retention test were the same as those in the post-test.

The stimuli were tape-recorded in a soundproof booth in the Cornell Phonetics Laboratory, using a cardioid microphone (Electrovoice RE20) and a cassette recorder (Carver TD-1700). They were then digitized at 11 kHz and low-pass filtered at 5 kHz using WAVES+/ESPS speech analysis software running on a SUN Sparc Station, after which they were transferred to a Swan 386/25 PC for the perceptual tests and training, using the BLISS software (Mertus, 1989).

Before the training program started, the intelligibility of the stimuli provided by the six talkers was assessed by one male and one female native speaker of Mandarin Chinese. Listeners indicated which tone they heard by pressing one of four response buttons. For both listeners, identification accuracy was 100% for all stimuli and all talkers.

C. Procedure

The training program consisted of a pretest phase, a training phase, and a post-test phase. Both the tests and training were conducted at the Cornell Phonetics Laboratory, where listeners were tested or trained in a sound-treated cubicle. Stimuli were presented binaurally at a comfortable sound level over Sony MDR-V6 headphones. Listeners were instructed to indicate their responses by pushing corresponding buttons representing each of the four tones. The four buttons were labeled from left to right by the numbers 1 to 4, as well as by the tonal diacritics (stylized pitch contours).

1. Pretest

Both the trainees and the controls took the pretest, in which they were presented with 100 randomized stimuli, with an inter-trial-interval of 3 s. The listeners were told to respond after each stimulus. They were encouraged to guess if unsure. No feedback was given at any time. The pretest lasted about 10 min, with no more than four listeners tested at any one time. All listeners were tested within a one-week period.

2. Training sessions

Immediately after the pretest, only the eight trainees participated in the two-week training program, consisting of eight sessions of 40 min each, during which the trainees were trained auditorily with the stimuli produced by four talkers.

The four tones were trained pairwise (i.e., Tones 1 and 2, Tones 1 and 3, Tones 1 and 4, Tones 2 and 3, Tones 2 and 4, and Tones 3 and 4). Pairwise presentation during training allowed for a systematic increase in difficulty of tone contrasts. The order of tone pair presentation was from easiest to most difficult, in accordance with the error analysis obtained from the trainees' pretest. That is, for each training set, the first session always started with Tones 1 and 3, followed by Tones 3 and 4, and Tones 1 and 4; the second session had Tones 1 and 2, Tones 2 and 4, and Tones 2 and 3 presented in succession. Three tone pairs were trained in each session, such that it took a training set of two successive sessions to complete one talker for a total of 180 stimuli. The order of presentation in terms of talker was counterbalanced for the eight trainees, but male and female talkers were always presented alternately.

During each session, the trainees' task was two-alternative forced-choice identification. They were to indicate (within 2 s) which tone of a certain tone pair they had heard by pressing the corresponding button. Immediate feedback was given after each stimulus, with a neutral voice indicating the correct response in English, and the talker repeating both tones in the tone pair. For example, for target stimulus *bei 3* (bearing Tone 3) in tone pair 3 and 4 training, stimulus presentation and feedback went as follows:

Talker: *bei 3*.

Trainee's response.

Neutral English voice: *That was Tone 3*.

Talker's repetition: *bei 3*.

Neutral English voice: *Tone 4 is:*

Talker: *bei 4*.

Thus the above block was considered a training trial, with an inter-trial-interval of 5 s. In addition, to focus the trainees' attention, each trial started with a 500-Hz pure tone. Each tone pair training (i.e., 30 trials) ended with a short break.

After each two consecutive sessions (i.e., a single talker), trainees were given a test of 60 selected trained stimuli produced by the same talker. No feedback was given. Since there were four different talkers for training, four assessments (training set 1–4) were administered.

3. Post-tests

Immediately after the training program, both the trained and the control listeners took the post-test, which was otherwise identical to the pretest, except that the stimuli were re-randomized. The listeners then took Gen 1, with 60 novel stimuli produced by one of the male talkers from training, and Gen 2, with an additional 60 novel stimuli produced by a new female talker; the procedures of both were comparable with the pretest. The post-tests were completed within a week's period.

4. Retention test

Six months after training, four trainees (Trainees 1, 3, 4, and 7 in Table I) and four controls (Controls 1, 2, 4, and 6 in Table I) were available for the long-term retention test, which involved the same stimuli and procedure as the post-

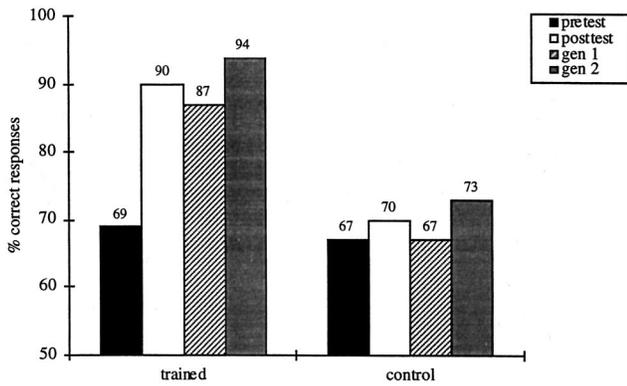


FIG. 2. Mean percent correct identification of the four Mandarin tones for trained ($n=8$) and control ($n=8$) subjects at pretest, post-test, generalization test 1 (Gen 1: old talker, new stimuli), and generalization test 2 (Gen 2: new talker, new stimuli).

test. All four trainees and two of the controls (Controls 2 and 4) had not been exposed to Mandarin for at least three months (summer break) before the retention test. The other two controls (Controls 1 and 6), however, had been in Taiwan for three months taking an intensive Mandarin course.

II. RESULTS

A. Overall improvement and generalization

Correct identification scores for the trained and the control groups at the pretest, post-test, generalization test 1, and generalization test 2 are displayed in Fig. 2. As shown in the left-hand bars, the trainees showed an improvement in their identification scores from pretest (69% correct identification) to post-test (90% correct identification), a substantial 21% increase in tone identification accuracy. Moreover, this increase in performance was also revealed in the two generalization tests (87% correct identification in Gen 1; and 94% correct identification in Gen 2), indicating tone contrasts gained in training were extended to novel talkers and stimuli.

In contrast, as the right-hand bars show, although the control listeners started at approximately the same level as the trainees in the pretest (67% correct identification), they exhibited little improvement in the three post-tests (70% in the post-test, 67% in Gen 1, and 73% in Gen 2).

The overall results were analyzed using a two-way ANOVA of Test (pretest, post-test, Gen 1, Gen 2) and Group (trained, control), with Test as the repeated measure. There was a significant main effect of Test [$F(1,14)=25.10$, $p<.0001$], Group [$F(1,14)=7.65$, $p<.015$], and a significant Group \times Test interaction [$F(3,42)=11.61$, $p<.0001$]. To further investigate these effects, two one-way ANOVAs were conducted. First, a one-way ANOVA was calculated for each test, with Group as factor. As expected, no reliable difference was obtained between the trained and control group at pretest [$F(1,14)=0.15$, $p>.703$]. However, the two groups were significantly different at the post-test [$F(1,14)=10.33$, $p<.006$], Gen 1 [$F(1,14)=10.59$, $p<.006$], and Gen 2 [$F(1,14)=12.25$, $p<.003$]. This indicates that the trained and control subjects' tone identification accuracy was comparable to start with, but their performance was different after training. Second, a one-way ANOVA

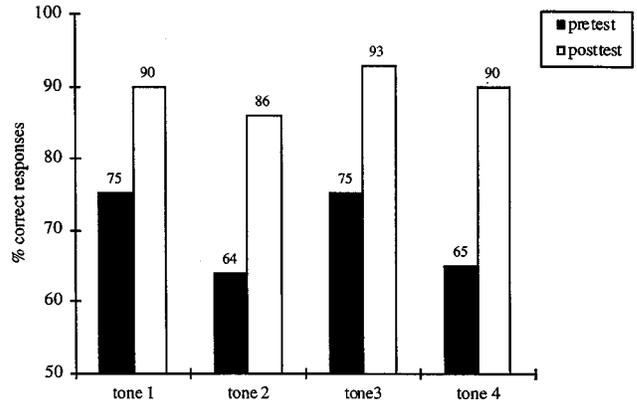


FIG. 3. Trained subjects' mean percent correct identification for each tone at pretest and post-test.

with Test as factor showed, for the trained group, a significant difference among the four tests [$F(3,28)=13.73$, $p<.0001$]. *Post hoc* comparison (Tukey-HSD) showed that the pretest score was significantly lower than that of either the post-test or Gen 1, or Gen 2. Moreover, there were no significant differences among post-test, Gen 1, and Gen 2. Conversely, for the control group, no reliable difference was found among the four tests [$F(3,28)=0.32$, $p>.812$]. Since no difference was found among the post-test, Gen 1, and Gen 2 for either the trained or the control group, subsequent analyses were conducted using the post-test as the representative of the three tests.

B. Individual tones and tone pairs

The trainees' performance for each individual tone is illustrated in Fig. 3, revealing that identification of each tone improved significantly from the pretest to the post-test: 15% improvement for Tone 1 [$F(1,14)=5.15$, $p<.006$]; 22% for Tone 2 [$F(1,14)=7.12$, $p<.001$]; 18% for Tone 3 [$F(1,14)=2.87$, $p<.05$]; and 25% for Tone 4 [$F(1,14)=6.20$, $p<.002$]. Interestingly, there was no significant difference among the four tones at either pretest [$F(1,30)=0.73$, $p>.545$], or post-test [$F(1,30)=0.62$, $p>.607$],

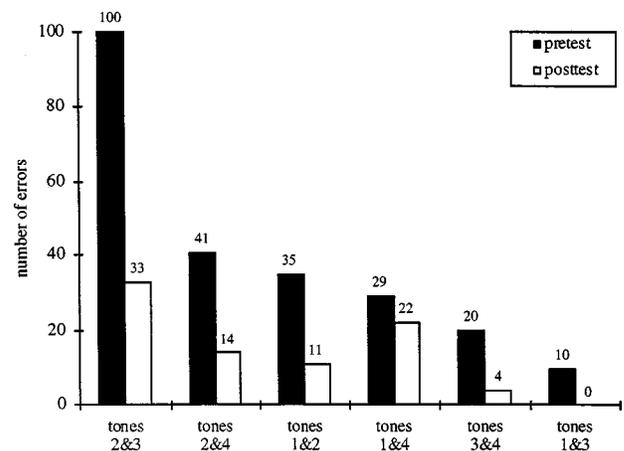


FIG. 4. Tone pair confusions for trained subjects at pretest and post-test. The number of errors (out of 400) for each tone pair refers to misperception of one tone as the other in the corresponding pair.

even though, at pretest, the trainees' identification of Tone 2 and Tone 4 appears poorer as compared to that of Tone 1 and Tone 3.

An analysis of tone confusions is shown in Fig. 4, which compares, for the pretest and the post-test, the number of errors the trainees made for each tone pair out of a total of 400 (25syllables×2tones×8trainees) (see Appendix for complete pretest and post-test confusion matrices). For example, the number of errors for tone pair 1 and 2 is the sum of misperceptions of both Tone 1 as Tone 2, and Tone 2 as Tone 1. In agreement with the overall data, a comparison of the errors made at the pretest and post-test shows a decrease of errors for each tone pair.

The tone pair confusion analysis demonstrated significant differences among the tone pairs for both tests (pretest: [$F(1,46)=9.70, p<.0001$]; post-test: [$F(1,46)=3.81, p<.006$]). *Post hoc* analyses reveal that at pretest, the most difficult tone pair was Tones 2 and 3, followed by Tones 2 and 4, Tones 1 and 2, Tones 1 and 4, Tones 3 and 4, and Tones 1 and 3 (as mentioned previously, this provided the rationale for the reversed order of tone pair presentation during training). However, at post-test, tone pair 1 and 4 became the second most confusing pair next to tone pair 2 and 3. Analysis of variance revealed a significant interaction of tone pair and test (pretest, post-test) [$F(3,92)=9.70, p<.0001$]. More specifically, while all other tone pairs showed a reliable decrease in errors from the pretest to the post-test, the difference between the two tests for Tones 1 and 4 was not significant [$F(1,14)=0.32, p>.577$]. Thus it appears that tone pair 1 and 4 was most resistant to improvement. Nonetheless, the rank order of the tone pairs at pretest and post-test was still highly correlated (Spearman $r=0.83, p<.04$), which indicates that the pattern of tone confusion before and after training is to a large extent comparable.

C. Performance during training

The results from the four assessments during training were analyzed as a function of training set and as a function of talker. Trainees' performance from training set 1 to training set 4 was not significantly different [$F(1,30)=0.61, p>.617$]. The trainees' scores were already very high after the first training set (88% correct identification), and were maintained in the following three assessments (92%, 88%, and 92%, respectively), revealing little progressive improvement as training went along. The high identification accuracy of the four assessments during training might be attributed to the fact that subjects were only tested on the stimuli that were just used in that training session. In addition, since each test represents a different talker, a progressive improvement may not necessarily be expected.

No reliable difference as a function of talker was observed [$F(1,30)=0.38, p>.770$], nor was there any significant difference between the male and female talkers [$F(1,30)=0.88, p>.355$]. Identification scores were 93% and 89% for the two female talkers, and 90% and 88% for the two male talkers.

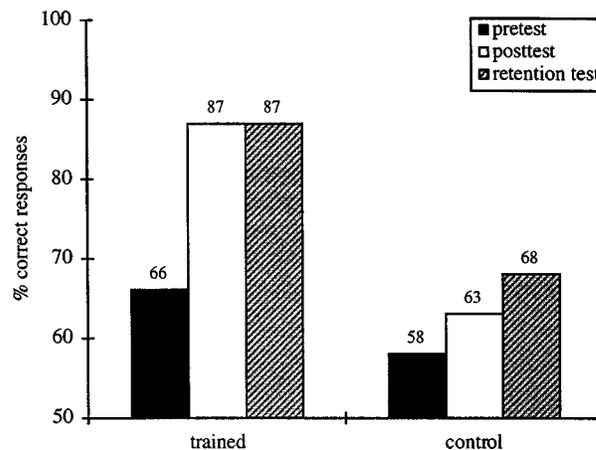


FIG. 5. Mean percent correct identification of the four Mandarin tones for the trained ($n=4$) and the control ($n=4$) subjects at pretest, post-test, and the retention test six months after training.

D. Long-term retention

Figure 5 illustrates the four trained and four control listeners' performance in the retention test as compared to that in the pretest and post-test, revealing that the trainees' improvement was maintained six months after training. The mean identification accuracy for the trainees in the retention test (87%) retains the post-test level (87%), both of which are higher than in the pretest (66%). By contrast, for the controls, the progression from the pretest (58%) to the post-test (63%) and retention test (68%) is much smaller. Detailed analysis of individual listeners revealed that the controls' mean retention score was boosted by the two listeners with three-month Mandarin exposure in Taiwan (32% and 7% improvement from the pretest). Omitting the data from these two listeners would result in a retention test score of 58% for the remaining control subjects, identical to their pretest scores.

A two-way ANOVA of Test (pretest, post-test, retention) and Group (trained, control), with Test as repeated measure, revealed a significant difference in both Test [$F(2,21)=12.44, p<.001$] and Group [$F(1,22)=5.79, p<.05$], but there was no significant Test x Group interaction [$F(2,21)=3.02, p>.087$]. More specifically, a one-way ANOVA was conducted for each group with Test as factor. For the trained group, an expected difference was observed for the three tests [$F(2,9)=9.89, p<.005$], with the pretest score significantly lower than the post-test and retention test (Tukey-HSD). Although the controls show a slight progression of the mean scores from the pretest to the post-test and retention test, there was no significant difference among these tests [$F(2,9)=0.50, p>.619$].

E. Individual trainees

Individual trainee and control performance at pretest, post-test, and retention test is summarized in Table II. Each trainee's identification accuracy improved after training (ranging from 6% to 33%), and the improvement was retained. It should also be noted that there is a large degree of variability among the eight trainees' initial levels, which seems to be reflected the extent of the training effects. Thus

TABLE II. Individual listeners' tone identification accuracy (%) at pretest, post-test, and retention test.

	Pretest	Post-test	Improvement	Retention
Trainee				
1	59	88	+29	81
2	62	83	+21	...
3	63	80	+17	76
4	67	87	+20	92
5	67	100	+33	...
6	73	90	+17	...
7	75	93	+18	98
8	89	95	+6	...
Control				
1	42	60	+18	74
2	55	48	-7	58
3	58	58	0	...
4	60	58	-2	58
5	75	82	+7	...
6	75	86	+11	82
7	85	90	+5	...
8	85	77	-8	...

while the listener with a lower initial score (e.g., listener 1: 59%) showed substantial improvement (29%) in the post-test, training effects were much smaller (6%) for the one who started high (e.g., listener 8: 89% at pretest). It appears a bit surprising that listener 5 reached 100% correct identification at post-test, given that her pretest score was comparatively low (67%). However, a closer inspection of her data showed that in her pretest, 90% of the errors was due to misperception of Tone 3 as Tone 2. Since her problem was limited to one tone pair, improvement may have been easier. The retention test shows that for each of the four trainees, the improvement gained from training was maintained after six months. In particular, the training effect does indeed appear robust, given that listeners were not exposed to Mandarin for as long as three months prior to the retention test.

The trainees' self-evaluation of their performance (obtained from debriefing) is summarized in Table III. Consistent with their actual performance, all listeners recognized some degree of improvement after training. Given that many of them did not claim to have other sources of input that specifically influenced their tone perception, their improvement could largely be attributed to the training. However, although some trainees reported a progressive improvement

during training, and many of them considered female talkers more intelligible, neither of these assessments was mirrored in the data.

Finally, in connection with the language background information of the trainees (cf. Table I), two other minor observations could be made based on the above individual analyses. First, neither trainee 2 nor trainee 4 was taking a Mandarin course during the time of training, yet their improvement (21% and 20%, respectively) was at the average level (21%), which further demonstrates the robustness of training. Second, two listeners with some experience with another tone language (Cantonese) were involved in the training program (trainee 4 and trainee 7). However, an examination of their overall improvement and tone confusion patterns shows that their performance was comparable with the other "nontonal" listeners.

III. DISCUSSION

The present study demonstrated that the perception of Mandarin tones can be improved using a simple training task, indicating that the procedure which has been adopted in training the acquisition of non-native segmental contrasts can also be applied at the suprasegmental level.

The results showed a robust effect of training by a substantial 21% increase in the trainees' overall tone perception accuracy, a significant improvement which also holds true for each of the four tones, and for each individual trainee. More importantly, the improvement gained in training was generalized to new stimuli (18% increase) and new talkers and stimuli (25% increase), and was retained by listeners six months after training (21% increase). These results are comparable to those obtained in the segmental training studies described previously (e.g., Jamieson and Morosan, 1986; Logan *et al.*, 1991; Lively *et al.*, 1994; Bradlow *et al.*, 1997).

Several aspects of tone training warrant discussion related to the general L2 acquisition domain. First, as discussed above, one of the ultimate goals of the acquisition of L2 is the construction of new phonetic categories of the target language. Logan *et al.* (1991) pointed out that the high-variability training procedure facilitates the formation of novel phonetic categories in that stimulus variability exposes learners to the full range of acoustic phonetic cues that char-

TABLE III. Trainees' self-reported performance in the training program.

Trainee	Degree of improvement after training	Progression ^a	Degree of attentiveness in training	More intelligible talker-voice	Other source of tone input
1	moderate	yes	attentive	female	no
2	great	yes	attentive	female	no
3	moderate	yes	occasionally not attentive	female	no
4	great	yes	very attentive	female	no
5	moderate	no	attentive	female	no
6	moderate	not known	attentive	higher voice	self-practice
7	moderate	no	occasionally not attentive	no difference	self-practice
8	moderate	not known	attentive	female	no

^aWas identification progressively easier from sessions 1 to 8?

acterize those categories, while talker variability enables listeners to overcome idiosyncrasies due to differences in talkers' vocal tract size, glottal source function, and speaking rate.

This training procedure was also adopted in the current study; that is, training stimuli were chosen to represent a variety of phonetic environments, and were produced by a number of talkers of both genders. Acoustic analysis has shown that the F_0 pattern for a particular tone is subject to change in different vowels (Howie, 1976). Therefore, it is important that various vowels are used in order for the physical stimuli to be mapped onto more abstract phonemic representations. Talker variability is particularly crucial in tone training, since different talkers (especially males and females) have different fundamental frequencies. It has been reported that native Mandarin speakers use changes in F_0 contours more than height to distinguish among tones, whereas native English listeners tend to attach more importance to height (Gandour, 1983, 1984). Thus by using different talkers, learners are trained to focus on detecting the pitch contour differences of the tones, and to normalize the differences in F_0 height of various talkers. In addition, since intrinsic duration also differs for the four tones (Lin, 1965), talker variability would enable listeners to normalize differences in speaking rate.

All these measures were taken to enhance the tonal category distinctions for the American trainees. As reviewed previously, English listeners' discrimination and identification of Mandarin tones tend to be less "categorical" as compared to Mandarin listeners (Leather, 1987; Stagray and Downs, 1993). Therefore, if training emphasizes those perceptual cues employed by native Chinese to categorize the four tones, the formation of these tonal categories by English learners should be expected. In the present study, the fact that the trainees' identification accuracy increased to a large extent for all the four tones independent of stimuli and talkers, and that the increase had been retained in the trainees' long-term memory, suggests that a separate category for each tone may have been formed and maintained after training.

These results strongly support the previous claim in the segmental domain that adult L2 learners can establish separate phonetic categories for those L2 sounds that are nonexistent in their L1 sound systems (e.g., Flege, 1992). While for native Mandarin speakers tonal pattern is an integral part of the lexicon, such functional association between segmental structure and F_0 contour does not exist in American learners' phonetic systems. In this sense, forming tonal categories is comparable with forming new segmental categories, which may be effortful, but attainable, for adult L2 learners.

However, since, for American listeners, acquiring the Mandarin tone system involves the integration of F_0 information at the lexical and sentential level, their knowledge of the function of pitch in the stress and intonation systems of English may be evident as well. In the present study, although the trainees exhibited an increase in the identification of all the four tones, their tone pair confusion patterns showed that these four tones were indeed differentially acquired. Tone pair 1 and 4 was most resistant to improvement

and was reported by many trainees as "confusing." These two tones were also found difficult for Americans to acquire by Shen (1989), who proposed that Tone 1 and Tone 4 are prosodically less marked for English listeners than Tone 2 and Tone 3. Similarly, White (1981) found that English listeners perceive Mandarin high tones as stressed, and the low Tone 3 as unstressed. Given these findings, it might be that, in this study, Tone 1 and Tone 4 are most resistant to improvement since they are both comparable to the English unmarked or stressed condition, while the other tone pairs each involve at least one tone that is novel or "unnatural" in English. While the initial difficulty in distinguishing Tones 2 and 3 has been attributed to their acoustic similarities (Chen, 1997; Moore and Jongman, 1997), Tones 2 and 3 were improved greatly after training. It might be speculated that since these two tones are so novel to the English listeners, these listeners are more attentive to their distinctions in training. That training can fine-tune distinctions as subtle as tones 2 and 3 may well be due to the novel nature of these two tones to the American listeners.

These findings are consistent with those in the studies of L2 segmental acquisition. For instance, in their study of English vowel acquisition, Bohn and Flege (1992) hypothesized that phonetic learning for similar sounds does not progress much along with L2 experience, whereas new sounds benefit from learning. Likewise, learners are more likely to perceive or produce new, rather than similar, L2 phones authentically (Flege, 1987). Taken together, the present results provide a piece of evidence that the pattern of L2 suprasegmental acquisition might be analogous to that of segmental acquisition, with respect to L1 interference. Although more studies on the comparison of English and Mandarin prosodic patterns are needed to provide a more definite interpretation for the present results, the potential mapping of the patterns of L2 acquisition at segmental and suprasegmental levels is indeed interesting.

IV. CONCLUDING REMARKS

In this study, auditory training at the suprasegmental level was demonstrated to be effective. That is, the perception of Mandarin tones by American learners can be improved with training. The contrasts can be generalized to novel words and talkers, and maintained in long-term memory.

These results raise the question of whether perceptual training can be transferred to production, so that training efforts could result in a facilitating effect (i.e., positive transfer) from one modality to the other (Leather and James, 1991). Since segmental training studies have found that learning gained perceptually can benefit production (Rochet, 1995; Bradlow, 1997), it is worthwhile to test if such transfer will also occur in tone training. Moreover, fine acoustic analysis of American listeners' tone production before and after training, as compared to the native norms, may also be beneficial to quantitatively judge the trainees' improvement after training. Finally, this study only presented training stimuli in isolation. Given that, more often than not, tones are to be perceived and produced in context, training at the

phrasal or sentential levels should also be involved in future studies. These future studies will allow further investigations of the acquisition of Mandarin tones as well as the interaction of L1 and L2 at a suprasegmental level.

ACKNOWLEDGMENTS

The authors wish to thank Qinhong Anderson and Xiaoxin Sun for recruiting listeners, Yufen Lee Mehta, Ning-kang Jiang, Yonghong Mao, Li Tang, Qiuyun Teng, and Dongming Zhang for providing the training stimuli, Eric Evans and Doug Stauffer for technical support, the 16 listeners for their persistent participation, and editor James M. Hillenbrand and three anonymous reviewers for valuable comments and suggestions. Portions of this study were reported at the 135th meeting of the Acoustical Society of America [J. Acoust. Soc. Am. **103**, 3090 (1998)].

APPENDIX

Confusion matrices for the trained group at (a) pretest and (b) post-test (25 stimuli \times 8 trainees = 200 responses for each tone).

(a) Pretest

Perceived as	Stimulus			
	Tone 1	Tone 2	Tone 3	Tone 4
Tone 1	152	7	4	15
Tone 2	28	126	37	37
Tone 3	6	63	152	13
Tone 4	14	4	7	135

(b) Posttest

Perceived as	Stimulus			
	Tone 1	Tone 2	Tone 3	Tone 4
Tone 1	180	4	0	9
Tone 2	7	172	11	11
Tone 3	0	21	185	0
Tone 4	13	3	4	180

Best, C. T. (1995). "A direct realist view of cross-language speech perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Baltimore), pp. 171–204.

Best, C. T., McRoberts, G. W., and Sithole, N. N. (1988). "Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants," *J. Exp. Psychol.* **14**, 345–360.

Blicher, D. L., Diehl, R. L., and Cohen, L. B. (1990). "Effects of syllable duration on the perception of the Mandarin tone2/tone3 distinction: Evidence of auditory enhancement," *J. Phonetics* **18**, 37–49.

Bluhme, H., and Burr, R. (1971). "An audio-visual display of pitch for teaching Chinese tones," *Stud. Linguistics* **22**, 51–57.

Bohn, O. S., and Flege, J. E. (1992). "The production of new and similar vowels by adult German learners of English," *Stud. Second Language Acquisition* **14**, 131–158.

Bradlow, A. R., Pisoni, D. B., Yamada, R. A., and Tohkura. (1997). "Training Japanese listeners to identify English /r/ and /l/ IV: Some effects of perceptual learning on speech production," *J. Acoust. Soc. Am.* **101**, 2299–2310.

Broselow, E., Hurtig, R. R., and Ringen, C. (1987). "The perception of second language prosody," in *Inter-language Phonology, The Acquisition of Second Language Sound System*, edited by G. Ioup and S. H. Weinberger (Newbury House, Cambridge), pp. 350–361.

Chao, Y. R. (1948). *Mandarin Primer* (Harvard University Press, Cambridge).

Chen, G. T. (1974). "The pitch range of English and Chinese speakers," *J. Chinese Linguistics* **2**, 159–171.

Chen, Q. (1997). "Toward a sequential approach for tonal error analysis," *J. Chinese Language Teachers Assoc.* **32**, 21–39.

Chuang, C. K., Hiki, S., Sone, T., and Nimura, T. (1972). "The acoustical features and perceptual cues of the four tones of standard colloquial Chinese," *Proceedings of the Seventh International Congress on Acoustics (Budapest)*, 297–300.

Flege, J. E. (1987). "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification," *J. Phonetics* **15**, 47–65.

Flege, J. E. (1992). "The intelligibility of English vowels spoken by British and Dutch talkers," in *Intelligibility in Speech Disorders: Theory, Measurement and Management*, edited by R. D. Kent (John Benjamins, Amsterdam).

Flege, J. E. (1995). "Second language speech learning: theory, findings, and problems," in *Speech Perception and Linguistic Experience*, edited by W. Strange (York, Baltimore), pp. 233–273.

Gandour, J. T. (1983). "Tone perception in Far Eastern languages," *J. Phonetics* **11**, 149–175.

Gandour, J. T. (1984). "Tone dissimilarity judgments by Chinese listeners," *J. Chinese Linguistics* **12**, 235–261.

Gårding, E., Kratochvil, P., Svantesson, J. O., and Zhang, J. (1986). "Tone 4 and Tone 3 discrimination in modern standard Chinese," *Language and Speech* **29**, 281–293.

Howie, J. M. (1976). *Acoustical Studies of Mandarin Vowels and Tones* (Cambridge University Press, Cambridge).

Jamieson, D. G., and Morosan, D. E. (1986). "Training non-native speech contrasts in adults: Acquisition of the English /ə/-/ɚ/ contrast by francophones," *Percept. Psychophys.* **40**, 205–215.

Jamieson, D. G., and Morosan, D. E. (1989). "Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques," *Can. J. Psychol.* **43**, 88–96.

Kiriloff, C. (1969). "On the auditory discrimination of tones in Mandarin," *Phonetica* **20**, 63–67.

Leather, J. (1987). "F0 pattern inference in the perceptual acquisition of second language tone," in *Sound Patterns in Second Language Acquisition*, edited by A. James and J. Leather (Foris, Dordrecht), pp. 59–81.

Leather, J., and James, A. (1991). "The acquisition of second language speech," *Studies in Second Language Acquisition* **13**, 305–341.

Lenneberg, E. (1967). *Biological Foundations of Language* (Wiley, New York).

Lin, M. C. (1965). "The pitch indicator and the pitch characteristics of tones in Standard Chinese," *Acta Acust. (China)* **2**, 8–15.

Lively, S. E., Logan, J. S., and Pisoni, D. B. (1993). "Training Japanese listeners to identify English /r/ and /l/ II: The role of phonetic environment and talker variability in learning new perceptual categories," *J. Acoust. Soc. Am.* **94**, 1242–1255.

- Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., and Yamada, T. (1994). "Training Japanese listeners to identify English /r/ and /l/ III: Long-term retention of new phonetic categories," *J. Acoust. Soc. Am.* **96**, 2076–2087.
- Logan, J. S., Lively, S. E., and Pisoni, D. B. (1991). "Training Japanese listeners to identify English /r/ and /l/: A first report," *J. Acoust. Soc. Am.* **89**, 874–886.
- Logan, J. S., and Pruitt, J. S. (1995). "Methodological issues in training listeners to perceive non-native phonemes," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Baltimore), pp. 351–377.
- Massaro, D. W., Cohen, M. M., and Tseng, C. (1985). "The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese," *J. Chinese Linguistics* **13**, 267–290.
- Mertus, J. (1989). *BLISS Manual* (Brown University, Providence).
- Moore, C. B., and Jongman, A. (1997). "Speaker normalization in the perception of Mandarin Chinese Tones," *J. Acoust. Soc. Am.* **102**, 1864–1877.
- McClaskey, C. L., Pisoni, D. B., and Carrell, T. D. (1983). "Transfer of training of a new linguistic contrast in voicing," *Percept. Psychophys.* **34**, 323–330.
- Pisoni, D. B., Aslin, R. N., Perey, A. J., and Hennessy, B. L. (1982). "Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants," *J. Exp. Psychol.* **8**, 297–314.
- Rochet, B. L. (1995). "Perception and production of second-language speech sounds by adults," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York, Baltimore), pp. 379–410.
- Shen, X. S. (1989). "Toward a register approach in teaching Mandarin tones," *J. Chinese Language Teachers Assoc.* **24**, 27–47.
- Shen, X. S., and Lin, M. C. (1991). "A perceptual study of Mandarin tones 2 and 3," *Language and Speech* **34**, 145–156.
- Stagray, J. R., and Downs, D. (1993). "Differential sensitivity for frequency among speakers of a tone and a nontone language," *J. Chinese Linguist.* **21**, 143–163.
- Strange, W., and Dittmann, S. (1984). "Effects of discrimination training on the perception of /r-l/ Japanese adults learning English," *Percept. Psychophys.* **36**, 131–145.
- White, C. M. (1981). "Tonal perception errors and interference from English intonation," *J. Chinese Language Teachers Assoc.* **16**, 27–56.

Frequency dependence of ultrasonic backscatter from human trabecular bone: Theory and experiment

Keith A. Wear^{a)}

U.S. Food and Drug Administration, Center for Devices and Radiological Health, HFZ-142,
12720 Twinbrook Parkway, Rockville, Maryland 20852

(Received 30 April 1999; accepted for publication 19 August 1999)

A model describing the frequency dependence of backscatter coefficient from trabecular bone is presented. Scattering is assumed to originate from the surfaces of trabeculae, which are modeled as long thin cylinders with radii small compared with the ultrasonic wavelength. Experimental ultrasonic measurements at 500 kHz, 1 MHz, and 2.25 MHz from a wire target and from trabecular bone samples from human calcaneus *in vitro* are reported. In both cases, measurements are in good agreement with theory. For mediolateral insonification of calcaneus at low frequencies, including the typical diagnostic range (near 500 kHz), backscatter coefficient is proportional to frequency cubed. At higher frequencies, the frequency response flattens out. The data also suggest that at diagnostic frequencies, multiple scattering effects on the average are relatively small for the samples investigated. Finally, at diagnostic frequencies, the data suggest that absorption is likely to be a larger component of attenuation than scattering. [S0001-4966(99)02612-0]

PACS numbers: 43.80.Cs, 43.80.Ev [FD]

INTRODUCTION

Osteoporosis is a major public health issue. In the United States, over 1.5 million fractures attributed to osteoporosis occur each year. These include fractures in vertebrae (500 000), hip (250 000), and distal forearm (240 000), with annual treatment-related cost exceeding \$10 000 000 000.¹

Bone mineral density (BMD) is one of the primary predictive risk factors for osteoporotic fracture.²⁻⁷ An expert panel of the World Health Organization (WHO) has proposed diagnostic criteria for osteoporosis based on BMD measurements. The WHO criteria were recently applied to femoral BMD measured by dual energy x-ray absorptiometry (DEXA) in men and women participating in the third National Health and Nutrition Examination Survey (NHANES III), 1988-1994.⁵ The prevalence of osteoporosis (defined by WHO as BMD below a threshold equal to 2.5 standard deviations below the mean of young, normal non-Hispanic Caucasian women, a threshold intentionally chosen so that the prevalence would approximate that of lifetime osteoporotic fractures in Caucasian women) in the USA is 13%-18% or 13-17 million.⁵

Common methods for assessing bone density include DEXA and quantitative computed tomography (QCT). These methods are expensive, involve ionizing radiation, and require relatively sophisticated equipment and are often unavailable in smaller hospitals with low caseloads. Finally, while these measurements correlate relatively well with bone mineral density, they are not necessarily very sensitive to *micro-architecture* or to the protein matrix of bone (which are also important determinants of fracture risk).

Approaches based on ultrasound offer advantages in-

cluding low cost, lack of ionizing radiation, speed, simplicity, and portability. Most current methods are based on broadband ultrasonic attenuation (BUA) and speed-of-sound (SOS) measurements in the calcaneus. It has been demonstrated that ultrasonic attenuation is highly correlated with calcaneal mass density.⁸⁻¹¹ Calcaneal ultrasonic measurements (BUA combined with SOS) have been shown to perform well for prediction of hip fractures in elderly women in prospective^{12,13} and retrospective¹⁴⁻¹⁶ studies. Despite this diagnostic utility, the fundamental mechanisms underlying the interaction between ultrasound and calcaneus are not well understood presently.

Measurements of ultrasonic backscattering properties may provide additional useful information. The decreased number and size of trabeculae (which may serve as scattering sites) within bone that accompany the aging process would be expected to reduce backscatter. The clinical feasibility and diagnostic promise of this measurement have already been demonstrated.¹⁷⁻¹⁹

The objective of this paper is to investigate mechanisms underlying ultrasonic scattering in trabecular bone. Toward this end, a theoretical model for ultrasonic scattering from trabecular bone is presented and subsequently tested experimentally. Tests are performed to assess the magnitude of multiple scattering effects in trabecular bone.

I. THEORY

Bone tissue may be classified into two types: cortical (or compact) bone and trabecular (or spongy) bone. Cortical bone is most abundant in the shafts of long bones such as the tibia, though it constitutes the outer shell of every bone in the body. Trabecular bone is concentrated in the vertebrae, pelvis, calcaneus, and at the ends of long bones, though it is present in the interiors of all bones. The calcaneus is predominantly composed of trabecular bone covered by a thin

^{a)}Electronic mail: kaw@cdrh.fda.gov

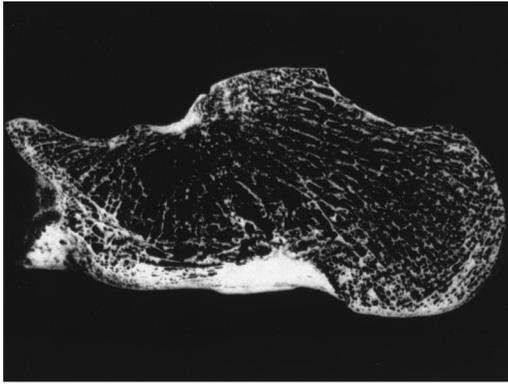


FIG. 1. Calcaneus sample with lateral cortical layers removed. Long thin trabeculae are apparent.

cortical shell. Trabecular bone consists of a three-dimensional lattice of branching spicules or plates. The spaces between the trabeculae are filled with marrow, which consists of fat and cellular components of blood constituents. In this paper, scattering from trabecular bone is investigated.

Trabeculae within bone are postulated to be the principle sources of ultrasound scattering and are modeled here as cylinders with diameters that are small relative to the ultrasonic wavelength. In addition, it is assumed that these cylinders are long relative to the ultrasound beam cross section and are oriented perpendicular to the ultrasound propagation direction. Figure 1 shows the presence of long thin trabeculae within the interior of a calcaneus, which is consistent with this model. Although the trabeculae exhibit a myriad of orientations, they are all approximately perpendicular to the ultrasound propagation direction provided that the bone is interrogated in the mediolateral (or lateromedial) direction (perpendicular to the plane of Fig. 1). A typical value for human calcaneal trabecular thickness is $120 \pm 10 \mu\text{m}$ (mean \pm standard deviation).²⁰ The wavelengths of ultrasound (in water) used in the experiment were much larger and ranged from 3.1 mm (at 500 kHz) to 0.7 mm (at 2.25 MHz).

The intensity of an acoustic wave scattered from a solid cylinder, I_s , is given by^{21,22}

$$I_s = \frac{Ia}{\pi r} |\varphi_s(\phi)|^2, \quad (1)$$

where I is the intensity of the incident plane wave, a is the radius of the cylinder, r is the distance from the cylinder axis to the observation point, and

$$|\varphi_s(\phi)|^2 = \frac{1}{ka} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \varepsilon_m \varepsilon_n \sin \eta_m \sin \eta_n \times \cos(\eta_m - \eta_n) \cos(m\phi) \cos(n\phi), \quad (2)$$

where ϕ is the angle between the incident and observation directions, $k=2\pi/\lambda$, λ is the wavelength in the fluid surrounding the scatterer, $\varepsilon_0=1$, $\varepsilon_m=2$ ($m>0$), and η_m is the phase angle for scattering from a cylinder given by²¹

$$\tan \eta_n = \tan \delta_n(x) \frac{\tan \Phi_n + \tan \alpha_n(x)}{\tan \Phi_n + \tan \beta_n(x)}, \quad (3)$$

where $x=ka$,

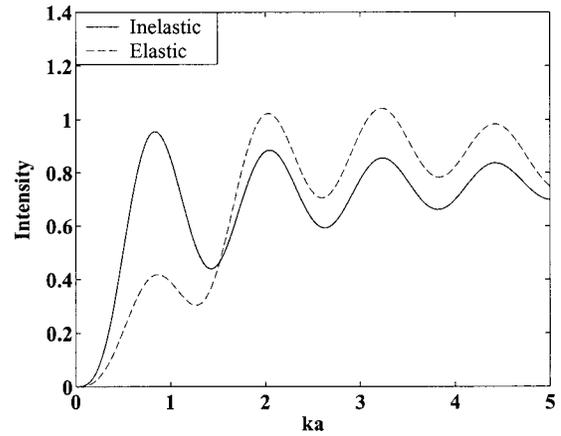


FIG. 2. Theoretical backscatter from a hydroxyapatite cylinder immersed in water as a function of ka where $k=2\pi/\lambda$ and a is the radius. The solid curve corresponds to the inelastic model and the dashed curve corresponds to the elastic model.

$$\delta_n(x) = \tan^{-1}[-J_n(x)/N_n(x)], \quad (4)$$

$$\alpha_n(x) = \tan^{-1}[-xJ'_n(x)/J_n(x)], \quad (5)$$

$$\beta_n(x) = \tan^{-1}[-xN'_n(x)/N_n(x)], \quad (6)$$

and

$$\tan \Phi_n = (-\rho/\rho_1) \tan \zeta_n(x_1, \sigma), \quad (7)$$

where J_m and N_m correspond to Bessel and Neumann functions, primes denote differentiation, ρ is the density of the fluid surrounding the scatterer, ρ_1 is the density of the cylindrical scatterer, and σ =Poisson's ratio. For inelastic scattering, $\tan \Phi_n=0$ for all n . In general, the scattering phase angle, $\zeta_n(x_1, \sigma)$ is given by

$$\tan \zeta_n(x_1, \sigma) = -\frac{x_2^2}{2} \frac{\frac{\tan \alpha_n(x_1)}{\tan \alpha_n(x_1) + 1} \frac{n^2}{\tan \alpha_n(x_2) + n^2 - x_2^2/2}}{\frac{\tan \alpha_n(x_1) + n^2 - x_2^2/2}{\tan \alpha_n(x_1) + 1} \frac{n^2 [\tan \alpha_n(x_2) + 1]}{\tan \alpha_n(x_2) + n^2 - x_2^2/2}}, \quad (8)$$

where $x_1=k_1a$, $x_2=k_2a$, $k_1=\omega/c_1$, $k_2=\omega/c_2$, c_1 =the compressional wave velocity in the scatterer, and c_2 =the shear wave velocity in the scatterer. Since the ratio of x_1 to x_2 is completely determined by σ , ζ_n may be expressed simply as a function of x_1 and σ . For backscatter, $\phi=180^\circ$.

Scattered intensities for both inelastic (solid line) and elastic (dashed line) scattering from a cylinder as functions of ka are shown in Fig. 2. For the elastic case, material properties for hydroxyapatite ($c_1=6790$ m/s, $\rho=3.22$ g/cm³, and $\sigma=0.28$)^{23,24} scatterers in water ($c_1=1480$ m/s at 20 °C and $\rho=1.00$ g/cm³) were assumed. Submicroscopic deposits of calcium phosphate, similar but not necessarily identical to hydroxyapatite, $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$, are the major inorganic constituents of bone.²⁵ Anderson, Soo, and Trahey, in modeling microcalcifications in the breast, performed a similar comparison in which they employed inelastic and elastic models for spherical scatterers composed of hydroxyapatite.²⁴ The relationship between inelastic and

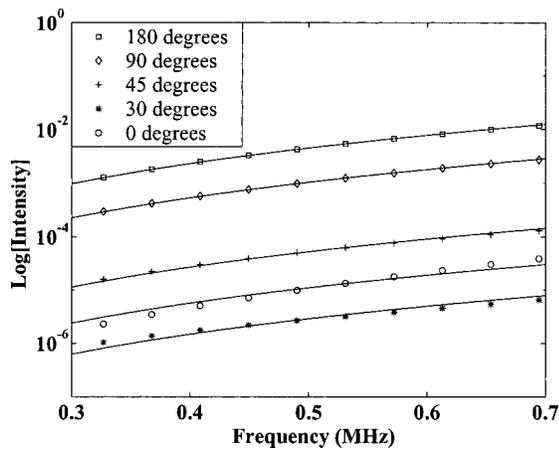


FIG. 3. The logarithm of the intensity of an elastically scattered wave from a hydroxyapatite cylinder versus frequency for various angles (ϕ). Also shown are cubic fits to the data. The apparent juxtaposition of the 0° and 30° data is due to minima in the angular distribution at $\pm 25^\circ$.

elastic scattered intensities is very similar for the two different shapes. For both cylinders and spheres, there are four roughly equally spaced maxima in the range $0 < ka < 5$. The locations of the maxima are approximately the same for the elastic and inelastic cases. For both cylinders and spheres, the first elastic peak is lower in magnitude than the first inelastic peak, while the second through fourth elastic peaks are higher in magnitude.

In the low frequency limit ($ka \ll 1$), it may be shown that the intensity of the inelastically scattered wave is given by²²

$$I_s \cong \frac{\pi k^3 a^4}{8r} I(1 - 2 \cos \phi)^2. \quad (9)$$

Thus, in the low-frequency limit, inelastic scattering becomes proportional to the cube of the ultrasonic frequency at all angles. This is at least approximately true for the elastic case, as can be seen in Fig. 3, where cubic fits agree well with scattered wave intensities at various angles. The angular distribution of scattering at 500 kHz is illustrated in Fig. 4. Scattering is most prominent in the backward direction.

This theory applies to a single cylindrical scatterer. It can also describe the frequency dependence of scattering

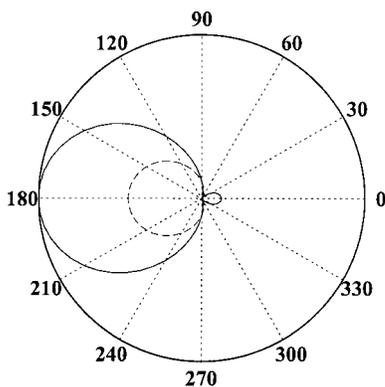


FIG. 4. Angular distribution of scattering for the inelastic case (solid line) and elastic case (dashed line). Scattering is most prominent in the backward direction.

from a superposition of scattered signals from many cylindrical sources, provided two conditions are satisfied. The first condition is that echoes from different sources add incoherently. This will occur if the cylindrical scatterers are positioned sufficiently randomly that the phases of their scattered waves may be taken to be uniformly distributed on the interval from 0 to 2π .^{26,27} The second condition is that the effects of multiple scattering are negligible.

II. EXPERIMENTAL METHODS

A. Biological methods

Sixteen human (gender and age unknown) calcaneus samples were obtained. They were defatted using a trichloro-ethylene solution. Defatting was presumed not to significantly affect measurements since attenuation^{11,28} and speed of sound^{28,29} of defatted trabecular bone have been measured to be only slightly different from their counterparts with marrow left intact. The cortical lateral sides were sliced off, leaving two parallel surfaces with direct access to trabecular bone. The thicknesses of the samples varied from 14 to 21 mm. 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 was measured for each experiment and ranged between 19.1 and 21.2 °C. The relative orientation between the ultrasound beam and the calcanei was the same as with *in vivo* measurements performed with commercial bone sonometers, in which sound propagates in the mediolateral (or lateromedial) direction.

Apparent density, the ratio of the dehydrated, defatted tissue mass to the total specimen volume,¹¹ was measured for each sample. Mass was measured using a balance. Volume was assessed from separate measurements of thickness and cross-sectional area. Thickness between the two parallel planar surfaces (cut by machine) was measured using calipers. Cross-sectional areas were measured by computer processing of scanned images of the samples.

B. Ultrasonic methods

In addition to the bone samples, a wire oriented perpendicular to the ultrasound propagation direction was interrogated. These measurements were used to validate the measurement methodology and the numerical computation for the frequency-dependent backscatter from a cylinder [Eq. (2) with $\phi = 180^\circ$]. Repeated measurements were performed in order to generate means and standard deviations.

A Panametrics (Waltham, MA) 5800 pulser/receiver was used. Samples were interrogated in a water tank using Panametrics circular, focused, broadband transducers with center frequencies of 500 kHz, 1 MHz, and 2.25 MHz. 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 general purpose interface bus (GPIB)] for off-line analysis.

Backscatter coefficients were measured using a reference phantom method.³⁰ 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 minimized. A reference phantom was used with known frequency-dependent backscatter coefficient, $\eta_R(f)$, and attenuation coefficient, $\alpha_R(f)$, placed in the water tank at the same distance as for the bone samples. The phantom consisted of glass spheres in agar with $\alpha_R(f) = 0.124f^{1.63}$ dB/cm, $c = 1556$ m/s, $\eta_R(f) = 1.15 \times 10^{-4}$, 1.46×10^{-3} , and $9.35 \times 10^{-3} \text{ cm}^{-1}\text{Sr}^{-1}$ at 500 kHz, 1 MHz, and 2.25 MHz, respectively. Power spectra from the bone samples, $I_B(f)$, and the phantom $I_R(f)$ were acquired. The backscatter coefficient, $\eta_B(f)$, is then obtained from

$$\eta_B(f) = \frac{1}{T^4} \frac{G[\alpha_B(f), l] I_B(f)}{G[\alpha_R(f), l] I_R(f)} \eta_R(f), \quad (10)$$

where T is the amplitude transmission coefficient at the water/bone interface. Due to the high degree of porosity of the samples, T was assumed to be 1.³¹ The function $G[\alpha(f), l]$ compensates for attenuation effects and is given by²⁶

$$G[\alpha(f), l] = \frac{4\alpha(f)l}{1 - e^{-4\alpha(f)l}}, \quad (11)$$

where l is the gate length. Good agreement between experimental measurements using this method and theoretical predictions based on Faran's theory of scattering from spheres²¹ for ultrasonic backscatter coefficients from phantoms has been previously reported by this laboratory.³² [Note that Eq. (11) differs from a previously used attenuation compensation function,¹⁹ since in the present experiment there was water rather than tissue or phantom between the transducer and the gated volume.]

Attenuation measurements were required in order to compensate signals prior to backscatter coefficient estimation. Attenuation was measured using a standard through-transmission 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. The bone samples were larger in cross-sectional area than the receiving transducer apertures. Attenuation coefficient was then estimated using a log spectral difference technique.³³ In principle, this substitution technique can exhibit appreciable error if the speed of sound differs substantially between the sample and the reference.³⁴ However, one study indicates that this diffraction-related error is negligible in the calcaneus.³¹ Evidently the speed of sound in the calcaneus, approximately 1475–1650 m/s,³¹ is sufficiently close to that of distilled water at room temperature, 1487 m/s,³⁵ that diffraction-related errors may be ignored.

III. RESULTS

The means and standard deviations for the measured properties of the calcaneus samples were 0.33 ± 0.11 g/ml (apparent density), 9.94 ± 5.74 dB/cm MHz (attenuation slope), and $0.062 \pm 0.066 \text{ cm}^{-1}\text{Sr}^{-1}$ (backscatter coefficient at 500 kHz).

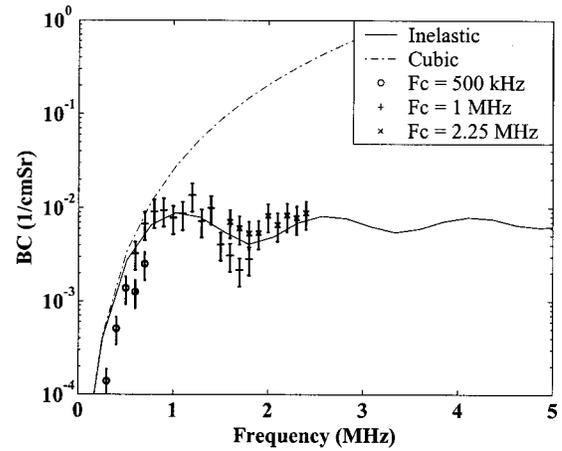


FIG. 5. Backscatter from a wire test object. The solid curve is the theoretical frequency-dependent backscatter from a rigid cylinder [Eq. (1) with $\phi = 180^\circ$] based on the actual radius of the wire ($a = 191 \mu\text{m}$). The chain-dashed curve is the low-frequency cubic approximation [Eq. (9) with $\phi = 180^\circ$]. The magnitudes of the theoretical curves were arbitrarily selected to fit the data. Good agreement in frequency dependence between theory and experiment may be seen. Error bars denote standard errors of measurements. The three symbols correspond to the center frequencies of the three transducers used: 500 kHz (circle), 1 MHz (+), and 2.25 MHz (\times).

Figure 5 shows experimental measurements of backscatter coefficient for the wire target. Also, shown are the general frequency-dependent inelastic backscatter [Eq. (1) with $\phi = 180^\circ$] and the low-frequency cubic approximation [Eq. (9) with $\phi = 180^\circ$] based on the actual radius of the wire ($a = 191 \mu\text{m}$). The magnitudes of the theoretical curves were arbitrarily adjusted to fit the data. Good agreement in frequency dependence between theory and experiment may be seen.

Figure 6 shows experimental measurements of average backscatter coefficient for the 16 calcaneus samples. Also

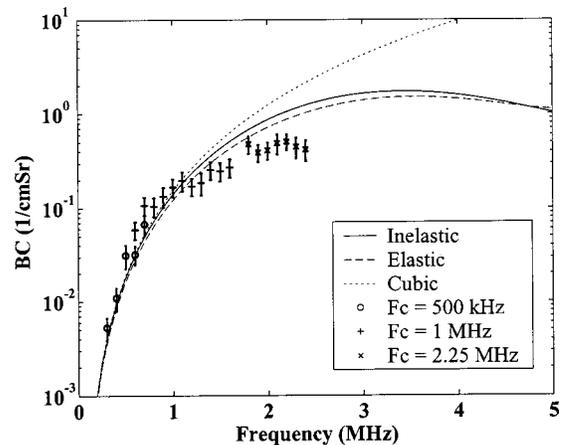


FIG. 6. Experimental measurements of average backscatter coefficient for the 16 calcaneus samples. Also shown are the inelastic (solid line) and elastic (dashed line) frequency-dependent backscatter coefficients [Eq. (1) with $\phi = 180^\circ$] based on a published value for the average radius of calcaneal trabeculae. The dotted line is the low-frequency cubic approximation [Eq. (9) with $\phi = 180^\circ$]. Again, the magnitudes of the theoretical curves were arbitrarily selected to fit the data. Good agreement in frequency dependence between theory and experiment may be seen. Error bars denote standard errors of measurements. The three symbols correspond to the center frequencies of the three transducers used: 500 kHz (circle), 1 MHz (+), and 2.25 MHz (\times).

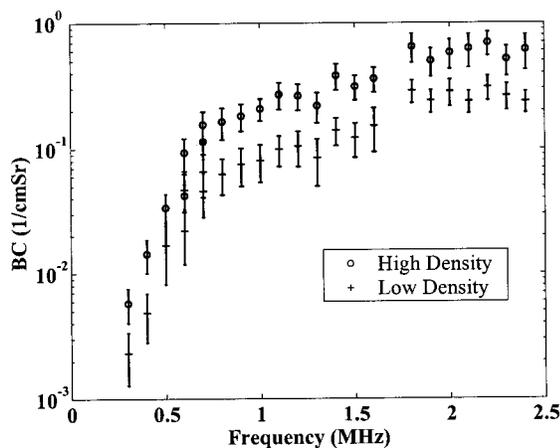


FIG. 7. Frequency-dependent experimental measurements for backscatter coefficients for two subgroups of the bone samples: the eight lowest density (circle) and the eight highest density (+). Both subgroups exhibit essentially the same frequency dependence. The high-density group backscatters more strongly in a uniform fashion throughout the set of frequencies employed. Error bars denote standard errors of measurements.

shown are the general frequency-dependent backscatter [Eq. (1) with $\phi = 180^\circ$] for both the inelastic model (solid line) and the elastic model (dashed line) as well as the low-frequency cubic approximation [Eq. (9) with $\phi = 180^\circ$], assuming the material properties for hydroxyapatite given above and $a = 60 \mu\text{m}$.²⁰ Again, the magnitudes of the theoretical curves were arbitrarily adjusted to fit the data. There is little difference in the frequency dependences for the inelastic and elastic models in the low-frequency range examined here. Good agreement in frequency dependence between both theoretical models and experiment may be seen. The low-frequency data (300–700 kHz) were least-squares fit to power-law functions, Af^n , where A is an arbitrary multiplicative constant and n is the exponent. For the 16 samples, n was found to be 3.26 ± 0.20 (mean \pm standard error), very close to the expected cubic dependence.

Figure 7 shows frequency-dependent experimental measurements for backscatter coefficients for the bone samples, now divided into two subgroups: the eight lowest density and the eight highest density. Both subgroups exhibit essentially the same frequency dependence. The high-density group backscatters more strongly in a uniform fashion throughout the set of frequencies employed. The means and standard errors of ratios of high-to-low backscatter coefficients were comparable for all three transducer center frequencies: 2.4 ± 0.6 (500-kHz data), 2.5 ± 0.3 (1-MHz data), and 2.3 ± 0.3 (2.25-MHz data).

IV. DISCUSSION

A model describing the frequency dependence of backscatter coefficient from trabecular bone has been presented. Scattering was assumed to originate from the surfaces of trabeculae, which were modeled as long thin cylinders with radii small compared with the ultrasonic wavelength. Experimental ultrasonic measurements at 500 kHz, 1 MHz, and 2.25 MHz from a wire target and from trabecular bone samples from human calcaneus *in vitro* have been presented. In both cases, measurements were in good agreement with

theoretical predictions. For insonification in the mediolateral direction at low frequencies ($ka \ll 1$), including the typical diagnostic range (near 500 kHz), backscatter coefficient is proportional to frequency cubed. At higher frequencies, the frequency response flattens out.

The data presented here have implications with regard to the relative roles of scattering and absorption in determining attenuation in trabecular bone at diagnostic frequencies. Attenuation in the range between 300 and 700 kHz has been found in numerous studies to be approximately proportional to frequency to the first power.^{8–11} The theoretical model presented here predicts that, at low frequencies, scattering in all directions is approximately proportional to frequency cubed. These two different frequency dependences could be consistent if absorption is a larger component of attenuation than scattering.

The data presented here suggest that, at typical diagnostic frequencies (near 500 kHz), the effects of multiple scattering on the average are relatively small for the samples investigated. This is evidenced in Fig. 6 as, in the low-frequency regime, the theoretically expected frequency-cubed dependence of backscatter coefficient was observed. Multiple scattering, if present, would have had the effect of applying additional high-pass (frequency-cubed) filtering to the echoes (via multiple scatters), resulting in a measured backscatter coefficient with a more rapid than cubic frequency dependence. Even if multiple scattering were present, the relatively high attenuation coefficient of trabecular bone would tend to suppress it as multiply scattered waves tend to traverse longer path lengths than singly scattered waves.

Measurements of the frequency dependence of scattering might in principle be used to infer scatterer size (here, trabecular width), similar to methodology successfully employed by Insana and co-workers to evaluate sizes of different scattering components in the kidney.^{36–38} However, several factors would make accurate, diagnostically useful determination of trabecular width extremely challenging. First, the size of the calcaneus is rather small, limiting the quantity of data that can be averaged for backscatter coefficient estimation. Second, the high attenuation coefficient of calcaneus limits the bandwidth of frequencies available for analysis.

Figure 7 illustrates the potential diagnostic utility of the backscatter coefficient measurement. That is, higher backscatter coefficients are associated with denser bones. This result reinforces a previously published finding in normal individuals *in vivo* in which a measure of the magnitude of ultrasonic backscatter was found to correlate highly with bone density assessed using QCT.¹⁹

ACKNOWLEDGMENTS

The author is grateful for funding provided by the US Food and Drug Administration Office of Women's Health. In addition, the suggestions of Matthew Myers, FDA, and William Sacks, FDA, are greatly appreciated. The author also thanks Tim Hall, University of Kansas, for providing the phantom used in this work.

- ¹L. Lenchik and D. J. Sartoris, "Current concepts in osteoporosis," *Am. J. Roentgenol.* **168**, 905–911 (1997).
- ²L. J. Melton *et al.*, "Long-term fracture prediction by bone mineral assessed at different skeletal sites," *J. Bone Miner. Res.* **8**, 1227–1233 (1993).
- ³D. M. Black *et al.*, "Axial and appendicular bone density predict fracture in older women," *J. Bone Miner. Res.* **7**, 633–638 (1992).
- ⁴S. R. Cummings *et al.*, "The study of osteoporotic fractures research group: Bone density at various sites for the prediction of hip fractures," *Lancet* **341**, 72–75 (1993).
- ⁵A. C. Looker *et al.*, "Prevalence of low femoral bone density in older U.S. adults from NHANES III," *J. Bone Miner. Res.* **12**, 1761–1768 (1997).
- ⁶J. L. Kanis *et al.*, "The diagnosis of osteoporosis," *J. Bone Miner. Res.* **9**, 1137–1141 (1994).
- ⁷World Health Organization, "Assessment of fracture risk and its application to screening for postmenopausal osteoporosis," Technical Report Series No. 843 (WHO, Geneva, Switzerland).
- ⁸C. M. Langton, S. B. Palmer, and R. W. Porter, "The measurement of broadband ultrasonic attenuation in cancellous bone," *Eng. Med. (Berlin)* **13**, 89–91 (1984).
- ⁹P. Rossman, 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).
- ¹⁰M. B. Tavakoli and J. A. Evans, "Dependence of the velocity and attenuation of ultrasound in bone on the mineral content," *Phys. Med. Biol.* **36**, 1529–1537 (1991).
- ¹¹C. M. Langton *et al.*, "Prediction of mechanical properties of the human calcaneus by broadband ultrasonic attenuation," *Bone (N.Y.)* **18**, 495–503 (1996).
- ¹²D. Hans *et al.*, "Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study," *Lancet* **348**, 511–514 (1996).
- ¹³D. C. Bauer *et al.*, "Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women," *Arch. Intern. Med.* **157**, 629–634 (1997).
- ¹⁴M. 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).
- ¹⁵C. H. Turner *et al.*, "Calcaneal ultrasonic measurements discriminate hip fracture independently of bone mass," *Osteoporosis Int.* **5**, 130–135 (1995).
- ¹⁶C. C. Glüer *et al.*, "Osteoporosis: Association of recent fractures with quantitative US findings," *Radiology* **199**, 725–732 (1996).
- ¹⁷K. A. Wear and B. S. Garra, "Assessment of bone density using broadband ultrasonic backscatter," Proceedings of the 22nd Int. Symposium on Ultrason. Imag. and Tissue Char., Washington, DC, 1997, p. 14 (Abstract).
- ¹⁸P. Giat, C. Chappard, C. Roux, P. Laugier, and G. Berger, "Preliminary clinical assessment of the backscatter coefficient in osteoporosis," Proceedings of the 22nd Int. Symposium Ultrason. Imag. and Tissue Char., Washington, DC, 1997, p. 16 (Abstract).
- ¹⁹K. A. Wear and B. S. Garra, "Assessment of bone density using ultrasonic backscatter," *Ultrasound Med. Biol.* **24**, 689–695 (1998).
- ²⁰K. D. Hausler, P. A. Rich, P. C. Smith, and E. B. Barry, "Relationships between static histomorphometry and ultrasound in the human calcaneus," *Calcif. Tissue Int.* **64**, 477–480 (1999).
- ²¹J. J. Faran, "Sound scattering by solid cylinders and spheres," *J. Acoust. Soc. Am.* **23**, 405–418 (1951).
- ²²P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, 1986).
- ²³D. E. Grenoble *et al.*, "The elastic properties of hard tissues and apatites," *J. Biomed. Mater. Res.* **6**, 221–233 (1972).
- ²⁴M. E. Anderson, M. S. C. Soo, and G. E. Trahey, "Microcalcifications as elastic scatters under ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 925–934 (1998).
- ²⁵R. P. Heany, "Calcium," in *Principles of Bone Biology*, edited by J. P. Bilezikian, L. G. Raisz, and Gideon A. Rodan (Academic, San Diego, 1996).
- ²⁶M. O'Donnell and J. G. Miller, "Quantitative broadband ultrasonic backscatter: An approach to nondestructive evaluation in acoustically inhomogeneous materials," *J. Appl. Phys.* **52**, 1056–1065 (1981).
- ²⁷E. L. Madsen, M. F. Insana, and J. A. Zagzebski, "Method of data reduction for accurate determination of acoustic backscatter coefficients," *J. Acoust. Soc. Am.* **76**, 913–923 (1984).
- ²⁸J. 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).
- ²⁹C. 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).
- ³⁰J. 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 Press, Boca Raton, 1993).
- ³¹P. Droin, G. Berger, and P. Laugier, "Velocity dispersion of acoustic waves in cancellous bone," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 581–592 (1998).
- ³²K. 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).
- ³³R. 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).
- ³⁴W. Xu and J. J. Kaufman, "Diffraction correction methods for insertion ultrasound attenuation estimation," *IEEE Trans. Biomed. Eng.* **40**, 563–570 (1993).
- ³⁵A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications* (McGraw-Hill, New York, 1981), p. 31.
- ³⁶M. F. Insana and D. G. Brown, "Acoustic scattering theory applied to soft biological tissues," in *Ultrasonic Scattering in Biological Tissues*, edited by K. K. Shung and G. A. Thieme (CRC Press, Boca Raton, 1993).
- ³⁷M. 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).
- ³⁸T. J. Hall, M. F. Insana, L. A. Harrison, and G. G. Cox, "Ultrasonic measurement of glomerular diameters in normal adult humans," *Ultrasound Med. Biol.* **22**, 987–997 (1996).

Simulation of ultrasonic pulse propagation, distortion, and attenuation in the human chest wall

T. Douglas Mast

Applied Research Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

Laura M. Hinkelman^{a)}

Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14627

Leon A. Metlay

Department of Pathology and Laboratory Medicine, University of Rochester Medical School, Rochester, New York 14642

Michael J. Orr

Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14627

Robert C. Waag

Departments of Electrical and Computer Engineering and Radiology, University of Rochester, Rochester, New York 14627

(Received 5 April 1999; accepted for publication 23 August 1999)

A finite-difference time-domain model for ultrasonic pulse propagation through soft tissue has been extended to incorporate absorption effects as well as longitudinal-wave propagation in cartilage and bone. This extended model has been used to simulate ultrasonic propagation through anatomically detailed representations of chest wall structure. The inhomogeneous chest wall tissue is represented by two-dimensional maps determined by staining chest wall cross sections to distinguish between tissue types, digitally scanning the stained cross sections, and mapping each pixel of the scanned images to fat, muscle, connective tissue, cartilage, or bone. Each pixel of the tissue map is then assigned a sound speed, density, and absorption value determined from published measurements and assumed to be representative of the local tissue type. Computational results for energy level fluctuations and arrival time fluctuations show qualitative agreement with measurements performed on the same specimens, but show significantly less waveform distortion than measurements. Visualization of simulated tissue-ultrasound interactions in the chest wall shows possible mechanisms for image aberration in echocardiography, including effects associated with reflection and diffraction caused by rib structures. A comparison of distortion effects for varying pulse center frequencies shows that, for soft tissue paths through the chest wall, energy level and waveform distortion increase markedly with rising ultrasonic frequency and that arrival-time fluctuations increase to a lesser degree. © 1999 Acoustical Society of America. [S0001-4966(99)03212-9]

PACS numbers: 43.80.Qf, 43.80.Cs, 43.58.Ta, 43.20.Fn [FD]

INTRODUCTION

Echocardiography is widely employed for diagnosis of cardiac diseases including valvular defects, pericardial effusion, and wall motion abnormalities.¹⁻³ Commonly, echocardiography is performed noninvasively through the chest (transthoracic) using an external probe placed on the chest wall. The chest wall, however, can considerably degrade image quality because acoustic paths between the skin and heart may contain ribs and cartilage as well as inhomogeneous muscle and fatty tissue. The result is that as many as 10-30% of patients cannot be successfully imaged with present transthoracic techniques.⁴ This limitation of transthoracic echocardiography has led to the development of transeophageal echocardiography, in which the heart is imaged by a transducer inserted into the esophagus.¹⁻⁴ Although transeophageal echocardiography provides superior image qual-

ity, resulting in high diagnostic sensitivity and specificity, the invasiveness of the procedure is accompanied by increased risk.³⁻⁶ For this reason, improvements in the noninvasive transthoracic approach are desirable, for example, by the development of methods to compensate for image degradation caused by the chest wall.

An understanding of ultrasonic aberration produced by the chest wall is important to the development of appropriate compensation methods for transthoracic ultrasonic imaging. Direct measurements of ultrasonic distortion produced by chest wall specimens^{7,8} have been helpful. Results reported in Ref. 7 show that propagation through the chest wall causes substantial beam distortion. However, that study did not distinguish the effect of soft tissue from effects caused by rib structures. In Ref. 8, a detailed study of distortion caused by soft tissue paths indicates that soft tissue distortion in the chest wall is substantially less than the corresponding distortion in the human abdominal wall. However, distortion caused by ribs was only treated qualitatively in the latter

^{a)}Present address: Department of Meteorology, The Pennsylvania State University, University Park, PA 16802.

study because the physical mechanisms of rib-induced distortion could not be adequately described by the method reported there. Although a model of ultrasound propagation in the chest wall has previously been described,⁹ that model is based on coarse depictions of chest wall morphology including homogeneous tissue layers and evenly-spaced, uniformly-shaped ribs. These previous experiments and simulations, therefore, have left gaps in the current knowledge about the physical causes of ultrasonic wavefront distortion caused by the chest wall.

Recent work on the simulation of ultrasonic pulse propagation^{10–12} has provided insight about the wavefront distortion caused by the human abdominal wall. Although these studies have provided specific information about the relationships between soft tissue morphology and ultrasonic wavefront distortion, the work is not fully applicable to distortion caused by the human chest wall. The morphology of chest wall soft tissue is different from that of the abdominal wall in ways that can affect ultrasonic wavefront distortion.⁸ Furthermore, imaging through the chest wall is complicated by ribs that limit the usable acoustic window size and cause scattering and reflection.

The study reported here applies quantitative simulation methods, similar to those presented in Refs. 10 and 12, to anatomically detailed chest wall models that include the ribs. Accurate depiction of rib–ultrasound interactions requires not only representation of the strong reflections associated with sound speed and density contrast between ribs and soft tissue (already accurately modeled by the finite difference method of Ref. 10), but also modeling of the strong losses associated with propagation through bone and cartilage. For this reason, the finite-difference method described in Ref. 10 has been extended to include tissue-dependent absorption. Quantitative descriptions of the distortion caused by soft tissues are obtained by statistical analysis of simulated distortion. Visualizations of wavefronts propagating through maps of chest cross sections provide evidence about physical relationships between wavefront distortion and the morphology of ribs and soft tissue structures in the chest wall. Further insight about wavefront distortion mechanisms is provided by a comparison of distortion results for incident pulses of different center frequencies.

I. THEORY

Ultrasonic pulse propagation through the human chest wall is modeled here using the equations of motion for a fluid of variable sound speed, density, and absorption. The tissue is assumed motionless except for small acoustic perturbations. Absorption is included using an adaptation of the Maxwell solid model,¹³ in which all absorption effects are represented by a single relaxation time. This assumption results in frequency-independent absorption characteristics. Equivalent treatments of tissue-dependent absorption have been employed by a number of previous models for ultrasonic propagation in biological tissues.^{14–16} For such a fluid, the linearized equations of mass conservation, momentum conservation, and state can be combined to obtain the first-order, two-dimensional, coupled propagation equations,

$$\frac{\partial p(x,y,t)}{\partial t} + \rho(x,y) c(x,y)^2 \nabla \cdot \mathbf{v}(x,y,t) = -\alpha(x,y) p(x,y,t), \quad (1)$$

$$\rho(x,y) \frac{\partial \mathbf{v}(x,y,t)}{\partial t} + \nabla p(x,y,t) = 0. \quad (2)$$

Here, $p(x,y,t)$ is the acoustic perturbation in fluid pressure, $\mathbf{v}(x,y,t)$ is the vector acoustic particle velocity, $\rho(x,y)$ is the ambient density, $c(x,y)$ is the ambient sound speed, and $\alpha(x,y)$ is an absorption coefficient that is equivalent to the inverse of a spatially-dependent relaxation time $\tau(x,y)$.

The absorption coefficient α , defined as a real quantity, is related to the energy lost per unit length as follows. The propagation equations (1) and (2) lead, for plane-wave propagation of the form $p = e^{i(kx - \omega t)}$, to the dispersion relation

$$k = \frac{\omega}{c} \sqrt{1 + \frac{i\alpha}{\omega}}, \quad (3)$$

where k is the complex wavenumber, ω is the (real) radial frequency $2\pi f$, and c is the (real) sound speed. The imaginary part of the wavenumber k is the absorption in nepers per unit length. Thus, the absorption parameter α can be obtained by a numerical solution of the equation

$$\text{Im}[k] = \frac{\text{loss (dB/length)}}{20 \log_{10}(e)} = \text{Im} \left[\frac{\omega}{c} \sqrt{1 + \frac{i\alpha}{\omega}} \right]. \quad (4)$$

Solution of Eq. (4) results in wavenumbers having a real part that differs from ω/c . Since this discrepancy is less than 1% over the range of tissue properties employed in the present study, use of absorption coefficients computed from Eq. (4) does not significantly affect propagation characteristics except by adding the specified absorption.

Equations (1) and (2) were solved numerically using the finite-difference time-domain (FDTD) method described in Refs. 10 and 17. This method is a two-step MacCormack predictor–corrector algorithm that is fourth-order accurate in space and second-order accurate in time. The computations employed a spatial step size of 15 points per wavelength at the pulse center frequency of 2.3 MHz. Time step sizes were computed using a Courant–Friedrichs–Lewy number of 0.25. Further details on this class of finite difference algorithms can be found in Refs. 18–20.

The initial condition was chosen to model the experimental configuration in Ref. 8, in which a spatially broad, nearly planar wavefront was emitted from a wideband, pulsed, unfocused source far from the tissue layer. The initial wavefront was represented in the present simulation as a plane wave pulse propagating in the $+y$ direction:

$$p(x,y,0) = -\sin[k_0(y-y_0)] e^{-(y-y_0)^2/(2\sigma^2)},$$

$$u(x,y,0) = 0,$$

and (5)

$$\mathbf{v}(x,y,0) = \frac{p(x,y,0)}{\rho c},$$

where the wavenumber k_0 is equal to $2\pi f_0/c$ for a center frequency of f_0 , σ is the Gaussian parameter of the pulse temporal envelope, and u and v are the x and y components of the vector acoustic particle velocity \mathbf{v} . The spatial Gaussian parameter σ was chosen to simulate the bandwidth of the pulse used in the experiments, as discussed below in the Method section.

The computational configuration is analogous to that described in Ref. 10. The domain of computation is two-dimensional, with the y direction taken to be parallel to the direction of propagation and the x direction parallel to the initial wavefront. As in Ref. 10, periodic boundary conditions were applied on the domain edges that were parallel to the direction of propagation, while radiation boundary conditions were applied on the edges perpendicular to the direction of propagation.

II. METHOD

This study employed six chest wall specimens obtained during the autopsies of four different donors between 79 and 85 years of age at death. One specimen (4L) was from a white female, while the others were from white males. After the specimens were obtained, they were stored unfixed at -20°C and thawed when needed for study. Wavefront distortion measurements were made on these and other specimens as part of a study described in Ref. 8. In those measurements, 2.3 MHz ultrasonic pulses generated by a 0.5-in. piston transducer propagated through individual chest wall specimens immersed in a 37°C water bath and the transmitted pulses were measured by a 96-element broadband cardiac array scanned to synthesize a two-dimensional aperture. Statistics describing wavefront distortion, including arrival time fluctuations, energy level fluctuations, and wave shape distortion, were computed for the measured pulses.

For the present study, six of the previously measured specimens were cut into ~ 7 -mm thick cross sections using the technique described in Ref. 10. The slices were then fixed and stained with a modified Gomori's trichrome stain according to the procedure detailed in Ref. 21, so that tissue types could be distinguished. This stain colored muscle tissue red and connective tissue blue while leaving the fat its natural color. Calcified tissue, including bone and cartilage in the current specimens, was not differentially stained by this technique, but the natural contrast between bone, cartilage, and marrow was sufficient to allow tissue mapping. Full-color 300 d.p.i. images of the cross sections were created by placing each stained tissue cross section directly onto the surface of a flatbed digital scanner. Image editing packages (Adobe Photoshop, Version 3.0, and the Gnu Image Manipulation Program, Version 1.0) were used to manually segment the cross sectional images, i.e., to map the images into regions that corresponded to one of six media. The media were water (representing water external to specimens or blood inside blood vessels), fat (including subcutaneous fat, fat interlaced within muscle layers, and marrow), muscle, connective tissue (including skin, septa, and fasciae), cartilage, and bone (including cortical bone and trabeculae within cancellous bone).

The nomenclature employed here for the cross sections corresponds to that of Ref. 8 for the whole specimens from which the cross sections were taken; each cross section is identified by a donor number together with "L" or "R" to indicate whether the corresponding specimen was taken from the left or right side of the breastplate. Additional numbers were used in Ref. 8 to indicate the intercostal space used in each measurement; here, lower-case letters are used to indicate independent acoustic paths. Wavefront distortion measurement results from four of the specimens employed here (4L, 5L, 7L, and 7R) were reported in Ref. 8. Distortion statistics for specimens 8L and 8R were not presented in Ref. 8 because of limited acoustic windows. No new measurements were made for the present study; statistics describing measured distortion are taken directly from Ref. 8.

The six segmented tissue maps are shown in Fig. 1. All of the cross sections contain a layer of septated subcutaneous fat below the skin. Most of the cross sections also include a layer composed primarily of the major pectoral muscles and their connective fasciae above the ribs. Between the ribs are regions of muscle (internal intercostal and external intercostal groups) interlaced with fat. In some cases, additional thin layers of fat between muscle layers are apparent. Cross sections 4L and 7R are cut along the intercostal spaces parallel to the ribs, so that in each a wide cross section of soft tissue appears. Cross sections 5L, 7L, and 8L are cut perpendicular to the ribs, so that each contains soft-tissue acoustic paths with width equal to the width of the corresponding intercostal spaces. Cross section 8R is cut perpendicular to the sternum at a location of large curvature in the ribs, so that the ribs are diagonally sectioned. Several blood vessels appear in cross sections 4L, 7L, 7R, and 8R; the largest of these is the internal mammary artery.

The basic structure of the cross sections is consistent with standard descriptions of chest wall anatomy.^{22,23} Ribs appear in each cross section; each rib is composed of a "costal cartilage" near the sternum (shown in most of the cross sections considered here) attached to a "true rib" (composed primarily of cancellous bone) at the edge farther from the sternum. In the cross sections considered here, the costal cartilages are primarily composed of calcified cartilage, surrounded by a thin layer of cortical bone (solid, dense bone with microscopic porous structure), which in turn is surrounded by the periosteum, a thin membrane of connective tissue. Cross sections 7L and 7R also appear to contain a small amount of cortical bone in the central portion of the ribs. This phenomenon may be associated with advanced calcification known to occur in aging humans.²⁴ Cancellous bone, composed of thin trabeculae that form macroscopic cells filled with marrow, is seen in all the ribs of cross section 5L, which was taken at a distance farther from the sternum so that the true ribs, rather than the costal cartilages, were included in this cross section. Some cancellous bone is also apparent within portions of the ribs of cross sections 4L and 8R. In each case, the cancellous bone is surrounded by a thin layer of cortical bone and by the periosteum. A portion of the sternum, composed of cancellous bone surrounded by cortical bone, is visible at the left side of cross section 4L.

The density and sound speed grids needed for the finite-

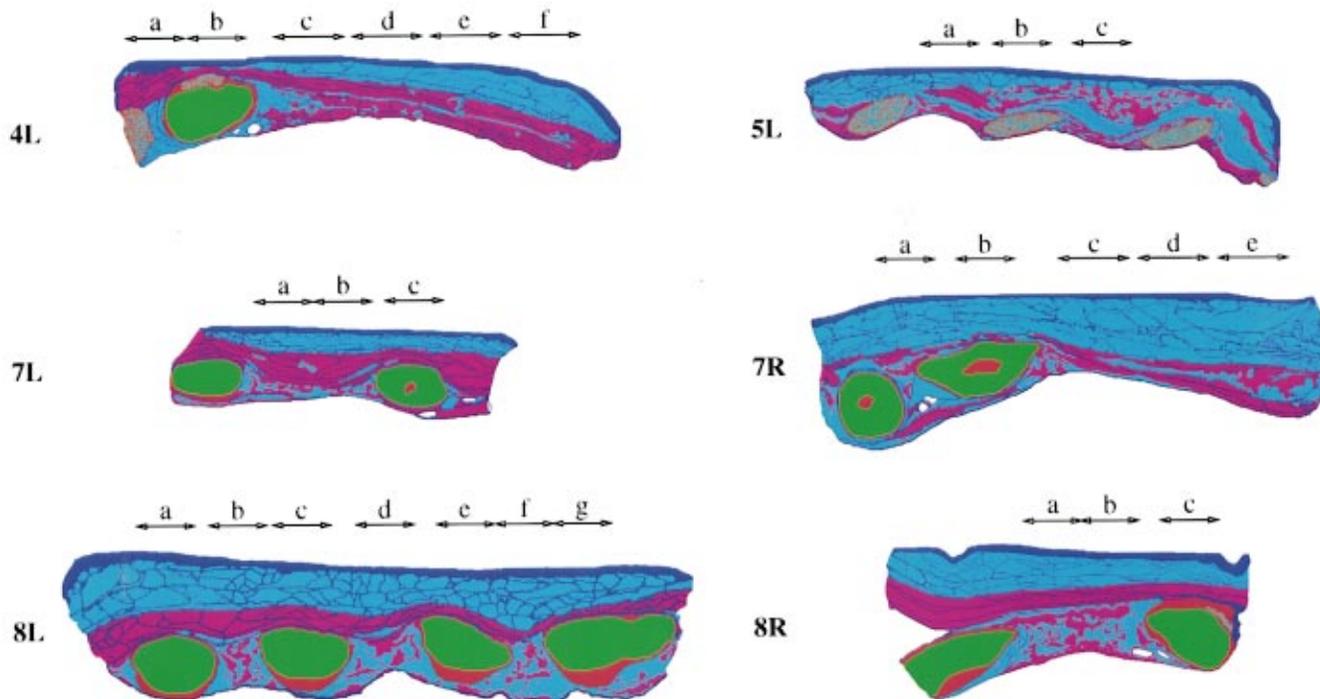


FIG. 1. Chest tissue maps used in simulations. In each map, blue denotes skin and connective tissue, cyan denotes fat, purple denotes muscle, orange denotes bone, and green denotes cartilage. Blood vessels appear as small water-filled (white) regions. Simulated apertures are indicated using lower-case letters for each cross section; the letters correspond to the acoustic path labels used throughout, while the length of the arrow beneath each letter corresponds to the extent of the simulated aperture. Smaller arrows indicate 55-element (11.60-mm) apertures while large arrows indicate 68-element (14.28-mm) apertures.

difference computation were created by mapping regions of the segmented tissue images to reference density and sound speed values for the five tissue types and water. The water sound speed and density employed are those of pure water at body temperature (37.0 °C).^{25,26} Sound speeds for muscle and fat were obtained by averaging values for human tissues given in Refs. 27 and 28. A representative sound speed for connective tissue was determined using an empirical formula relating collagen content to ultrasonic sound speed²⁹ together with a measured value for the collagen content of human skin.³⁰ The sound speed employed for bone was obtained from an average of values reported in Ref. 31 for longitudinal-wave propagation in human cortical bone. The sound speed used here for cartilage is that given in Ref. 32 as quoted in Ref. 27. Density values for soft tissues were determined from Ref. 33 by averaging values reported for adipose tissue, skeletal muscle, and skin, respectively. Density values employed for bone and cartilage are average values from Ref. 31.

Absorption values were determined from attenuation measurements summarized in Ref. 27 for human fat at 37 °C, human bicep muscle at 37 °C, human skin at 40 °C, human and bovine cartilage at 23 °C, and human skull (temperature not reported). Attenuation values reported at other ultrasonic frequencies were interpolated (or, for the skull data, extrapolated) to obtain values for 2.3 MHz (corresponding to the pulse center frequency employed here and in Ref. 8) assuming a linear dependence of attenuation on frequency. This assumed linear dependence is a simplifying approximation; tissue measurements show that attenuation varies approximately as $\alpha \propto f^\beta$, where β is typically between 0.9 and 1.5

for various human soft tissues.³⁴ The absorption for water was estimated by extrapolating frequency- and temperature-dependent absorption values summarized in Ref. 35 to 2.3 MHz and 37.0 °C. The values of tissue parameters employed in the present study are given in Table I.

The finite-difference program was employed to compute propagation of a plane wave pulse through each scanned cross section from the skin to the peritoneal membrane, mimicking the propagation path employed in the distortion measurements of Ref. 8. The spatial step size of the finite-difference grid was chosen to be 0.0442 mm, or 1/15 wavelength in water at the center frequency of 2.3 MHz. The temporal step size was chosen to be 0.00725 μ s, for an optimal Courant–Friedrichs–Lewy number $c\Delta t/\Delta x$ of 0.25.²⁰ The Gaussian parameter σ of the source pulse was chosen to be 0.4766 mm in accordance with the experimentally measured pulse bandwidth (for pulses transmitted through a water path) of 1.2 MHz. A visual comparison confirmed that the

TABLE I. Assumed physical properties for each tissue type employed in the simulations.

Tissue type	Sound speed (mm/ μ s)	Density (g/cc)	Absorption (dB/mm)
Water	1.524	0.993	0.0007
Fat	1.478	0.950	0.12
Muscle	1.547	1.050	0.21
Connective	1.613	1.120	0.37
Cartilage	1.665	1.098	0.97
Bone	3.540	1.990	4.37

simulated pulse closely matched the measured pulses in shape and length.

Each simulation was performed on a workstation with 128 MB of random-access memory. Finite-difference grids on the order of 1500×1000 points were employed. At each time step, the wave field was updated on a grid subset chosen to include the entire support of the acoustic wave but to exclude quiescent regions. The entire pressure field was saved as a raster image at intervals of $0.725 \mu\text{s}$ for later visualization. The computation time for each simulation was on the order of five hours.³⁶

Signals were recorded for $8.62 \mu\text{s}$ at a sampling frequency of 138 MHz by simulated apertures with dimensions close to those in the experimental study of Ref. 8. Positions of all simulated apertures employed are sketched in Fig. 1. The simulation of receiving elements was performed by integrating the locally-computed pressure over the element pitch of 0.21 mm. For cross sections cut parallel to the ribs, the simulated apertures contained 68 elements for an aperture width of 14.28 mm. For cross sections cut perpendicular to the ribs, 55 simulated elements were used to form 11.55 mm apertures. Element directivity effects were implicitly incorporated by the integration of acoustic fields over the width of each element; the resulting directivity functions correspond to those for an idealized line element of width 0.21 mm.

A one-dimensional version of the reference waveform method^{10,37} was used to calculate the arrival time of the pulse at each receiving position in the simulation data. In this method, the relative arrival time of each received waveform is computed by cross-correlation with a reference waveform. The arrival time fluctuations across the receiving aperture are then calculated by subtracting a linear fit from these calculated arrival times, and the root-mean-square value of these fluctuations is computed. Energy level fluctuations in the data were calculated by summing the squared amplitudes of each waveform over a $2.4\text{-}\mu\text{s}$ window that isolated the main pulse, converting to decibel units, and subtracting the best linear fit from the resulting values. As for polynomial fits previously employed in wavefront distortion measurements,⁸ the purpose of the linear fit removal in each case was to compensate for gross changes in tissue thickness across the array. Variations in pulse shape across the aperture were evaluated using the waveform similarity factor;³⁷ this quantity, which can be considered a generalized cross-correlation coefficient, has a maximum of unity when all received waveforms are identically shaped.

To test the frequency dependence of chest wall wavefront distortion, propagation through eight portions of specimens, each containing only soft tissue, was also computed for wavefronts having center frequencies of 1.6 and 3.0 MHz. In each case, the initial wavefront was chosen to have the same temporal envelope as above. The absorption coefficient at these frequencies for each tissue type was extrapolated from the value employed at 2.3 MHz using the assumption that absorption depended linearly on the center frequency. The spatial and temporal sampling rates were also varied in inverse proportion to the pulse center frequency. All runs were otherwise identical in configuration and processing to those described above.

TABLE II. Statistics of simulated wavefront distortion caused by thirteen soft tissue paths within chest wall cross sections. The "Path" column shows the cross section label and aperture letter for each path; these labels correspond to those shown in Fig. 1. The statistics shown include the average specimen thickness for the tissue path considered, rms values and correlation lengths (CL) of the arrival time fluctuations (ATF) and the energy level fluctuations (ELF), the waveform similarity factor (WSF), and the total attenuation.

Path	Thickness (mm)	ATF		ELF		WSF	Attenuation (dB)
		rms (ns)	CL (mm)	rms (dB)	CL (mm)		
4L-c	15.4	32.0	0.60	1.98	1.68	0.981	5.62
4L-d	12.7	10.0	2.58	0.46	1.23	0.999	4.08
4L-e	16.0	10.0	1.37	1.61	1.74	0.998	5.26
4L-f	17.0	17.3	2.48	0.92	1.61	0.999	5.33
5L-a	11.0	11.6	0.95	1.51	1.13	0.991	4.29
5L-c	15.0	14.8	1.03	1.15	1.19	0.996	5.01
7L-a	16.2	16.8	2.64	0.95	1.29	0.999	5.46
7L-b	14.9	22.5	2.66	1.19	1.61	0.998	4.91
7R-c	17.7	17.4	1.77	2.52	2.07	0.997	5.83
7R-d	21.0	8.3	1.10	0.85	1.79	0.999	7.07
7R-e	24.7	13.7	1.37	1.06	1.62	0.997	8.69
8R-a	23.8	26.6	1.78	2.58	1.40	0.992	7.76
8R-b	22.2	29.9	1.44	1.95	1.11	0.989	6.09
Mean	17.5	17.8	1.67	1.44	1.50	0.995	5.80
St. Dev.	4.2	7.8	0.71	0.66	0.30	0.005	1.33

III. RESULTS

Simulated wavefront distortion results for 13 soft tissue paths (i.e., paths in which wavefront distortion was not significantly influenced by the ribs) are shown in Table II. These results indicate that soft tissue paths cause a wide range of wavefront distortion effects depending on the specific morphology of each path. For instance, path 7R-c causes arrival time and energy level fluctuations that are more than twice the magnitude of those caused by the adjacent path 7R-d. This difference is thought to arise from morphological features, including muscle tissue with interlaced fat and a large amount of connective tissue, of the tissue within path 7R-c. Also notable is that the specimen thickness does not closely correspond to variations in distortion. The largest rms arrival time fluctuation and lowest waveform similarity factor, for example, are caused by path 4L-c, which has an average thickness less than the mean for all the tissue paths.

Wavefront distortion statistics for the 13 soft tissue paths are graphically summarized in Fig. 2 together with corresponding statistics for all of the soft tissue measurements reported in Ref. 8. This comparison indicates that wavefront distortion caused by soft tissues in the chest wall simulations is comparable to measured distortion. Arrival time fluctuations and energy level fluctuations for simulated distortion are slightly less than measured values, but mean values of both fluctuations for the simulations fall well within one standard deviation of the corresponding mean fluctuation for the measurements. The waveform similarity factor, however, is substantially higher for simulations than measurements, indicating that simulated waveforms were distorted considerably less than measured waveforms. Correlation lengths for the simulated distortions are somewhat less

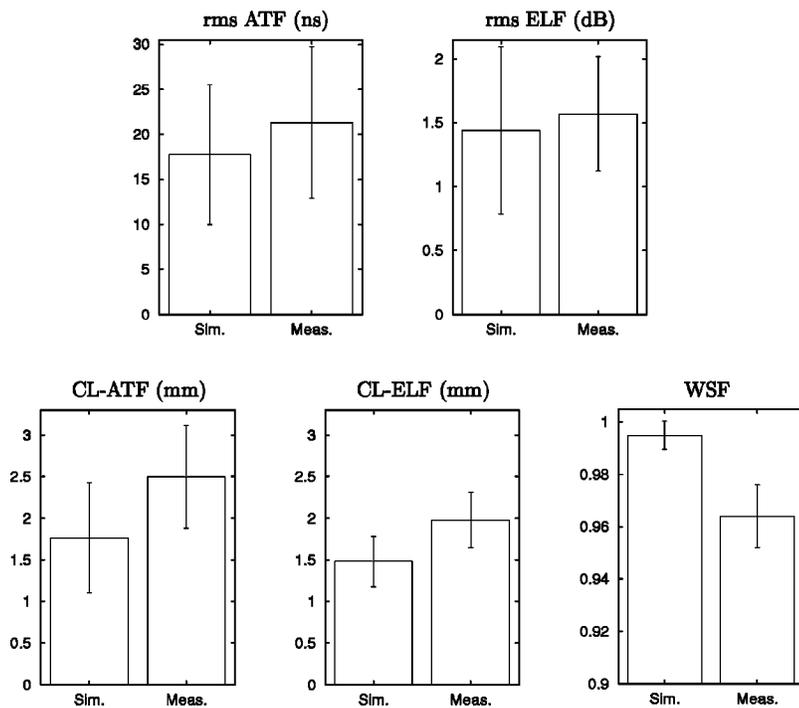


FIG. 2. Summary of distortion statistics for soft tissue paths. The bar chart shows mean values of the rms arrival time fluctuations (ATF), rms energy level fluctuations (ELF), correlation lengths (CL) of these fluctuations, and waveform similarity factors (WSF) for the simulations performed in the present paper and the experiments reported in Ref. 8. Error bars indicate a range of plus or minus one standard deviation from the mean.

than measured values. However, consistent with measurements, the mean correlation length of the simulated arrival time fluctuations is greater than that for the simulated energy level fluctuations.

As in Ref. 8, rib structures were found to cause much more distortion than soft tissue alone. The varied nature of distortion caused by rib effects is illustrated in Fig. 3, which shows three representative sets of measured signals for specimen 8L. These measurements were made during the study reported in Ref. 8. The first panel shows 96 adjacent measured signals, along the array direction (approximately parallel to the ribs) for propagation through a tissue path within an intercostal space. The signals are not severely distorted; secondary arrivals are discernible, but are of lower amplitude than the main arrival. The second panel shows 96 measured signals for an elevation over a rib. Here, all signals are severely distorted. Multiple arrivals, as well as high-amplitude spatially-random fluctuations, are seen. The third panel shows 50 measured signals along the elevation direction (perpendicular to the ribs), centered over the soft tissue between the ribs. Here, the main wavefront is curved rather than straight, an additional arrival behind the main wavefront is seen, and portions of the signals from over the ribs (at both edges of the panel) are advanced relative to the signals from the central soft tissue region.

The present simulations allow more detailed qualitative and quantitative investigation of rib effects than were possible from the previous measurements. Propagation through two rib-influenced paths is illustrated in Figs. 4 and 5, in which computed ultrasonic pulses are superimposed on portions of the tissue maps from Fig. 1. (Similar visualizations of propagation through soft human body wall tissue were shown in Ref. 10.)

Figure 4 shows propagation through a thin rib, composed chiefly of cancellous bone, in cross section 5L (corresponding approximately to path 5L-b). A strong reflection

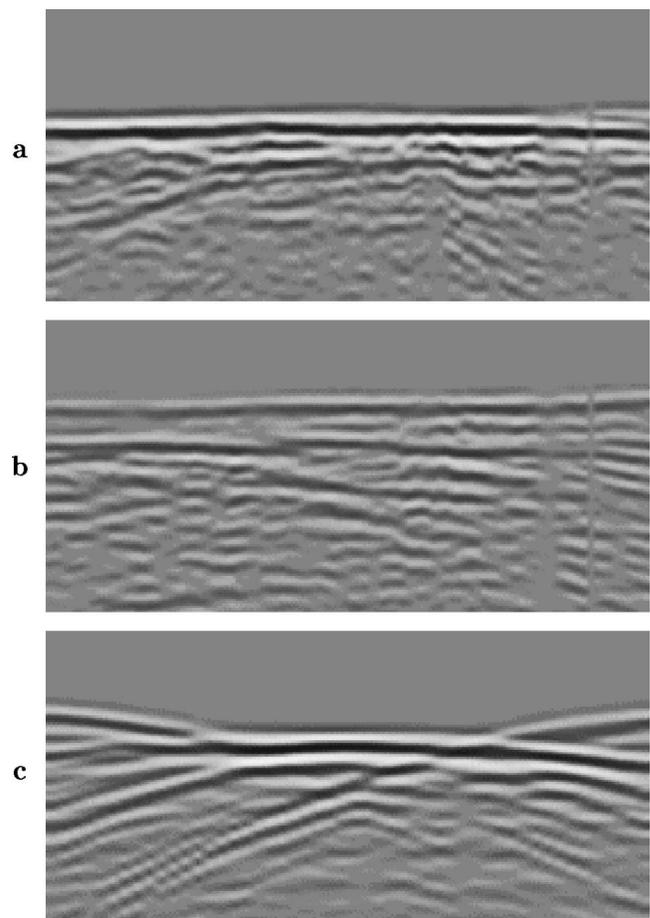


FIG. 3. Measured waveforms for three propagation paths in specimen 8L. Each panel shows received waveforms on a bipolar logarithmic gray scale with a dynamic range of 40 dB. The horizontal range shown in each panel is 20 mm and the vertical range shown is $6.4 \mu\text{s}$. (a) Tissue path between two ribs, in azimuth direction (parallel to ribs). (b) Path including a rib, azimuth direction. (c) Tissue path including intercostal space between two ribs, elevation direction (perpendicular to ribs).

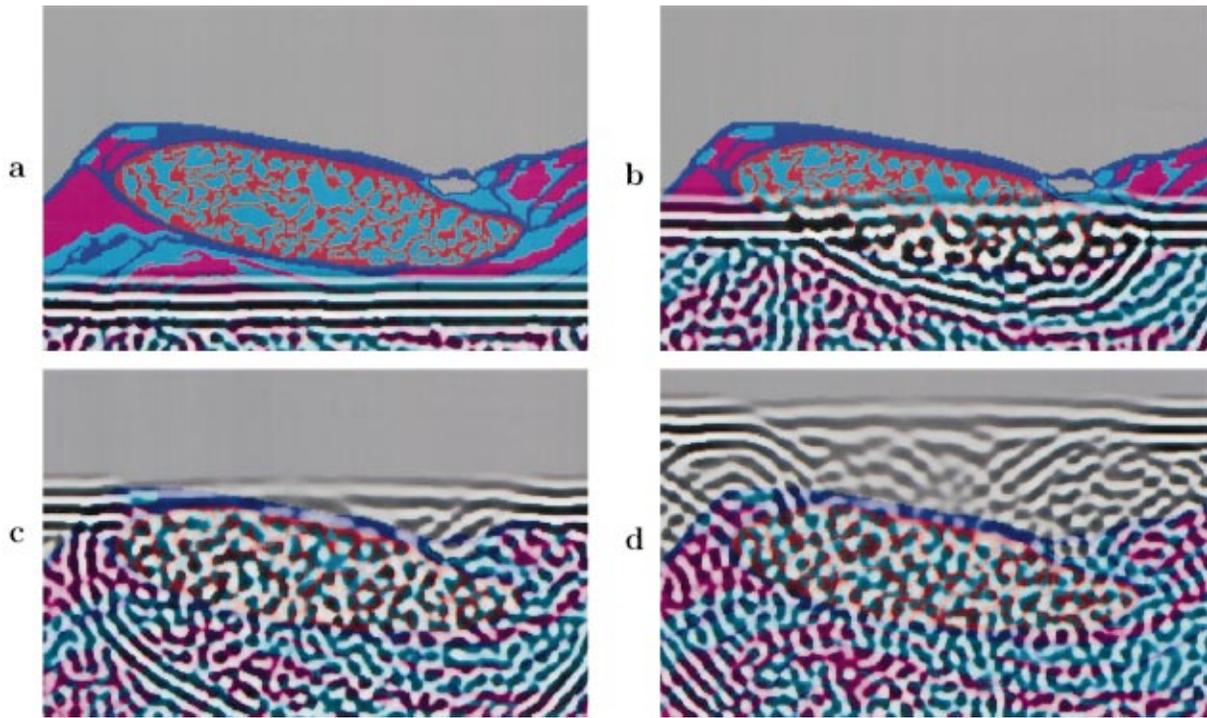


FIG. 4. Simulated propagation through the central rib in cross section 5L (path 5L-b). Panels (a)–(d) show instantaneous acoustic pressure fields at successive intervals of $2.17 \mu\text{s}$. Each panel shows an area that spans 20.32 mm horizontally and 14.58 mm vertically. Logarithmically compressed wavefronts are shown on a bipolar scale with black representing minimum pressure, white representing maximum pressure, and a dynamic range of 57 dB.

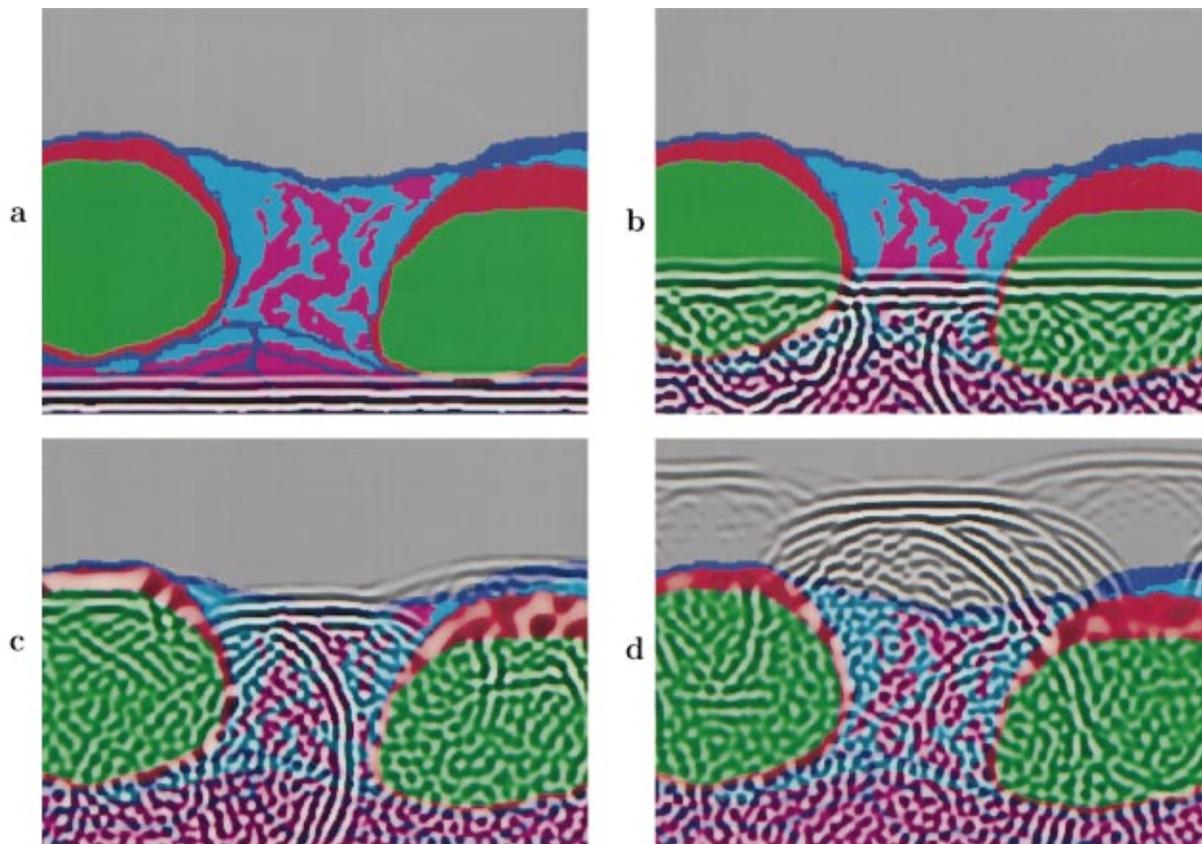


FIG. 5. Simulated propagation through an intercostal space in cross section 8L (path 8L-b). Panels (a)–(d) show instantaneous wavefields at successive intervals of $3.62 \mu\text{s}$. Each panel shows an area that spans 28.27 mm horizontally and 21.20 mm vertically. Wavefronts are shown using the same format as in Fig. 4.

occurs at the first interface between bone and soft tissue, removing a substantial amount of energy from the main wavefront. The small, high-contrast trabeculae within the rib cause considerable scattering, as can be observed in panel (b) of Fig. 4. The scattering causes random fluctuations behind the main wavefront; these fluctuations somewhat resemble those seen in the measured data of Fig. 3(b). After passing through the rib, as seen in panels (c) and (d) of Fig. 4, the central portion of the wavefront shows substantial attenuation and distortion. However, the average arrival time of the wavefront is not greatly changed by propagation through the rib, but is advanced by only about one-half period. This phenomenon apparently occurs because the influence of the “slow” marrow (modeled here as fat) counteracts the influence of the “fast” trabeculae. Noteworthy is that the predominant ultrasonic wavelength has increased after propagation through the rib, so that the effective center frequency of the wavefront has been lowered. Since the absorption model used in the present study includes only frequency-independent absorption, the loss of short-wavelength components in this simulation results only from frequency-dependent scattering caused by the trabeculae.

Propagation within path 8L-b, which includes two larger ribs and the corresponding intercostal space, is illustrated in Fig. 5. At the position of the cross section, these ribs are composed primarily of cartilage and surrounded by a thin layer of cortical bone. Since the cartilage and bone of these ribs are modeled as homogeneous structures, small-scale scattering within these tissues did not occur in this simulation. Instead, the wavefront is reflected from interfaces between cartilage, bone, and soft tissue.

The visualization shown in Fig. 5 provides physical reasons for all the rib-related distortion phenomena seen in the measured data of Fig. 3(c). The wavefronts propagating through the ribs show greater attenuation than that in Fig. 4, both because of the high absorption of the ribs and because of the reflections noted above. These wavefronts are also advanced relative to the wavefront propagating through the intercostal space, because of the higher sound speed of both bone and cartilage. The wavefront propagating through the intercostal space is distorted somewhat by the inhomogeneous soft tissue path, as can be observed in panels (b) and (c). However, much greater distortion results from interaction between the wavefront and the ribs. A rightward-propagating reflection, seen in panels (b) and (c), combines with the main wavefront in panel (d) to result in severe distortion at the right side of the central wavefront. A leftward-propagating reflection from the other rib is also apparent. Furthermore, diffraction from the edges of the ribs results in large curvature of the soft tissue wavefront.

Distortion and attenuation statistics for a variety of simulations employing rib-influenced paths are shown in Table III. Footnotes in Table III indicate physical causes of distortion present within each path. A variety of distortion and attenuation mechanisms are illustrated. Propagation through small intercostal spaces (paths 4L-a, 8L-b, 8L-f, and 7R-a) causes diffraction effects that introduce substantial curvature into the wavefront, as seen in Fig. 5. This large-scale wavefront curvature is associated with large arrival

TABLE III. Statistics of simulated wavefront distortion caused by fourteen tissue paths including rib structures. The footnotes associated with the label for each path indicate morphological features and physical phenomena that affected the wavefront distortion computed for that path. The format is analogous to that in Table II.

Path	Thickness (mm)	ATF		ELF		WSF	Attenuation (dB)
		rms (ns)	CL (mm)	rms (dB)	CL (mm)		
4L-a ^{a,b,c,d}	21.0	260.3	3.00	2.58	2.72	0.968	15.33
4L-b ^{b,c}	17.6	161.9	1.90	4.16	1.49	0.641	43.35
5L-b ^b	14.2	92.5	0.69	3.06	1.92	0.775	26.87
7L-c ^{c,e}	17.8	47.2	1.58	5.33	2.04	0.958	19.66
7R-a ^{a,c,d}	30.4	123.1	2.12	3.80	1.78	0.960	16.57
7R-b ^{c,e}	24.3	165.6	2.71	6.88	2.07	0.274	43.06
8L-a ^c	25.3	113.9	1.18	7.75	2.29	0.907	32.44
8L-b ^{a,d}	22.8	109.7	2.05	3.43	1.22	0.974	10.28
8L-c ^c	28.8	134.0	2.75	3.04	1.57	0.944	40.47
8L-d ^d	23.6	78.9	0.64	3.06	1.55	0.950	6.78
8L-e ^c	26.4	208.8	1.91	3.62	1.50	0.810	44.27
8L-f ^{a,d}	28.5	169.9	1.79	5.02	1.95	0.916	10.70
8L-g ^c	27.6	210.8	1.40	3.36	1.35	0.892	44.22
8R-c ^{b,c}	24.9	81.4	2.08	2.76	1.25	0.962	44.32

^aSmall intercostal spaces.

^bCancellous bone.

^cCortical bone and cartilage.

^dStrong rib reflections.

^eCortical bone within cartilage.

time fluctuation values although the wavefronts generally appear to be locally smooth. Interference between directly-transmitted and rib-reflected wavefronts (paths 4L-a, 8L-b, 8L-d, 8L-f, and 7R-a) introduces arrival time, energy level, and waveform distortion substantially greater than that for soft tissue paths without ribs. Propagation through cancellous bone (paths 4L-a, 4L-b, 5L-b, and 8R-c) results in considerable attenuation and large waveform distortion, while propagation through cortical bone and cartilage (paths 4L-a, 4L-b, 8L-a, 8L-c, 8L-e, 8L-g, 7L-c, 7R-a, 7R-b, and 8R-c) results in even larger attenuation but smaller distortion. Where bone is embedded within cartilage (paths 7L-c and 7R-b), additional scattering also occurs. For the path including a large bone inclusion (path 7R-b), this scattering results in an extremely high energy level and waveform distortion.

Computed frequency-dependent wavefront distortion statistics are summarized in Fig. 6. Tissue paths used for these computations, none of which include rib structures, are those labeled 4L-d, 4L-f, 5L-a, 5L-c, 8R-a, 8R-b, 7L-a, and 7L-b in Fig. 1. The results shown in Fig. 6 indicate that arrival time fluctuations, energy level fluctuations, and waveform distortion all become more severe with increasing pulse frequency. The most dramatic change is in the energy level distortion; on average, the rms energy level fluctuations for the 3.0-MHz signals are 2.3 times those for the 1.6-MHz signals. Correlation lengths of both arrival time and energy level fluctuations decrease with frequency, so that the predominant length scales of ultrasonic wavefront distortion are seen to decrease with the ultrasonic wavelength. As with the rms distortion statistics, the most dramatic frequency-dependent change is in the energy level fluctuations. Still, even the high-frequency pulses here show substantially

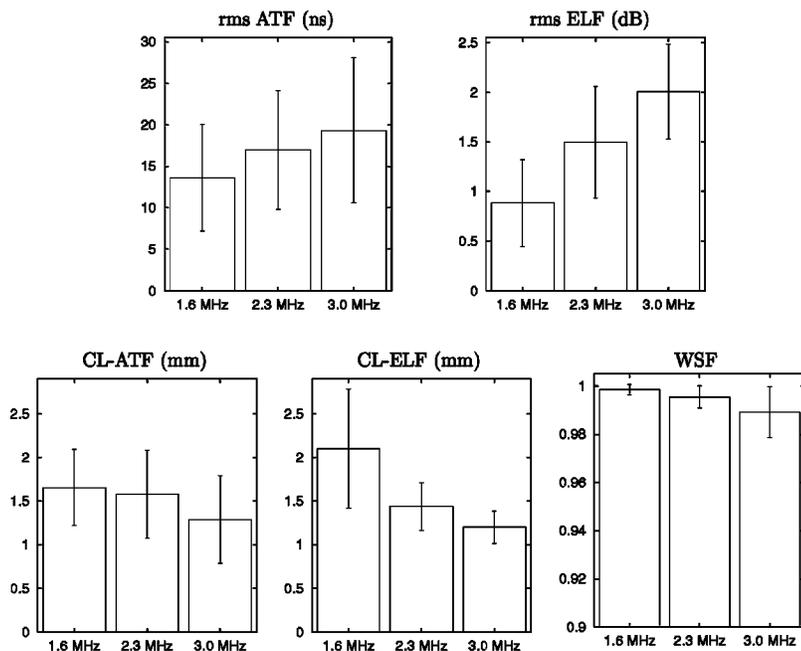


FIG. 6. Summary of simulated frequency-dependent distortion results. Mean rms arrival time fluctuations (ATF), energy level fluctuations (ELF), correlation lengths (CL) of these fluctuations, and waveform similarity factors (WSF) are shown for each of the three pulse frequencies investigated. Error bars indicate a range of plus or minus one standard deviation from the mean.

smaller distortion than that previously observed in experiments and simulations for the human abdominal wall.^{10–12,38}

IV. DISCUSSION

As with earlier simulations of propagation through tissue,^{10,12} the current study shows qualitative agreement with measured wavefront distortion results for similar specimens.⁸ However, the accuracy of the present model is limited by simplifications of true tissue structure. In particular, the computational model here does not account for property variations within tissue types, tissue microstructure, or three-dimensional tissue structure. Each of these simplifications limits the ability of the present model to precisely mimic experimentally measured ultrasonic wavefront distortion. These limitations are discussed, with respect to soft tissues, in Ref. 10.

The modeling of ribs adds additional complication. In the current study, individual trabeculae were assumed to be composed of tissue having properties identical to cortical bone, an assumption known as Wolff's hypothesis.³⁹ The validity of this hypothesis has been questioned,^{40,41} however, measured elastic properties of individual trabeculae vary widely^{40,41} and recent work⁴² has provided support for Wolff's hypothesis. Thus, the properties employed here for trabecular bone can be regarded as reasonable order-of-magnitude estimates. Likewise, the modeling of marrow as fat tissue is a simplifying assumption that may have limited validity, although available data suggest that the density and sound speed of marrow are close to those for other adipose tissues.³¹ In addition, the present model for cartilage is based on measurements of normal cartilage, while the cartilage present in the specimens employed here was calcified due to the age of the donors. However, density measurements made on eight representative samples of calcified cartilage (two from specimen 7R, four from specimen 1R,⁸ and two from an unused specimen) resulted in an average density of 0.00111 kg/m³, which is different by only 1% from the den-

sity assumed here. Since sound speed in calcified tissue has been empirically shown to be directly related to density,^{43,44} this small change in density suggests that the acoustic properties of the calcified cartilage in our specimens is close to that for normal cartilage.

The computations reported here model the chest wall as a fluid of variable sound speed, density, and compressibility. This model implicitly neglects shear wave propagation. The neglect of shear waves in soft tissues is believed to be justified because the absorption of shear waves in soft tissues is much greater than absorption of longitudinal waves.^{45,46} In calcified tissues, however, significant shear waves are known to be generated.^{47,48} In the current scattering configuration, some shear waves are likely generated wherever the rib surface is far from parallel to the wavefront. However, since shear wave absorption has been found to be somewhat larger than longitudinal wave absorption for ultrasonic propagation in bone,⁴⁷ the significance of shear-wave propagation within bone on transmitted ultrasonic wavefronts is questionable. For this reason, omission of nonlongitudinal waves in the present study, as in another computational study of ultrasonic scattering from bone,⁴⁹ is believed to be justified; however, further study would be required to confirm this assumption.

The absence of frequency-dependent absorption is a possible source of error in the present estimates of total tissue attenuation, energy level fluctuations, and waveform distortion. However, since absorption in tissue increases approximately linearly with frequency, lower absorption for frequency components below the pulse center frequency would nearly cancel higher absorption for frequency components above the center frequency, so that the average absorption incurred by a wideband pulse should still be computed with fair accuracy. For this reason, the absence of frequency-dependent absorption in the calculations reported here is not considered to be a significant source of error in the computed attenuation or energy level fluctuation curves. Still, the inclusion of frequency-dependent absorption would result in

additional waveform distortion effects. The lack of this effect is a likely reason for the lower waveform distortion (higher waveform similarity factors) obtained from simulations as compared to measurements. However, the absence of frequency-dependent absorption effects allowed frequency-dependent scattering effects to be clearly quantified separately from absorption effects.

Although the simulations were planned to match the measurements of Ref. 8 closely, a number of differences remain. The most important of these, as discussed in Ref. 10, is that the simulations were performed using a two-dimensional tissue model while the measurements were inherently three-dimensional. Other differences include details of the source waveform and wavefront shape, variations in the specimen orientations and the regions interrogated, and variations in the distance between the specimen and the real or simulated receiving aperture. All of these differences could contribute to discrepancies between measurements and simulations.

In general, most of the simplifying assumptions in the present tissue model are likely to result in underestimation of wavefront distortion produced by the human chest wall. The incorporation of tissue microstructure, spatially-dependent acoustic properties for each tissue type, shear wave propagation in bone and cartilage, three-dimensional propagation, and frequency-dependent absorption could all result in greater spatial and temporal variations in the propagating acoustic fields, so that these features could produce simulated distortion with characteristics closer to measurements. For this reason, distortion statistics computed using the present tissue model should be interpreted as lower limits for the statistics of distortion occurring in real chest wall tissue.

Additionally, some of the discrepancy between simulated and measured distortion may be explained by the non-uniform characteristics of the receiving transducer employed in the measurements.⁸ The water-path measurements reported in Ref. 8 show arrival time fluctuations (mean 2.21 ns) and energy level fluctuations (mean 0.36 dB); although small, these fluctuations are comparable to the difference between the average measured and simulated fluctuations. Thus, compensation for arrival time and energy level fluctuations due to transducer irregularities could reduce measured distortion to levels closer to the simulations. Also, the waveform similarity factor for water path measurements was 0.991,⁸ which indicates greater waveform distortion than the average value of 0.995 computed here for soft tissue paths. Thus, compensation of the measured data for transducer impulse-response variations could raise the measured waveform similarity factor to a value in closer agreement with simulations.

Previous experimental measurements of wavefront distortion caused by the human chest wall⁸ have suggested that distortion caused by chest wall soft tissues is less severe than that caused by the human abdominal wall.^{11,38} This difference has been observed to occur even though average specimen thicknesses were comparable in chest wall⁸ and abdominal wall^{11,38} measurements. The present results provide support for these results; arrival time and energy level distortion by the chest wall was found here to be smaller than that produced by the abdominal wall in previous simulation

studies.^{10,12} For the simulations, this difference may be partially explained by the fact that the chest wall specimens employed here are thinner on average (mean thickness 17.5 mm) than the abdominal wall cross sections employed in Refs. 10 and 12 (mean thickness 26.7 mm). Another possible partial explanation is that the pulse center frequency employed in abdominal wall measurements and simulations was 3.75 MHz, significantly higher than the center frequency of 2.3 MHz for the chest wall measurements and simulations. Differences in pulse frequency and specimen thickness may explain the discrepancy in energy level distortion between the abdominal wall and chest wall, but do not fully explain the discrepancy in arrival time distortion results. For instance, the mean arrival time and energy level fluctuations per unit length are 1.02 ns/mm and 0.083 dB/mm for the present study vs 1.96 ns/mm and 0.105 dB/mm for the abdominal wall cross sections of Ref. 10 and 12. Arrival time distortion was shown here to increase only subtly with increasing pulse frequency, so that this discrepancy in arrival time fluctuations is not fully explained by pulse frequency differences. However, energy level fluctuations increase markedly with frequency for chest wall tissue. Thus, for equal ultrasonic pulse frequencies, chest wall tissue should cause energy level distortion per unit length comparable to that caused by abdominal wall tissue.

It was suggested in Ref. 8 that chest wall morphology may differ from abdominal morphology in a manner that results in smaller ultrasonic wavefront distortion. The cross sections employed here can be compared with those employed in Refs. 10 and 12 to evaluate the importance of morphological differences between chest wall and abdominal wall tissue. One difference between the two groups of cross sections is the nature of the subcutaneous fat layers. The abdominal wall cross sections generally contain thicker fat layers, containing many more lobular structures than the chest wall cross sections. Since the high contrast between septa and fat causes substantial ultrasonic scattering,¹⁰⁻¹² this morphological difference is likely to result in lower overall energy level and waveform distortion for chest wall tissue (although, as discussed above, the energy level distortion per unit propagation length should be comparable). Also, the abdominal wall and chest wall cross sections have a markedly different structure within the muscle layers that occur below the subcutaneous fat. The abdominal wall cross sections have many large-scale features due to aponeuroses (interfaces between muscle groups, composed of connective tissue and fat) and large fatty regions. These large-scale features cause large wavefront fluctuations that are associated with large rms arrival time fluctuations.^{10,12} In contrast, muscle layers of the chest wall cross sections considered here contain primarily smaller-scale structures associated with small islands of interlaced fatty tissue. This morphological difference may result in lower large-scale arrival time fluctuations but significant energy level fluctuations associated with scattering, consistent with the differences between distortion caused by soft tissues in the abdominal wall and the chest wall.

The present results for the frequency dependence of distortion provide further insight into the importance of scatter-

ing effects relative to large-scale structure in wavefront distortion caused by soft tissues. If wavefront distortion in the chest wall were caused only by large-scale tissue structures, the distortion would be roughly independent of frequency, since propagation effects are independent of frequency in the geometric acoustics limit. However, distortion caused by scattering effects should increase with the pulse frequency for inhomogeneities of size comparable to the wavelength. Previous simulation and experimental studies¹⁰⁻¹² on distortion caused by the human abdominal wall have suggested that energy level fluctuations and waveform distortion are generally associated with scattering effects, while arrival time fluctuations are predominantly caused by large-scale path length differences. The present results, while consistent with those conclusions, indicate that scattering plays a role in all types of distortion considered here. Since energy level fluctuations and waveform similarity factors exhibit more dramatic increases in distortion with increasing pulse frequency, the present results suggest that scattering is of primary importance in causing energy level and waveform distortion and of secondary importance in causing arrival time distortion.

These results can be employed to evaluate the potential of various approaches to improve echocardiographic imaging. Available acoustic windows for transthoracic imaging are severely limited by the presence of the ribs, so that image quality cannot be significantly improved by an increase of aperture size. The present results also indicate that use of higher-frequency probes may provide less benefit than expected because of frequency-dependent scattering in the chest wall.

For these reasons, aberration correction methods are potentially important in transthoracic echocardiography, particularly for higher-frequency imaging. The frequency-dependent distortion results reported here suggest that distortion models employing single phase screens may be of some benefit for aberration correction in echocardiography through soft tissue paths. The relatively weak dependence of arrival time fluctuations on pulse frequency suggests that a large portion of arrival time variations are caused by tissue structures too large to cause significant frequency-dependent scattering effects. Similar conclusions regarding the importance of large-scale structures to arrival time fluctuations have also been drawn from results presented in Refs. 10 and 12.

Still, the present results, like those from earlier studies,¹⁰⁻¹² suggest that single phase screens will not provide complete correction for distortion caused by soft tissues. In particular, methods employing single phase screens will not completely remove distortion caused by scattering. The sharp increase of amplitude and waveform distortion with frequency, as well as the moderate increase of arrival time distortion with frequency, indicate that scattering effects become much more important to ultrasonic aberration as imaging frequencies increase. Furthermore, phase screen models do not inherently account for distortion caused by rib structures, shown here to produce diffraction, reflection, and scattering. Thus, any attempted correction using only phase

screen models is likely to provide little improvement in the presence of strong rib-induced effects.

Other correction models that incorporate rib structures may provide greater image improvements for the distortions most important to echocardiography. Processing wavefronts with techniques such as angular spectrum filtering can remove some spurious arrivals,⁵⁰ although such computations may be difficult to incorporate into a general correction algorithm. Other possible methods include those incorporating models of tissue structure. Models incorporating ray acoustics⁹ may provide improvement, but implicitly neglect diffraction and scattering effects, so that aberration correction would be incomplete, particularly for small intercostal spaces. A more complete aberration correction method could employ synthetic focusing using full-wave numerical computation of acoustic fields within sufficiently accurate models of tissue structure. This method has been implemented, within the context of a quantitative frequency-domain inverse scattering method, in Ref. 51. However, the results presented here indicate that distortion caused by soft tissue and rib structures varies widely based on morphological variations between (and within) individuals. Thus, for any general correction method employing models of tissue structure, separate models of tissue structure must be constructed for each region of interest.

V. CONCLUSIONS

A computational study of ultrasonic propagation through the chest wall, including tissue-dependent absorption as well as detailed anatomical cross sections, has been presented. For soft tissue paths, computational results for arrival time distortion, energy level distortion, and correlation lengths of these distortions are comparable to those reported in previous chest wall measurements. Both simulations and measurements indicate that arrival time distortion and energy level distortion caused by soft tissues in the human chest wall is smaller than that caused by the human abdominal wall. Differences in morphology between the abdominal wall and the chest wall provide a probable explanation for this difference.

Distortion caused by rib structures is much more severe than that caused by soft tissues. Reflections and diffraction from rib structures complicate wavefronts that travel through soft tissue paths adjacent to ribs and can cause arrival time and energy level fluctuations much greater than those induced by soft tissue structures. Wavefronts propagating directly through rib structures are attenuated by both internal absorption and reflection at interfaces between bone, cartilage, and soft tissue. Internal scattering within rib structures causes distortion phenomena that include severe waveform and energy level distortion, additional attenuation, and lowering of the effective frequency for the transmitted pulse. The strong dependence of distortion on the morphological details of rib structures presents a major challenge for aberration correction in echocardiography.

Simulation of propagation through soft tissue paths using three different pulse frequencies has indicated that the distortion types investigated here have different frequency dependence. Arrival time fluctuations increase subtly with frequency, while energy level and waveform distortion in-

crease greatly. Thus, a substantial portion of arrival time fluctuations produced by the chest wall may be explained by large-scale tissue variations, but some arrival time distortion and most energy level and waveform distortion apparently result from scattering. Thus, correction of wavefront distortion caused by soft tissues should become both more important and more challenging as pulse frequencies employed in imaging systems are increased.

ACKNOWLEDGMENTS

The authors thank Cari Kelly, Tara Jones, Tracy David, and Michael Pirri for assistance in creating accurate tissue maps of the chest wall cross sections. The original two-dimensional version of the distortion estimation software using the reference waveform method was developed by D.-L. Donald Liu. Funding for this investigation was provided by NIH grants No. DK 45533, No. HL 50855, and No. CA 74050, U.S. Army Grant No. DAMD17-98-1-8141, DARPA Grant N00014-96-0749, and the University of Rochester Diagnostic Ultrasound Research Laboratory Industrial Associates. Some computations were performed at the Cornell National Supercomputing Facility, which was supported in part by the National Science Foundation, New York State, and the IBM Corporation.

¹H. Feigenbaum, *Echocardiography* (Lea and Febiger, Philadelphia, 1994).
²A. E. Weyman, *Principles and Practice of Echocardiography*, 5th ed. (Lea and Febiger, Philadelphia, 1994).
³J. T. T. Chen, *Essentials of Cardiac Imaging*, 2nd ed. (Lippincott-Raven, Philadelphia, 1997).
⁴F. M. Clements and N. P. de Bruijn, *Transesophageal Echocardiography* (Little, Brown, and Company, Boston, 1991).
⁵K.-L. Chan, G. I. Cohen, R. A. Sochowski, and M. G. Baird, "Complications of transesophageal echocardiography in ambulatory adult patients: analysis of 1500 consecutive examinations," *J. Am. Soc. Echocardiogr.* **4**, 577–582 (1991).
⁶W. G. Daniel, R. Erbel, W. Kasper, C. A. Visser, R. Engberding, G. R. Sutherland, E. Grube, P. Hanrath, B. Maisch, K. Dennig, M. Scharlt, P. Kremer, C. Angermann, S. Iliceto, J. M. Curtius, and A. Mügge, "Safety of transesophageal echocardiography: a multicenter survey of 10,419 examination," *Circulation* **83**, 817–821 (1991).
⁷A. D. Savakus, K. K. Shung, and N. B. Miller, "Distortions of ultrasonic field introduced by the rib cage in echocardiography," *J. Clin. Ultrasound* **10**, 413–419 (1982).
⁸L. M. Hinkelman, T. L. Szabo, and R. C. Waag, "Measurements of ultrasonic pulse distortion produced by the human chest wall," *J. Acoust. Soc. Am.* **101**, 2365–2373 (1997).
⁹Y. Y. Botros, E. S. Ebbini, and J. L. Volakis, "Two-step hybrid virtual array-ray (VAR) technique for focusing through the rib cage," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**, 989–999 (1998).
¹⁰T. D. Mast, L. M. Hinkelman, M. J. Orr, V. W. Sparrow, and R. C. Waag, "Simulation of ultrasonic pulse propagation through the abdominal wall," *J. Acoust. Soc. Am.* **102**, 1177–1190 (1997).
¹¹L. M. Hinkelman, T. D. Mast, L. A. Metley, and R. C. Waag, "The effect of abdominal wall morphology on ultrasonic pulse distortion. Part I: Measurements," *J. Acoust. Soc. Am.* **104**, 3635–3649 (1998).
¹²T. D. Mast, L. M. Hinkelman, M. J. Orr, and R. C. Waag, "The effect of abdominal wall morphology on ultrasonic pulse distortion. Part II: Simulations," *J. Acoust. Soc. Am.* **104**, 3651–3664 (1998).
¹³H. Kolsky, *Stress Waves in Solids* (Clarendon, Oxford, UK, 1953), pp. 106–129.
¹⁴S. Leeman, L. Hutchins, and J. P. Jones, "Bounded pulse propagation," in *Acoustical Imaging*, edited by P. Alais and A. F. Metherell (Plenum, New York, 1982), Vol. 10, pp. 427–435.
¹⁵S. Finette, "Computational methods for simulating ultrasound scattering in soft tissue," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **34**, 283–292 (1987).

¹⁶C. W. Manry and S. L. Broschat, "FDTD simulations for ultrasound propagation in a 2-D breast model," *Ultrason. Imaging* **118**, 25–34 (1996).
¹⁷V. W. Sparrow and R. Raspet, "A numerical method for general finite amplitude wave propagation and its application to spark pulses," *J. Acoust. Soc. Am.* **90**, 2683–2691 (1991).
¹⁸R. W. MacCormack, *Lecture Notes in Physics* (Springer-Verlag, Berlin, 1971), Vol. 8, p. 151.
¹⁹D. Gottlieb and A. Turkel, "Dissipative two-four methods for time-dependent problems," *Math. Comput.* **30**, 703–723 (1976).
²⁰E. Turkel, "On the practical use of high-order methods for hyperbolic systems," *J. Comput. Phys.* **35**, 319–340 (1980).
²¹L. M. Hinkelman, L. A. Metlay, C. J. Churukian, and R. C. Waag, "Modified Gomori trichrome stain for macroscopic tissue slices," *J. Histotech.* **19**, 321–323 (1996).
²²H. Gray, *Gray's Anatomy*, edited by T. P. Pick and R. Howden (Gramercy, New York, 1977), pp. 358–364 (facsimile of 1901 American edition of *Anatomy, Descriptive and Surgical*).
²³R. Warwick and P. L. Williams, *Gray's Anatomy*, 35th ed. (Saunders, Philadelphia, 1973), pp. 488–490, 519–527.
²⁴A. Elkeles, "Sex differences in the calcification of the costal cartilages," *J. Am. Geriatr. Soc.* **14**, 456–461 (1966).
²⁵N. 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).
²⁶*Handbook of Chemistry and Physics*, edited by R. C. Weast (CRC Press, Boca Raton, 1985), p. F-10.
²⁷S. 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).
²⁸S. A. Goss, R. L. Johnston, and F. Dunn, "Compilation of empirical ultrasonic properties of mammalian tissues II," *J. Acoust. Soc. Am.* **68**, 93–108 (1980).
²⁹J. E. Olerud, W. O'Brien, M. A. Riederer-Henderson, D. Steiger, F. K. Forster, C. Daly, D. J. Ketterer, and G. F. Odland, "Ultrasonic assessment of skin and wounds with the scanning laser acoustic microscope," *J. Invest. Dermatol.* **8**, 615–623 (1987).
³⁰R. E. Neuman and M. A. Logan, "The determination of collagen and elastin in tissues," *J. Biol. Chem.* **186**, 549–556 (1950).
³¹F. A. Duck, *Physical Properties of Tissue: A Comprehensive Reference Book* (Academic, New York, 1990).
³²K. T. Dussik and D. J. Fritch, "Determination of sound attenuation and sound velocity in the structure constituting the joints, and of the ultrasonic field distribution within the joints on living tissues and anatomical preparations, both in normal and pathological conditions," Progress Report to Public Health Service, National Institutes of Health Project A454, 15 September 1956.
³³H. Q. Woodard and D. R. White, "The composition of body tissues" *Br. J. Radiol.* **59**, 1209–1219 (1986).
³⁴International Commission on Radiation Units and Measurements, *ICRU Report 61: Tissue Substitutes, Phantoms, and Computational Modelling in Medical Ultrasound* (ICRU, Bethesda, MD, 1998), pp. 43–51.
³⁵K. H. Herzfeld and T. A. Litovitz, *Absorption and Dispersion of Ultrasonic Waves* (Academic, New York, 1959), pp. 353–361.
³⁶Simulations were performed on a Linux workstation with an AMD K6 processor running at 200 MHz and 128 MB of random-access memory. The simulation code was written in Fortran 77 and compiled using the front end fort77 and the Fortran-to-C converter f2c with the Gnu C compiler gcc.
³⁷D.-L. Liu and R. C. Waag, "Correction of ultrasonic wavefront distortion using backpropagation and a reference waveform method for time-shift compensation," *J. Acoust. Soc. Am.* **96**, 649–660 (1994).
³⁸L. M. Hinkelman, D.-L. Liu, L. A. Metlay, and R. C. Waag, "Measurements of ultrasonic pulse arrival time and energy level variations produced by propagation through abdominal wall," *J. Acoust. Soc. Am.* **95**, 530–541 (1994).
³⁹J. Wolff, *Das Gesetz der Transformation der Knochen* (Hirschwald, Berlin, 1982).
⁴⁰J. C. Rice, S. C. Cowin, and J. A. Bowman, "On the dependence of the elasticity and strength of cancellous bone on apparent density," *J. Biomech.* **21**, 155–168 (1988).
⁴¹J. Y. Rho, R. B. Ashman, and C. H. Turner, "Young's modulus of trabecular and cortical bone material: ultrasonic and tensile measurements," *J. Biomech.* **26**, 111–119 (1993).
⁴²C. H. Turner, J. Rho, Y. Takano, T. Y. Tsui, and G. M. Pharr, "The

- elastic properties of trabecular and cortical bone tissues are similar: results from two microscopic measurement techniques," *J. Biomech.* **32**, 437–441 (1999).
- ⁴³S. Lees, J. M. Ahern, and M. Leonard, "Parameters influencing the sonic velocity in compact calcified tissues of various species," *J. Acoust. Soc. Am.* **74**, 28–33 (1983).
- ⁴⁴S. Lees, "Sonic properties of mineralized tissues," in *Tissue Characterization with Ultrasound*, edited by J. F. Greenleaf (CRC Press, Boca Raton, 1986), pp. 207–226.
- ⁴⁵L. A. Frizzell, E. L. Carstensen, and J. F. Dyro, "Shear properties of mammalian tissues at low MHz frequencies," *J. Acoust. Soc. Am.* **60**, 1409–1411 (1976).
- ⁴⁶E. L. Madsen, H. J. Sathoff, and J. A. Zagzebski, "Ultrasonic shear wave properties of soft tissues and tissue-like materials," *J. Acoust. Soc. Am.* **74**, 1346–1355 (1983).
- ⁴⁷L. Adler and K. V. Cook, "Ultrasonic properties of freshly frozen dog tibia," *J. Acoust. Soc. Am.* **58**, 1107–1108 (1975).
- ⁴⁸S. S. Kohles, J. R. Bowers, A. C. Vailas, and R. Vanderby, "Ultrasonic wave velocity measurement in small polymeric and cortical bone specimens," *J. Biomech. Eng.* **119**, 232–236 (1997).
- ⁴⁹K. Chandra and C. Thompson, "Ultrasonic characterization of fractal media," *Proc. IEEE* **81**, 1523–1533 (1993).
- ⁵⁰L. M. Hinkelman and D.-L. Liu, "Measurement and analysis of ultrasonic pulse wavefront distortion produced by chest wall," *Proceedings of the 12th Annual University of Rochester Diagnostic Ultrasound Research Laboratory Industrial Associates Meeting*, 1995, pp. 8–25.
- ⁵¹T. D. Mast, A. I. Nachman, and R. C. Waag, "Focusing and imaging using eigenfunctions of the scattering operator," *J. Acoust. Soc. Am.* **102**, 715–725 (1997).

Excitation and propagation of surface waves on a viscoelastic half-space with application to medical diagnosis

T. J. Royston^{a)}

University of Illinois at Chicago, Chicago, Illinois 60607

H. A. Mansy and R. H. Sandler

Rush Medical College, Chicago, Illinois 60612

(Received 15 June 1999; accepted for publication 10 September 1999)

An analytical solution is developed for the problem of surface wave generation on a linear viscoelastic half-space by a finite rigid circular disk located on the surface and oscillating normal to it. The solution is an incremental advancement of theoretical work reported in articles focused on seismology. Since the application of interest here is medical diagnostics, the solution is verified experimentally using a viscoelastic phantom with material properties comparable to biological soft tissue. Findings suggest that prior estimates in the literature of the shear viscosity in human soft tissue may not be accurate in the low audible frequency range. Measurement of wave motion on the skin surface caused by internal biological functions or external stimuli has been studied by a few researchers for rapid, noninvasive diagnosis of a variety of specific medical ailments. It is hoped that the developments reported here will advance these techniques and also provide insight into related diagnostic methods, such as sonoelastic imaging and other methodologies that utilize disease-related variations in tissue shear elasticity or variations in density due to gaseous inclusions.

© 1999 Acoustical Society of America. [S0001-4966(99)07712-7]

PACS numbers: 43.80.Cs, 43.80.Ev, 43.80.Qf, 43.20.Bi [FD]

INTRODUCTION

There has been limited research in the area of medical diagnosis using surface wave behavior on human soft tissue. Such an approach potentially can be noninvasive, rapid, and inexpensive. For example, Lee¹ and Hong and Fox^{2,3} have investigated noncontact characterization of cardiovascular dynamics using optical interferometry. Diagnosis of certain types of edema conditions and other skin diseases, including skin cancer, via alterations in surface wave propagation has been studied by several groups.⁴⁻¹¹ Other research has focused on characterizing lung edema by studying surface wave propagation directly on the inflated pulmonary parenchyma¹² or on the chest wall.¹³

Exploiting changes in subsurface shear wave propagation as a means of diagnosis has also received much attention in recent years. Several groups of researchers have utilized low-frequency excitation of shear waves, e.g., 20 to a few hundred Hz, coupled with Doppler ultrasonic imaging.¹⁴⁻¹⁶ It is sometimes referred to as "sonoelastic imaging." This method has been proposed as a technique for locating tumors, which typically represent significant changes in stiffness properties in otherwise acoustically homogeneous regions. A localized stiffness, such as a tumor, will distort the shear wave vibration pattern as imaged by Doppler ultrasound.

To support further development of these promising diagnostic methodologies, a better understanding of surface wave behavior on soft biological tissue is needed in the low audible frequency regime, 20 to 100 Hz. Above these fre-

quencies, surface waves become negligible more than a few centimeters from the source due to the high viscosity of soft biological tissue. Below these frequencies (the lower-frequency limit dependent on location), the human body is more aptly described by a lumped parameter model¹⁷ (or, at least, is not reasonably approximated as a half-space). It is noted that surface wave behavior is relevant to subsurface measurement schemes like sonoelastic imaging as the effects of surface (Rayleigh) waves penetrate below the surface with significant amplitude to frequency-dependent depths comparable to their wavelength on the surface. Surface waves, like shear waves, are predominantly dependent on shear viscoelastic properties and density.

A critical aspect of many of these proposed diagnostic techniques is an understanding of how the skin surface responds to external stimulation. While the general nature of free Rayleigh (surface) wave propagation on a viscoelastic medium has been investigated, the problem of surface wave generation from a typical finite surface source has not been explicitly addressed in the literature, to the best of the authors' knowledge. A basic pedagogical problem is that of determining the surface wave response to a disk-shaped transducer located on the surface of a viscoelastic half-space with human soft-tissue properties and vibrating normal to it. A solution to the related elastic problem can be found in the seismological literature. Seminal papers by Miller and Pursey^{18,19} address this problem and, despite their date, are frequently cited in many contemporary works. Integral representations of the general solution for the stated disk problem were offered. Asymptotic expressions for the far field were given in explicit form.

It is the objective of the research reported in this article to first theoretically derive a more accurate surface wave

^{a)} Author to whom correspondence should be addressed. Electronic mail: troyston@uic.edu

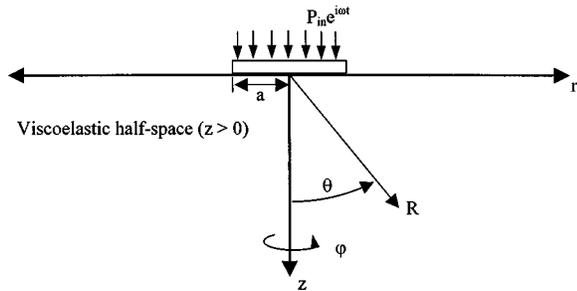


FIG. 1. Ideal viscoelastic half-space problem.

response function, valid for the near and far field, of an isotropic viscoelastic half-space to a finite disk of radius “ a ” on the surface and harmonically vibrating normal to it. The derivation is then validated based on experimental studies of a viscoelastic phantom with material properties comparable to human soft tissue. Generated theoretical results are also used to assess the accuracy of material parameter values previously reported in the literature for human soft tissue.

I. THEORY

The problem considered here is that of a circular disk of finite radius ‘ a ’ vibrating normally to the surface of an isotropic, viscoelastic medium. The geometry of the problem is depicted in Fig. 1. The theoretical derivations below represent incremental advancements to the derivations reported in Miller and Pursey.^{18,19} For the medium, the equation of wave propagation may be expressed in the following form:¹⁸

$$(\lambda + 2\mu)\nabla\nabla\cdot\mathbf{u} - \mu\nabla\times\nabla\times\mathbf{u} = \rho\frac{\partial^2\mathbf{u}}{\partial t^2}. \quad (1)$$

Here, \mathbf{u} is the displacement vector, ρ is the density of the medium, $\partial/\partial t$ denotes a derivative with respect to time, and λ and μ are the Lamé constants of the medium. For a linear viscoelastic material, the rate-dependent Lamé “constants” are expressible as $\lambda(t) = \lambda_1 + \lambda_2\partial/\partial t$ and $\mu(t) = \mu_1 + \mu_2\partial/\partial t$ where λ_1 , λ_2 , μ_1 , and μ_2 are coefficients of volume compressibility, volume viscosity, shear elasticity, and shear viscosity, respectively.²⁰ The problem is solved in the polar coordinate system depicted in Fig. 1. Thus, we have $\mathbf{u} = [u_r, u_\phi, u_z]^T$ and it is noted that $u_\phi = 0$ and derivatives with respect to ϕ vanish due to axisymmetry. Hence,

$$\nabla\cdot\mathbf{u} = \frac{1}{r}\frac{\partial}{\partial r}(ru_r) + \frac{\partial u_z}{\partial z} \quad (2a)$$

and

$$\nabla\times\mathbf{u} = \left(\frac{\partial u_r}{\partial z} - \frac{\partial u_z}{\partial r}\right)\mathbf{j}. \quad (2b)$$

Consider the case of harmonic force excitation of the massless disk of amplitude per unit area P_{in} and circular frequency ω such that it applies a uniform stress on the surface of the medium in the circular region of $r < a$. (It is noted that this assumption of uniform stress may limit the accuracy of the model to values of a less than or comparable to surface wavelength.) Miller and Pursey¹⁸ derive the following integral expression for wave propagation in the r and z directions

at any location in and on the surface of the medium, i.e., $z \geq 0$. Note that time dependence $e^{i\omega t}$ is omitted where $i = \sqrt{-1}$.

$$\frac{u_z}{P_{in}} = \frac{a}{\mu} \int_0^\infty \frac{J_1(\zeta ak_1)\sqrt{\zeta^2-1}}{F_0(\zeta)} \{2\zeta^2 e^{-zk_1\sqrt{\zeta^2-\eta^2}} + (\eta^2 - 2\zeta^2) e^{-zk_1\sqrt{\zeta^2-1}}\} J_0(\zeta rk_1) d\zeta, \quad (3a)$$

$$\frac{u_r}{P_{in}} = \frac{a}{\mu} \int_0^\infty \frac{J_1(\zeta ak_1)\zeta}{F_0(\zeta)} \{2\sqrt{\zeta^2-1}\sqrt{\zeta^2-\eta^2} e^{-zk_1\sqrt{\zeta^2-\eta^2}} + (\eta^2 - 2\zeta^2) e^{-zk_1\sqrt{\zeta^2-1}}\} J_1(\zeta rk_1) d\zeta, \quad (3b)$$

where

$$F_0(\zeta) = (2\zeta^2 - \eta^2)^2 - 4\zeta^2\sqrt{\zeta^2-\eta^2}\sqrt{\zeta^2-1}, \quad (3c)$$

$$\eta = k_2/k_1, \quad (3d)$$

$$k_1 = \omega\sqrt{\rho/(\lambda + 2\mu)}, \quad (3e)$$

and

$$k_2 = \omega\sqrt{\frac{\rho}{\mu}}. \quad (3f)$$

In the above equations, k_1 and k_2 denote the wavenumbers for compression and shear wave propagation, respectively, and J_0 and J_1 refer to Bessel functions of the first kind. The dummy variable ζ is used to denote integration over the wavenumber domain that has been normalized with respect to k_1 .

Equation (3a) can be rewritten as follows:

$$\frac{u_z}{P_{in}} = \frac{a}{\mu} \int_0^\infty \{\chi_1(\zeta) e^{-zk_1\sqrt{\zeta^2-\eta^2}} + \chi_2(\zeta) e^{-zk_1\sqrt{\zeta^2-1}}\} J_0(\zeta rk_1) d\zeta, \quad (4a)$$

where

$$\chi_1(\zeta) = 2\zeta^2 J_1(\zeta ak_1)\sqrt{\zeta^2-1}/F_0(\zeta) \quad (4b)$$

and

$$\chi_2(\zeta) = J_1(\zeta ak_1)\sqrt{\zeta^2-1}(\eta^2 - 2\zeta^2)/F_0(\zeta). \quad (4c)$$

To solve this equation, we make use of the integral expression for the Bessel function that is valid when $|\arg(r\zeta k_1)| < \pi/2$, a reasonable assumption in this case.¹⁸ It is given by

$$J_0(\zeta rk_1) = \frac{2}{\pi} \int_0^\infty e^{-\xi^2} \{u(\zeta rk_1, \xi) + \bar{u}(\zeta rk_1, \xi)\} d\xi, \quad (5a)$$

where

$$u(\zeta rk_1, \xi) = e^{i(rk_1\xi - (1/4)\pi)} (2\zeta rk_1 + i\xi^2)^{-1/2}, \quad (5b)$$

$$\begin{aligned} \bar{u}(\zeta rk_1, \xi) &= e^{-i(rk_1\xi - (1/4)\pi)} (2\zeta rk_1 - i\xi^2)^{-1/2} \\ &= -u(-\zeta rk_1, \xi), \end{aligned} \quad (5c)$$

and ξ is a dummy variable of the integration. Substituting these relations into Eq. (4a) and using the fact that $\chi_1(\zeta)$ and $\chi_2(\zeta)$ are odd functions, we obtain

$$\frac{u_z}{P_{in}} = \frac{2a}{\pi\mu} \int_{-\infty}^{\infty} \left\{ \chi_1(\zeta) e^{-z k_1 \sqrt{\zeta^2 - \eta^2}} + \chi_2(\zeta) e^{-z k_1 \sqrt{\zeta^2 - 1}} \right\} \int_0^{\infty} e^{-\xi^2} u(\zeta r k_1, \xi) d\xi d\zeta. \quad (6)$$

Using MATHEMATICA, a solution to the integral over ξ can be found

$$\begin{aligned} & \int_0^{\infty} e^{-\xi^2} u(\zeta r k_1, \xi) d\xi \\ &= \int_0^{\infty} e^{-\xi^2} e^{i(\zeta r k_1 - \pi/4)} (2\zeta r k_1 + i\xi^2)^{-1/2} d\xi \\ &= -\frac{1}{2} (-1)^{3/4} e^{i(\zeta r k_1 - \pi/4)} e^{-i\zeta r k_1} K_0(-i\zeta r k_1). \end{aligned} \quad (7)$$

Here, K_0 denotes the modified Bessel function of the second kind of order 0. Thus, Eq. (3a) has been transformed into the following equivalent expression:

$$\begin{aligned} \frac{u_z}{P_{in}} &= -(-1)^{3/4} e^{-i\pi/4} \frac{a}{\pi\mu} \int_{-\infty}^{\infty} \frac{J_1(\zeta a k_1) \sqrt{\zeta^2 - 1}}{F_0(\zeta)} \\ &\quad \times \{ 2\zeta^2 e^{-z k_1 \sqrt{\zeta^2 - \eta^2}} + (\eta^2 - 2\zeta^2) e^{-z k_1 \sqrt{\zeta^2 - 1}} \} \\ &\quad \times K_0(-i\zeta r k_1) d\zeta. \end{aligned} \quad (8)$$

The primary advantage of this expression over that of Eq. (3) is that the Cauchy principal value theorem is now easily applied to obtain its solution. It is assumed that the principal values of the radicals are used in the above expressions throughout the range of integration. This is consistent with waves traveling away from the source when a time factor $e^{i\omega t}$ is used and is necessary for the convergence of the integrals at infinity. To ensure continuity of the integrands, the contour of integration must be indented to pass below the branch point at -1 and above the branch point at $+1$. Under harmonic excitation conditions, we have

$$\eta = \sqrt{\{\lambda_1 + 2\mu_1 + i\omega(\lambda_2 + 2\mu_2)\} / (\mu_1 + i\omega\mu_2)}. \quad (9)$$

For realistic values of the coefficients in this expression, $+\eta$ will be located below the real axis in the 4th quadrant and $-\eta$ will be located in the 2nd quadrant above the real axis.

The poles of the integral in Eq. (8) are the zeros of the expression $F_0(\zeta)$ which is a cubic equation in ζ^2 . Hence, there are six roots. It is found, however, that for realistic values of η the only zeros of $F_0(\zeta)$ which correspond to principal values of the radicals $\sqrt{\zeta^2 - 1}$ and $\sqrt{\zeta^2 - \eta^2}$ are those for which ζ^2 is greater in magnitude than η^2 . These poles p will be located in like quadrants as η and will have magnitude slightly greater than η . Physically it follows that, in this normalized wavenumber domain, poles p associated with propagating surface or Rayleigh waves will have magnitude greater than η , which is associated with the propagating shear wave in the medium. This leads to a surface wave of phase speed slightly less than the phase speed of the shear wave. For the viscoelastic case with nonzero damping, contour integration along the real axis with the noted indentation about ± 1 is needed. In the limit as damping is neglected, the contour integration still follows the same route below nega-

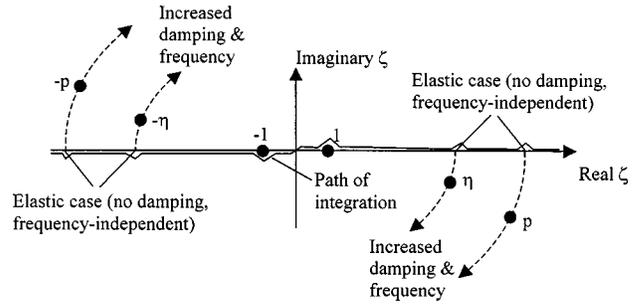


FIG. 2. Location of poles and determination of contour.

tive poles and above positive poles, but it must be further indented as indicated in Fig. 2. See Miller and Pursey¹⁸ for further discussion of this issue.

For the case that $\theta = \pi/2$, i.e., along the surface, the integral for u_z/P_{in} given in Eq. (8) is negligible except for the residue associated with the pole p . The contour integration can be taken counterclockwise around the 1st and 2nd quadrants, thus encircling the pole at $-p$. For the linear elastic case, the values of p are independent of harmonic excitation frequency ω and are located on the real axis, requiring indentation of the integral as noted above. For the linear viscoelastic case, the poles are no longer on the real axis, but are a function of harmonic excitation frequency, with the imaginary part increasing with increased excitation frequency.

In Ref. 19, an asymptotic expression was used to approximate the integral of Eq. (7). This led to results that are asymptotically valid for “large r ” and “small a .” In this paper, the exact solution of the integral is used, which has been obtained via MATHEMATICA. Thus, more general expressions are given for surface waves that should be valid for small r and large a , as well. We have the following:

$$\begin{aligned} \frac{u_z}{P_{in}} &= -(-1)^{3/4} e^{-i\pi/4} \frac{a}{\mu} 2i \left[\text{residue of } \frac{J_1(\zeta a k_1) \sqrt{\zeta^2 - 1}}{F_0(\zeta)} \right. \\ &\quad \times \{ 2\zeta^2 e^{-z k_1 \sqrt{\zeta^2 - \eta^2}} + (\eta^2 - 2\zeta^2) e^{-z k_1 \sqrt{\zeta^2 - 1}} \} \\ &\quad \left. \times K_0(-i\zeta r k_1) \text{ at } \zeta = -p \right] \\ &= (-1)^{3/4} e^{-i\pi/4} \frac{2ia}{\mu} \frac{J_1(p a k_1) \sqrt{p^2 - 1}}{F'_0(-p)} \\ &\quad \times \{ 2p^2 e^{-z k_1 \sqrt{p^2 - \eta^2}} + (\eta^2 - 2p^2) e^{-z k_1 \sqrt{p^2 - 1}} \} \\ &\quad \times K_0(ip r k_1), \end{aligned} \quad (10a)$$

where

$$F'_0(-p) = \left. \frac{\partial F_0}{\partial \zeta} \right|_{\zeta = -p}. \quad (10b)$$

This expression agrees asymptotically with Ref. 19. The asymptotic solution of the reference is given here for comparison:

$$\frac{u_z}{P_{in}} \approx \frac{a^2 e^{-(1/4)i\pi}}{\mu F'_0(p)} \sqrt{\frac{\pi k_1 p (p^2 - 1)}{2r}} \times e^{-ik_1 p r} \{2p^2 e^{-k_1 z \sqrt{p^2 - \eta^2}} + (\eta^2 - 2p^2) e^{-k_1 z \sqrt{p^2 - 1}}\}. \quad (11)$$

Note the approximations in Eq. (11) are with respect to the Bessel functions in Eq. (10). The function $J_1(pak_1)$ is approximated with $pak_1/2$. This may be acceptable in many seismological applications dealing with surface wave propagation on the earth caused by a highly localized excitation. But, in the medical application, it is conceivable that $|pak_1| > 1$ or ≈ 1 . In other words, the radius of the transducer disk a may be of the same order of magnitude as the length of the surface waves, given by $L_{surf} = 2\pi/\text{real}(k_1 p)$. (This fact does also call into question the assumption made at the beginning of the theoretical development that the disk will impart a uniform stress in the region $r < a$. Hence, the improved solution here may still have its limitations.) The approximation to $K_0(iprk_1)$ includes the exponential e^{iprk_1} and a few other terms and is only valid for large r . For the parameter values used in this article in Sec. III with $|pak_1| > 1$ and $r > a$, this approximation is reasonable.

To obtain a solution to Eq. (3b), it is observed that $\zeta k_1 J_1(\zeta r k_1) = -\partial/\partial r [J_0(\zeta r k_1)]$. It is also noted that

$$\frac{d}{dr} [K_0(-i\zeta r k_1)] = \frac{1}{2} i \zeta k_1 \{K_{-1}(-i\zeta r k_1) + K_1(-i\zeta r k_1)\} = i \zeta k_1 K_1(-i\zeta r k_1). \quad (12)$$

Here, K_{-1} and K_1 denote the modified Bessel functions of the second kind of order, -1 and 1 , respectively. Thus, following the arguments given above, the following equivalent expression to Eq. (3b) is obtained (assuming $|\arg(\zeta k_1)| < \pi/2$):

$$\frac{u_r}{P_{in}} = (-1)^{3/4} e^{-i\pi/4} \frac{ia}{\pi \mu} \int_{-\infty}^{\infty} \frac{J_1(\zeta a k_1) \zeta}{F_0(\zeta)} \times \{2\sqrt{\zeta^2 - 1} \sqrt{\zeta^2 - \eta^2} e^{-\zeta k_1 \sqrt{\zeta^2 - \eta^2}} + (\eta^2 - 2\zeta^2) e^{-\zeta k_1 \sqrt{\zeta^2 - 1}}\} K_1(-i\zeta r k_1) d\zeta. \quad (13)$$

Evaluating the contour integral (finding the residue) in a similar fashion to the case of u_z yields the following for the case of surface waves:

$$\frac{u_r}{P_{in}} = -(-1)^{3/4} e^{-i\pi/4} \frac{2a}{\mu} \frac{J_1(pak_1) p}{F'_0(-p)} \times \{2\sqrt{p^2 - 1} \sqrt{p^2 - \eta^2} e^{-\zeta k_1 \sqrt{p^2 - \eta^2}} + (\eta^2 - 2p^2) e^{-\zeta k_1 \sqrt{p^2 - 1}}\} K_1(iprk_1). \quad (14)$$

This expression agrees asymptotically with the far-field solution provided in Ref. 19. The asymptotic solution of the reference is given here for comparison:

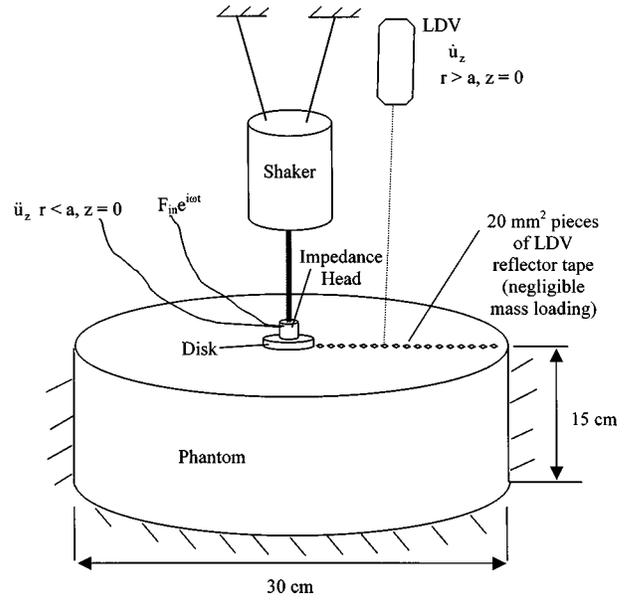


FIG. 3. Schematic of experimental setup.

$$\frac{u_r}{P_{in}} \approx \frac{a^2 e^{(1/4)i\pi}}{\mu F'_0(p)} \sqrt{\frac{\pi k_1 p^3}{2r}} e^{-ik_1 p r} \times \{2\sqrt{p^2 - 1} \sqrt{p^2 - \eta^2} e^{-k_1 z \sqrt{p^2 - \eta^2}} + (\eta^2 - 2p^2) e^{-k_1 z \sqrt{p^2 - 1}}\}. \quad (15)$$

II. EXPERIMENT

To assess the validity of the theoretical derivations presented in Sec. I an experiment was constructed that is schematically depicted in Fig. 3. The viscoelastic medium consists of a gel mixture of the following composition (per liter of water): 70 grams gelatin, 40 grams *n*-propanol, and 4 grams formaldehyde (37% solution). This ‘‘recipe’’ for a soft-tissue phantom is based on prior investigations.²¹ By taking velocity measurements using the laser Doppler vibrometer (Polytec model CLV-800-FF 1000), vertical velocity \dot{u}_z can be measured for any $r > a$ at the surface, $z = 0$. An impedance head (PCB model 288B02) records acceleration and force input to the Plexiglas disk at $z = 0$ and $0 < r < a$. The vibratory excitation is delivered through a stinger by an electromagnetic shaker (Labworks model ET-132-2) that is flexibly suspended above the phantom. Sensor output signals are recorded with a Hewlett-Packard 35 670 Dynamic Signal (FFT) Analyzer, which also performs preliminary analysis. The analyzer also provides the chirp excitation signal to the amplifier that drives the shaker. Chirp (rapidly swept sinusoidal) excitation signals were used to enable extracting frequency–amplitude–phase–time information. This allowed detection of outgoing waves from the source independent from reflections due to the finite boundaries of the phantom. Beyond basic frequency-response functions available in real time using the 35 670, more detailed analyses of the experimental data were conducted using MATLAB software.

TABLE I. Soft tissue and gel phantom parameter values for the theoretical and experimental studies.

Parameter	Soft tissue ^a	Gel phantom
λ_1 (N/m ²)	2.6×10^{9b}	2.6×10^{9b}
λ_2 (Ns/m ²)	0 ^c	0 ^c
μ_1 (N/m ²)	2.5×10^3	4.5×10^{3d}
μ_2 (Ns/m ²)	15	4 ^e
ρ (kg/m ³)	1100	1000
η, p at $\omega/2\pi=40$ Hz	668–359 <i>i</i> , 699–375 <i>i</i>	746–82.3 <i>i</i> , 781–86.2 <i>i</i>
c_{surf} (m/s), L_{surf} (cm)	2.20, 5.51	2.06, 5.16
η, p at $\omega/2\pi=60$ Hz	543–354 <i>i</i> , 569–370 <i>i</i>	731–119 <i>i</i> , 765–125 <i>i</i>
c_{surf} (m/s), L_{surf} (cm)	2.70, 4.50	2.11, 3.52
η, p at $\omega/2\pi=80$ Hz	464–335 <i>i</i> , 486–351 <i>i</i>	710–151 <i>i</i> , 744–159 <i>i</i>
c_{surf} (m/s), L_{surf} (cm)	3.17, 3.96	2.17, 2.71
η, p at $\omega/2\pi=200$ Hz	278–244 <i>i</i> , 291–255 <i>i</i>	567–253 <i>i</i> , 593–265 <i>i</i>
c_{surf} (m/s), L_{surf} (cm)	5.28, 2.64	2.72, 1.36
η, p when $\mu_2=0$ (no damping)	1020, 1067	760, 796
c_{surf} (m/s), L_{surf} (cm)	1.44, $2\pi c_{\text{surf}}/\omega$	2.03, $2\pi c_{\text{surf}}/\omega$

^aLame constants and density based on von Gierke *et al.* (Ref. 22).

^bSame value as water (Ref. 20). Primarily affects compression waves, not shear waves.

^cNegligible below 20 kHz (Ref. 20).

^dBased on matching surface wave speed with that observed experimentally.

^eBased on matching attenuation rate with that observed experimentally.

III. RESULTS AND DISCUSSION

Material parameter values used to theoretically model the gel phantom are provided in Table I. Volume compressibility and density are comparable to water. Shear viscoelastic constants are based on comparison to experimental surface wave speed and attenuation measurements. Also given are typical parameter values estimated by von Gierke *et al.*²² to model human soft tissue. All theoretical results presented in this section are based on the material parameter values given in Table I gel or tissue, unless otherwise noted.

A. Gel phantom model—theory and experiment

The experimentally measured vertical particle velocity amplitudes on the surface normalized to the disk velocity amplitude $|\dot{u}_{z=0}(r, \omega)/\dot{u}_{z=0}(r=0, \omega)|$ as a function of distance from the source r and excitation frequency ω are shown in Fig. 4. The displayed results are based on calculating the transfer function of the complete time record, during which the excitation frequency was swept from 0 to 100 Hz. Due to the finite dimensions of the phantom, resonant behavior is detectable at the lower frequencies that cannot be predicted using the theory developed in the previous section. This behavior could be due both in part to reflections from the outer radial boundary of the phantom or from its finite depth. Above about 40 Hz, due to increasing attenuation and decreased wavelength, resonant behavior is less apparent.

In Figs. 5 and 6 experimental waveform measurements of \dot{u}_z/F_{in} at selected frequencies are directly compared to theoretical predictions of Sec. I based on the developed “improved theory” and the “asymptotic theory” of Miller and Pursey.^{18,19} Also shown are the predictions when damping is neglected in the “improved theory.” Here, $F_{\text{in}}=P_{\text{in}}\pi a^2$ is the force input of the disk of radius a . The experimental results validate the improved theory. The limitations of the asymptotic theory are evident, particularly as frequency in-

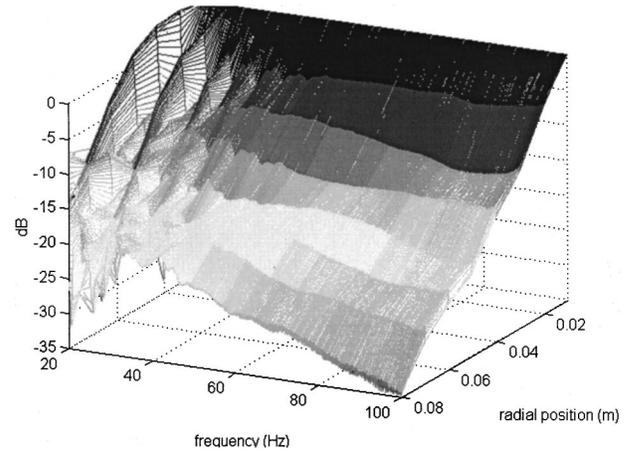


FIG. 4. Experimental measurements on the phantom of the vertical particle velocity amplitude on the surface normalized to the disk velocity as a function of excitation frequency and radial position. Amplitude in dB = $20 \log_{10}$.

creases and the approximation that $J_1(pak_1) \approx pak_1/2$ becomes less valid. The limitation of the approximation for $K_0(iprk_1)$ is not as evident for this set of parameter values. Also illustrated in these figures is the influence of the viscosity term, as the rate of attenuation per radial distance is significantly increased when viscosity is taken into account. Referring to Table I, some additional parameters that describe the surface wave behavior are provided. The phase speed of the surface wave, c_{surf} can be calculated from the following expression:

$$c_{\text{surf}} = \omega / \text{real}(k_1 p) = \text{real}(p \sqrt{(\lambda_1 + 2\mu_1 + i2\omega\mu_2)/\rho}). \quad (16)$$

If the system is viscoelastic, $\mu_2 \neq 0$, and p is complex, then it is dispersive with c_{surf} dependent on frequency, ω . If the system is elastic, $\mu_2 = 0$, and p is real, then it is nondispersive, with phase speed independent of frequency. Note that the dispersive nature of the phantom medium is evident, al-

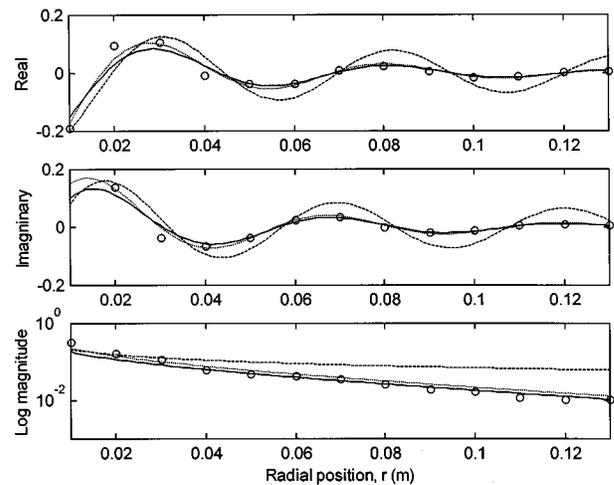


FIG. 5. Comparison of gel phantom experimental measurements with theoretical predictions of vertical particle velocity. Velocity is taken with respect to excitation force input \dot{u}_z/F_{in} (m/sN) on the surface as a function of radial position at excitation frequency $\omega/2\pi=40$ Hz. Key: $\circ \circ \circ$ experiment; — improved theory with damping, --- improved theory without damping ($\mu_2=0$); ··· asymptotic theory with damping.

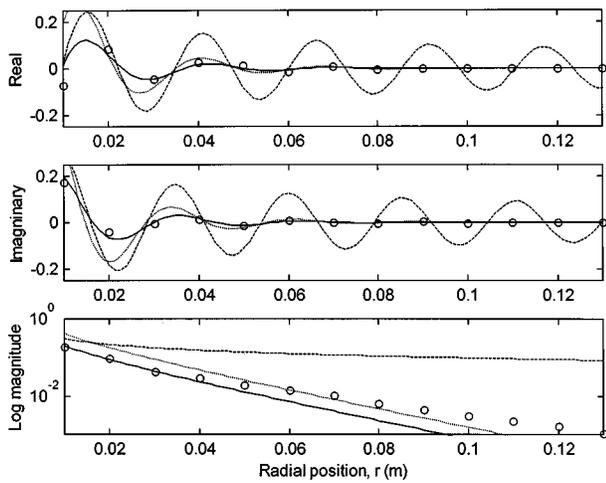


FIG. 6. Comparison of gel phantom experimental measurements with theoretical predictions of vertical particle velocity. Velocity is taken with respect to excitation force input \dot{u}_z/F_{in} (m/s/N) on the surface as a function of radial position at excitation frequency $\omega/2\pi=80$ Hz. Key: $\circ \circ \circ$ experiment; — improved theory with damping; --- improved theory without damping ($\mu_2=0$); ··· asymptotic theory with damping.

though its effect on phase speed is minimal for the range of 40 to 80 Hz. In terms of absolute value and trend, the phase speed of surface waves on the phantom agrees with those reported in the literature for human soft tissue.^{6,7,10,22} Very similar differences between the improved theory and asymptotic theory were predicted for horizontal particle velocity on the surface using the different theoretical models. Unfortunately, experimental comparison was not possible with the available instrumentation.

Having established the validity of the theoretical model, it is now used to investigate wave propagation below the surface. In Figs. 7 and 8, the predicted wave propagation profiles are shown for vertical and horizontal motion as a function of depth below the surface normalized by surface wavelength, L_{surf} . This facilitates comparison of results at

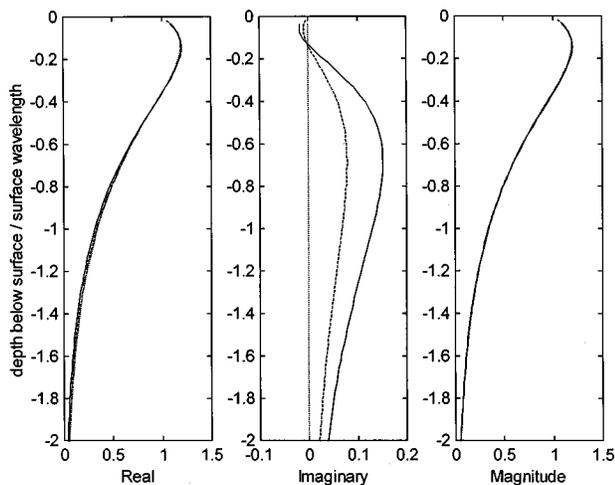


FIG. 7. Comparison of gel phantom theoretical predictions of vertical particle velocity below the surface normalized to the vertical velocity amplitude at the surface $\dot{u}_z(z)/\dot{u}_z(z=0)$ at selected excitation frequencies. Key: — improved theory with damping at $\omega/2\pi=80$ Hz; --- improved theory with damping at $\omega/2\pi=40$ Hz; ··· improved theory without damping at any frequency.

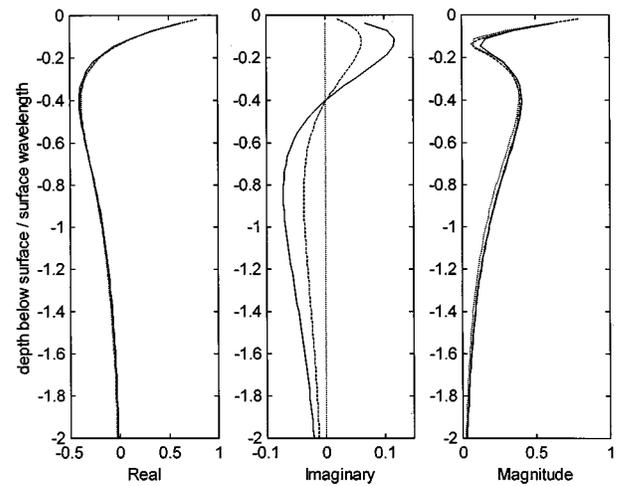


FIG. 8. Comparison of gel phantom theoretical predictions of horizontal particle velocity below the surface normalized to the vertical velocity amplitude at the surface $\dot{u}_r(z)/\dot{u}_z(z=0)$ at selected excitation frequencies. Key: — improved theory with damping at $\omega/2\pi=80$ Hz; --- improved theory with damping at $\omega/2\pi=40$ Hz; ··· improved theory without damping at any frequency.

different frequencies. It also renders predictions using the improved theory and asymptotic theory identical, since approximations in the asymptotic theory are not related to depth z but rather to radial position r and the radius of the disk. While for the purely elastic case the normalized quantities are independent of excitation frequency, for the viscoelastic case they have dependence on excitation frequency that is pronounced in the imaginary portion of the response. Note that vertical velocity amplitude actually increases slightly below the surface. Horizontal motion reverses direction. Both types of motion exponentially approach zero as depth increases. These predictions can be used to help assess the limitations of the experimental phantom model in simulating infinite half-space conditions. They can also be used to determine what frequency of surface waves may be altered by material anomalies located at different depths below the surface and at different radial distances from the source. In Fig. 9, calculations are shown of the amplitude of surface wave vertical-particle velocity \dot{u}_z as a function of normalized

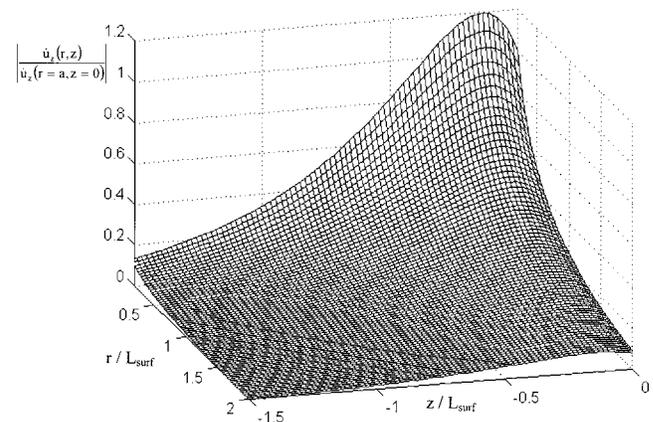


FIG. 9. Gel phantom theoretical prediction of the normalized amplitude of vertical particle velocity $|\dot{u}_z(r,z)/\dot{u}_z(r=a,z=0)|$ as a function of depth z and radial position r . Depth and radial position normalized with respect to surface wavelength L_{surf} . Results shown for $\omega/2\pi=40$ Hz.

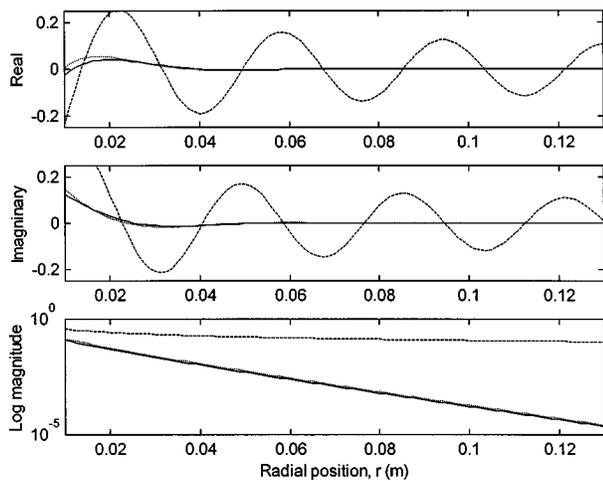


FIG. 10. Comparison of soft-tissue theoretical predictions of vertical particle velocity. Velocity is taken with respect to excitation force input \dot{u}_z/F_{in} (m/s/N) on the surface as a function of radial position at excitation frequency $\omega/2\pi=40$ Hz. Key: — improved theory with damping; --- improved theory without damping ($\mu_2=0$); ··· asymptotic theory with damping.

radial distance r and depth z at $\omega/2\pi=40$ Hz. Due to the dispersive nature of the viscoelastic medium, results at frequencies other than 40 Hz will not be identical but will be similar since dimensions have been normalized with respect to surface wavelength, L_{surf} .

B. Theory using human soft-tissue material values from literature

Human soft-tissue shear parameters estimated by von Gierke *et al.*²² and used by other researchers^{17,20} are different than those of the phantom used in the experimental study. Oestreicher's²⁰ material values are used in the theoretical model to generate the results shown in Figs. 10–14, which are analogous to Figs. 5–9 for the gel phantom but without the experimental data. The increased damping causes a sig-

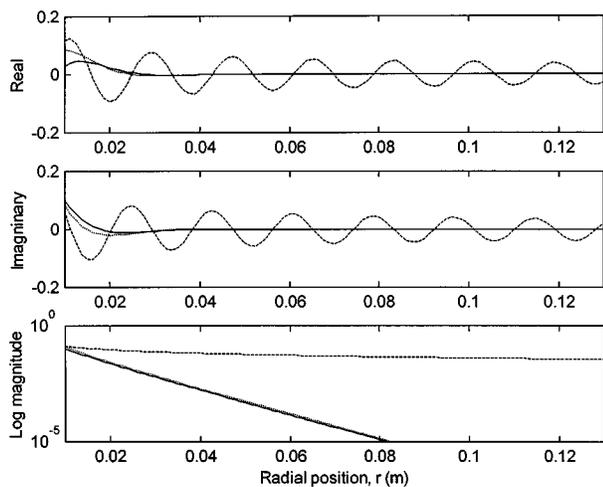


FIG. 11. Comparison of soft-tissue theoretical predictions of vertical particle velocity. Velocity is taken with respect to excitation force input \dot{u}_z/F_{in} (m/s/N) on the surface as a function of radial position at excitation frequency $\omega/2\pi=80$ Hz. Key: — improved theory with damping; --- improved theory without damping ($\mu_2=0$); ··· asymptotic theory with damping.

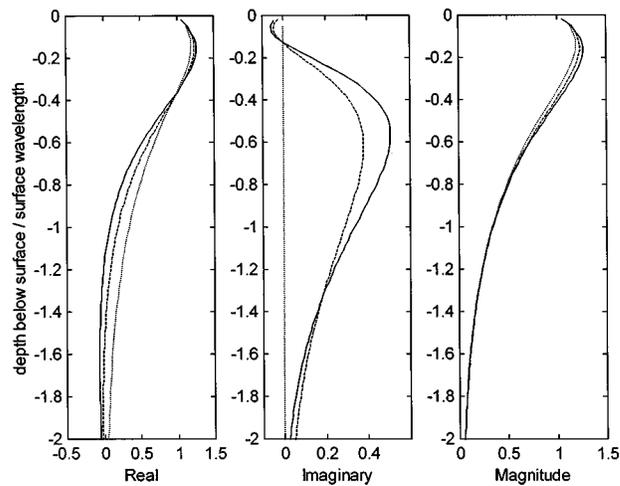


FIG. 12. Comparison of soft-tissue theoretical predictions of vertical particle velocity below the surface normalized to the vertical velocity amplitude at the surface $\dot{u}_z(z)/\dot{u}_z(z=0)$ at selected excitation frequencies. Key: — improved theory with damping at $\omega/2\pi=80$ Hz; --- improved theory with damping at $\omega/2\pi=40$ Hz; ··· improved theory without damping at any frequency.

nificantly sharper attenuation in response amplitude of the surface wave as one moves away from the source. While the magnitude of surface wave particle motion below the surface relative to the surface motion is comparable to that of the gel phantom, the imaginary portion of the response is increased. Referring to Table I, it is noted that the increased shear viscosity now has a much more significant impact on surface wave phase velocity c_{surf} and corresponding wavelength L_{surf} . The human-tissue model is significantly more dispersive.

These results do not agree quantitatively or qualitatively with some of the experimental measurements reported in the literature on human soft tissue. Indeed, in von Gierke *et al.*²² measurements taken over the thigh showed very little change in surface wave phase speed over the range of 10 to 180 Hz.

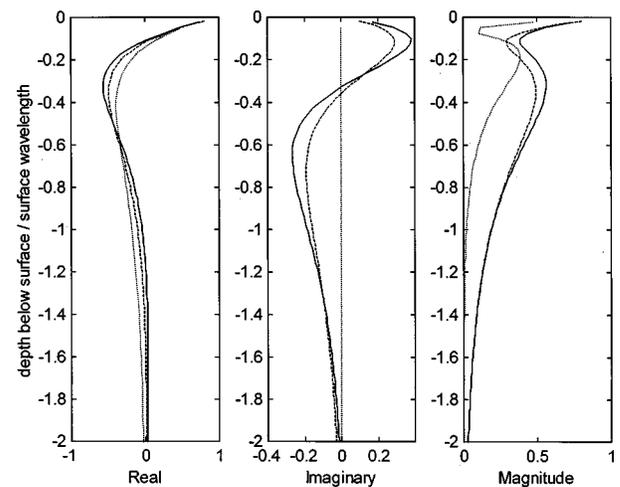


FIG. 13. Comparison of soft-tissue theoretical predictions of horizontal particle velocity below the surface normalized to the vertical velocity amplitude at the surface $\dot{u}_r(z)/\dot{u}_z(z=0)$ at selected excitation frequencies. Key: — improved theory with damping at $\omega/2\pi=80$ Hz; --- improved theory with damping at $\omega/2\pi=40$ Hz; ··· improved theory without damping at any frequency.

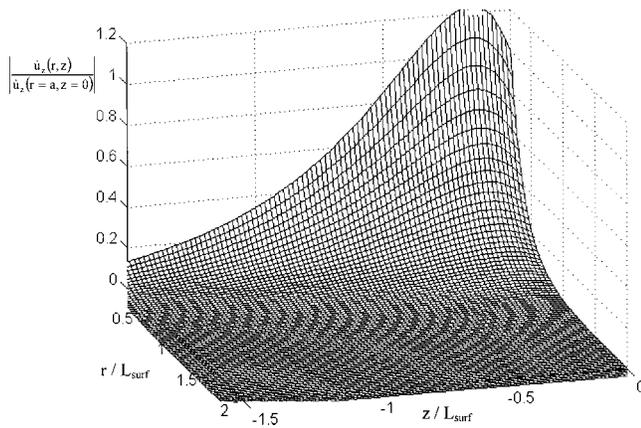


FIG. 14. Soft-tissue theoretical prediction of the normalized amplitude of vertical particle velocity $|\dot{u}_z(z)/\dot{u}_z(r=a, z=0)|$ as a function of depth z and radial position r . Depth and radial position normalized with respect to surface wavelength L_{surf} . Results shown for $\omega/2\pi=40$ Hz.

Additionally, in a stroboscopic photograph of response to excitation at 62 Hz, three crests and troughs of a surface wave propagating away from a piston source are clearly visible. Von Gierke *et al.*²² used several different approximate techniques that led to estimates of μ_2 ranging from 2.5–50 Ns/m². These estimates were primarily based on impedance measurements at the surface, not surface wave behavior. In other studies of surface waves in soft tissue, values of $\mu_2 = 2.5$ to 5 Ns/m² have been employed to obtain a good match between computation and experiment.^{4,10} Some estimates based on sonoelastic imaging have also been in this range.¹⁵ In other words, if one does use the soft-tissue parameters reported in von Gierke *et al.*²² and given in Table I of this paper, one does not get the experimental results reported by von Gierke *et al.*²² and others, at least in regard to surface wave speed and attenuation measurements.

Perhaps a more valid measure is the ratio of μ_2 to μ_1 . Experimental data found in von Gierke *et al.*²² and elsewhere¹⁵ support this. While values of μ_1 and μ_2 may vary by several factors for soft tissue from different parts of the anatomy or from the same location but based on different testing methods, the ratio of μ_2 to μ_1 is somewhat consistent (though not exact given the variability of biological tissue). It is this ratio that is an indicator of attenuation per cycle. For a value of $\mu_1 = 2500$ N/m², the observation made here that $\mu_2 = 15$ Ns/m² is too high is consistent with the experimental findings reported in the literature, even those of von Gierke *et al.*²²

Finally, the limited applicability of this study to the biological condition is noted. The model of a viscoelastic isotropic half-space does not take into account the layered variation of viscoelastic properties one encounters when moving from the skin surface through its layers to underlying tissue. Incorporation of these complicating effects is currently under investigation.

IV. CONCLUSION

An analytical solution has been developed and experimentally verified for the problem of surface wave generation on a viscoelastic half-space by a finite rigid circular disk

located on the surface and oscillating normal to it. The analytical solution is an incremental advancement of the theoretical work reported in Miller and Pursey^{18,19} that resulted in an asymptotic solution. Parametric studies reported here suggest that, for surface wave propagation on soft biological tissue, the limitations of the asymptotic solution may be more relevant than for the seismological applications for which the theoretical work of the reference was originally intended.

Since the application focus here is medical diagnostics, the analytical solution was verified experimentally using a viscoelastic phantom with material properties comparable to biological soft tissue. Surface vibration measurements were made using a noncontacting laser Doppler vibrometer. The reported results furnish the reader with some physical appreciation of the dimensions of surface waves on biological tissues in the low audible frequency range, 20–100 Hz. As the frequency increases further, these waves are attenuated at a higher and higher rate with respect to distance from the source. Reported results have also suggested that prior estimates in the literature for shear viscosity in human soft tissue are not accurate in the low audible frequency range.

It is hoped that the developments reported here will help to advance novel medical diagnostic techniques based on measuring variations in surface and shear wave behavior caused by elastic and/or density anomalies that may be indicators of particular illnesses or signify the presence of a gaseous inclusion. Research is in progress to extend the work to more realistic models that account for layered variations in material properties that are more indicative of the normal biological condition. The reported results may also help in the interpretation of surface-based measurements of vibration caused by subsurface biological sources, such as cardiovascular function, pulmonary disorders, or gastrointestinal activity.

ACKNOWLEDGMENTS

The financial support of the Whitaker Foundation under the BME grant program is acknowledged. S.-H. Kim is acknowledged for assistance with the experimental study. X. Zhang is acknowledged for assistance with the theoretical work.

- ¹J. Lee, "On the coupling and detection of motion between an artery with a localized lesion and its surrounding tissue," *J. Biomech.* **7**, 403–409 (1974).
- ²H.-D. Hong and M. D. Fox, "No touch pulse measurement by optical interferometry," *IEEE Trans. Biomed. Eng.* **41**, 1096–1099 (1994).
- ³H.-D. Hong and M. D. Fox, "Noninvasive detection of cardiovascular pulsations by optical Doppler techniques," *J. Biomed. Opt.* **2**, 382–390 (1997).
- ⁴R. O. Potts, D. A. Chrisman, and E. M. Buras, "The dynamic mechanical properties of human skin *in vivo*," *J. Biomech.* **16**, 365–372 (1983).
- ⁵M. Mridha and S. Odman, "Characterization of subcutaneous edema by mechanical impedance measurements," *J. Invest. Dermatol.* **85**, 575–578 (1985).
- ⁶M. Mridha, S. Odman, and P. A. Oberg, "Mechanical pulse wave propagation in gel, normal, and oedematous tissues," *J. Biomech.* **25**, 1213–1218 (1992).
- ⁷V. V. Kazakov and B. N. Klochkov, "Low frequency mechanical properties of the soft tissue of the human arm," *Biophysics* **34**, 742–747 (1989).

- ⁸J. M. Pereira, J. M. Mansour, and B. R. Davis, "Analysis of shear wave propagation in skin; application to an experimental procedure," *J. Biomech.* **23**, 745–751 (1990).
- ⁹J. M. Pereira, J. M. Mansour, and B. R. Davis, "Dynamic measurement of the viscoelastic properties of skin," *J. Biomech.* **24**, 157–162 (1991).
- ¹⁰B. N. Klochkov and A. V. Sokolov, "Waves in a layer of soft tissue overlying a hard-tissue half-space," *Acoust. Phys.* **40**, 244–248 (1994).
- ¹¹A. Sarvazyan, "Shear acoustic properties of soft biological tissues in medical diagnostics," *J. Acoust. Soc. Am.* **93**, 2329 (1993).
- ¹²S. Ganesan, C.-S. Man, and S. J. Lai-Fook, "Generation and detection of lung stress waves from the chest surface," *Respir. Physiol.* **110**, 19–32 (1997).
- ¹³C. Man, M. Jahed, S. J. Lai-Fook, and P. K. Bhagat, "Effect of pleural membrane on the propagation of Rayleigh-type surface waves in inflated lungs," *J. Appl. Mech.* **58**, 731–737 (1991).
- ¹⁴L. Gao, K. J. Parker, S. K. Alam, and R. M. Lerner, "Sonoelasticity imaging: Theory and experimental verification," *J. Acoust. Soc. Am.* **97**, 3875–3886 (1995).
- ¹⁵S. Catheline, F. Wu, and M. Fink, "A solution to diffraction biases in sonoelasticity: The acoustic impulse technique," *J. Acoust. Soc. Am.* **105**, 2941–2950 (1999).
- ¹⁶Y. Yamakoshi, J. Sato, and T. Sato, "Ultrasonic imaging of internal vibration of soft tissue under forced vibration," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **37**, 45–53 (1990).
- ¹⁷D. E. Goldman and H. E. Von Gierke, *The Effects of Shock and Vibration on Man* (ANSI, New York, 1960).
- ¹⁸G. F. Miller and H. Pursey, "The field and radiation impedance of mechanical radiators on the free surface of a semi-infinite isotropic solid," *Proc. R. Soc. London, Ser. A* **223**, 521–541 (1954).
- ¹⁹G. F. Miller and H. Pursey, "On the partition of energy between elastic waves in a semi-infinite solid," *Proc. R. Soc. London, Ser. A* **223**, 55–69 (1955).
- ²⁰H. L. Oestreicher, "Field and impedance of an oscillating sphere in a viscoelastic medium with an application to biophysics," *J. Acoust. Soc. Am.* **23**, 707–714 (1951).
- ²¹T. Hall, M. Bilgen, M. F. Insana, and T. A. Krouskop, "Phantom materials for elastography," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **44**, 1355–1365 (1997).
- ²²H. E. von Gierke, H. L. Oestreicher, E. K. Franke, H. O. Parrack, and W. W. von Wittern, "Physics of vibrations in living tissues," *J. Appl. Physiol.* **4**, 886–900 (1952).

Low-frequency whale sounds recorded on hydrophones moored in the eastern tropical Pacific

Kathleen M. Stafford

CIMRS/COAS, Oregon State University, Hatfield Marine Science Center, 2030 South Marine Science Drive, Newport, Oregon 97365

Sharon L. Nieukirk

CIMRS, Oregon State University, Hatfield Marine Science Center, Newport, Oregon 97365

Christopher G. Fox

NOAA, Pacific Marine Environmental Laboratory, Newport, Oregon 97365

(Received 13 May 1999; revised 28 August 1999; accepted 3 September 1999)

An array of autonomous hydrophones moored in the eastern tropical Pacific was monitored for one year to examine the occurrence of whale calls in this region. Six hydrophones which recorded from 0–40 Hz were placed at 8° N, 0°, and 8° S along longitudes 95° W and 110° W. Seven types of sounds believed to be produced by large whales were detected. These sound types were categorized as either moan-type (4) or pulse-type (3) calls. Three of the moan-type calls, and probably the fourth, may be attributed to blue whales. The source(s) of the remaining calls is unknown. All of the call types studied showed seasonal and geographical variation. There appeared to be segregation between northern and southern hemispheres, such that call types were recorded primarily on the northern hydrophones in the northern winter and others recorded primarily on the southern hemisphere hydrophones in the southern winter. More calls and more call types were recorded on the eastern hydrophones than on the western hydrophones. [S0001-4966(99)06012-9]

PACS numbers: 43.80.Ka [WA]

INTRODUCTION

The eastern tropical Pacific (ETP) is a biologically productive area (Wyrski, 1966) that supports a diverse community of fauna, including cetaceans (Reilly and Fiedler, 1994). Productivity is driven by several strong surface currents and fronts that result in substantial upwelling, especially east of the Galapagos and west of the Costa Rica Dome (Wyrski, 1964; Fiedler *et al.*, 1991). Previous studies have correlated zones of high cetacean abundance with these areas of high productivity (Volkov and Moroz, 1977; Reilly and Thayer, 1990; Wade and Gerrodette, 1993).

Blue whales (*Balaenoptera musculus*), Bryde's whales (*B. edeni*), and sperm whales (*Physeter macrocephalus*) are the most common large whales in the ETP (Wade and Friedrichsen, 1979; Reilly and Thayer, 1990). Other balaeopterid species that have been reported in this area are the humpback (*Megaptera novaeangliae*), minke (*B. acutorostrata*), fin (*B. physalus*), and sei (*B. borealis*) whales (Ramirez, 1989; Wade and Gerrodette, 1993). The populations of these species in the ETP were monitored sporadically by whaling nations and more regularly by efforts to monitor dolphin populations (Perrin *et al.*, 1985; Polacheck, 1987; Ramirez, 1988; Reilly, 1990; Reilly and Fiedler, 1994; Miyashita *et al.*, 1995). This monitoring effort has been almost exclusively by visual means, with little input from acoustic methods. Whitehead *et al.* (Whitehead and Arnborn, 1987; Whitehead *et al.*, 1989) used acoustics to monitor sperm whale populations around the Galapagos Islands, but this represented a highly localized, species-specific directed effort.

Acoustic surveys of cetacean habitats are a powerful means of identifying the species present, locating and tracking individuals, identifying stocks from regional dialects, and determining patterns of seasonal distribution and relative abundance (Clark and Ellison, 1989; Thompson *et al.*, 1992; Clark *et al.*, 1996; Clark and Fristrup, 1997; Clark and Charif, 1998; Moore *et al.*, 1998; Stafford *et al.*, 1998). In May 1996, the National Oceanic and Atmospheric Administration's Pacific Marine Environmental Laboratory (PMEL) deployed an array of six autonomous moored hydrophones along the East Pacific Rise (EPR) to monitor underwater seismicity in this area. These hydrophones have proved to be useful for monitoring low-frequency (0–40 Hz) whale calls in the vicinity of the deployments. A preliminary analysis of this data set revealed numerous low-frequency sounds that may be attributable to baleen whales.

METHODS

Moored hydrophones

Six hydrophone packages were moored along the EPR, an area known to be seismically active. The hydrophones were placed in the regions of 8° N, 0°, and 8° S along longitudes 95° W and 110° W (Fig. 1). The spacing between the hydrophones, designed around seismic parameters, was sufficiently great that it was unlikely that two or more hydrophones would simultaneously record the same sound from one animal. The hydrophones were deployed May 1996 and the data were recovered every 5–6 months. Each mooring package consisted of an anchor, an acoustic release, an au-

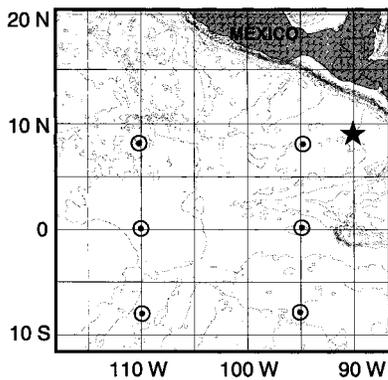


FIG. 1. Location of six autonomous hydrophones moored in the eastern tropical Pacific. Hydrophones are shown as \odot . The general vicinity of the Costa Rica Dome is shown as \star .

tonomous hydrophone logging system composed of an International Transducer Corporation 1032 hydrophone, preamplifier/filter (designed to prewhiten ocean ambient noise spectra from 1–40 Hz), and a digital recorder in a pressure-resistant titanium case, and flotation. The instruments were moored such that the hydrophones were suspended in the deep sound channel at depths of 650–750 m. The packages were designed to record for up to six months at a sampling rate of 100 Hz with lowpass filters set at 40 Hz. This sample rate was sufficient for recording energy from the very lowest frequency sounds of baleen whales. The data were archived within the instrument until it was recovered. They were then downloaded to 8-mm tape and the instrument redeployed.

Call detection

After discarding 198 hours of data from the hydrophone at $0^\circ 95^\circ W$ and 360 hours from the hydrophone at $0^\circ 110^\circ W$ due to extreme ambient noise (caused by cable strumming but also by periods of almost continual seismicity) that precluded any call detection, 47 682 total hours of data were examined for whale calls. Running spectrograms [fast Fourier transform (FFT) 256 points, overlap 50%] for each day for one year were examined visually for each of the six hydrophones (mid-May 1996–mid-May 1997). An exception to this data analysis was the hydrophone at $0^\circ 95^\circ W$, which did not log any data during the first deployment (mid-May 1996 to end of October 1996). Spectrograms were displayed by use of a program written in IDL[®] (Interactive Data Language, Research Systems, Inc., Boulder, Colorado).

Calls were identified as being potentially of biological origin based on both temporal and frequency characteristics of the sounds recorded. Sounds showing nonrandom temporal patterns [such as a series of similar sounds repeated at set intervals and interrupted by longer periods of silence (inferred to be breathing gaps) and/or calls that showed seasonality in detection] were considered good candidates for a biological source. Frequency characteristics used to identify possible biological sources included noncontinuous narrow-band sounds, frequency- or amplitude-modulated sound, and frequency content generally above 10 Hz. Most geologically produced sounds are very low frequency (<15 Hz), may last

over a minute, and are somewhat random in temporal occurrence (Dziak *et al.*, 1997). Shipping noise produces continuous bands of noise at narrow frequencies over relatively long periods of time, and geophysical survey noise such as air-guns tends to produce very broadband regularly spaced pulses without the irregular gaps seen in biological pulses (Richardson *et al.*, 1995). Using these criteria, seven different sound types were identified that suggested biological sources. Based on frequency content, these sounds were thought to have been produced by large whales.

These sound types were classified as either pulse-type calls or moan-type calls, and then further subdivided based on frequency and time characteristics. In this manner, three “known” moan-type calls, one unknown moan-type call, and three pulse-type calls were identified.

Presence or absence of the seven identified call types was recorded as a one or a zero for each hour of data examined. Measurements of time and frequency characteristics of each call type were digitized from a graphic workstation for calls with high signal-to-noise ratio and little coincident ambient noise. Characteristics recorded included call duration, beginning and ending frequencies, and frequency range. When appropriate, intercall intervals (the time from the end of one call to the beginning of the following tone or pulse), intergroup intervals (time between groups of pulses that appear to be in the same series), and long intervals (time between call series, usually greater than one minute, during which the calling animal is presumed to be at the surface) were measured. Call statistics were reported with standard deviations. Measured calls were then compared to published calls of species known to occur in the ETP.

RESULTS

Seven distinct low-frequency sounds potentially of biological origin were identified from data recorded on the six hydrophones. All of these sounds showed seasonal and regional variation (Table I). Although three of the moan-type calls can be attributed to blue whales, the origins of the other four calls are unknown to these authors. More calls and more types of calls were recorded on the hydrophones along longitude $95^\circ W$ than along $110^\circ W$. The most call types were recorded on the hydrophone at $8^\circ S 95^\circ W$ (six of the seven call types), and the most hours with calls was recorded at $8^\circ N 95^\circ W$.

Known sounds

Known call types detected on the hydrophones in the ETP consisted of three types of calls that have been attributed to blue whales. The first of these calls has been recorded in the northeast Pacific (Thompson *et al.*, 1996; Rivers, 1997; Stafford *et al.*, 1999), the second near Oahu (Thompson and Friedl, 1982), and the third off Chile (Cummings and Thompson, 1971).

Northeast Pacific blue whale calls

Northeast Pacific blue whale A–B calls [Fig. 2(a); reviewed in Rivers, 1997] were the most commonly recorded

TABLE I. Call occurrence by hydrophone. The number of hours during which at least one of each of the seven call types was recorded on each of the six hydrophones. The total number of hours of data collected for each hydrophone is given in the first line. Numbers in parentheses are percentages of total hours of data at each hydrophone in which a particular call was recorded.

	8° N 95° W	0 95° W	8° S 95° W	8° N 110 W	0 110 W	8° S 110 W
Total # hours of data	8735	4810	8771	8315	8347	8704
NE Pacific blue	4188 (50%)	215 (5%)	22 (0.2%)	49 (1%)	5 (<1%)	25 (<1%)
Southern blue	12 (<1%)	535 (11%)	1659 (20%)	56 (<1%)	129 (<1%)	202 (<1%)
NW Pacific blue	0	0	3 (<1%)	1 (<1%)	1 (<1%)	0
28-Hz moan	0	0	121 (1%)	7 (<1%)	4 (<1%)	35 (<1%)
Pulse series	250 (3%)	76 (2%)	1961 (22%)	271 (3%)	198 (2%)	435 (5%)
Pulse train	48 (<1%)	1 (<1%)	0	0	0	0
Short pulse	13 (<1%)	504 (10%)	1664 (19%)	2 (<1%)	1 (<1%)	2 (<1%)

call at the 8° N 95° W hydrophone and were detected to varying degrees on the remaining five hydrophones (Table I).

Series of repeated A–B calls were detected, as were A calls followed by multiple B calls. On a small number of occasions, only A calls were detected. C calls (Clark and Fristrup, 1997) were also detected, but less often. Figure 2(a) illustrates an A–B call with a C call. On average, the A calls measured from the ETP data lasted 21 s and had lowest components at 16 Hz (Table II). The B call swept from 17.7 to 16.1 Hz over 18.7 s. Time between the two call parts was 26.1 ± 3.0 s on average ($n=163$). When an A call was followed by more than one B call, the mean time between B calls was 31.8 ± 3.7 s ($n=48$). C calls swept up from 11.3 to 11.9 Hz over 7.7 s and followed A calls by 16.8 ± 2.3 s ($n=11$). The discrepancy between the C call and the A–C call sample sizes is due to a series of B–C calls where no A was visible.

These types of calls were recorded most often from November through May and less often from May through October [Fig. 2(b)]. However, A–B calls were detected in the area year-round (Stafford *et al.*, 1999).

Southern blue whales

The most complex calls recorded were two variations of calls which are similar to those reported for blue whales off Chile (Cummings and Thompson, 1971). These calls were recorded primarily on the 8° S 95° W hydrophone. The first variation was almost identical to the report of Cummings and Thompson (1971) and the second was similar enough to be classified with the first call [Fig. 2(c) and (d)].

The first call variation consisted of four parts; the first two were pulsive constant frequency calls with sidebands, the third was a simple short upswEEP, and the fourth was a pulsive frequency modulated (FM) call, also with sidebands. The average duration of the entire call was 38.8 ± 1.8 s ($n=23$). To maintain consistency with Cummings and Thompson (1971), the first two parts were labeled A and B, the fourth was labeled C, and the upswEEP was labeled D. (No upswEEP was reported in the original publication.) Part A's strongest component was at 21 Hz and lasted 13.4 s (Table II). Sidebands were at 7-Hz intervals ($n=25$). Part B followed part A by 2.3 ± 0.7 s, had its strongest component at 20.2 Hz, and lasted 9.1 s. Upper sidebands were also at 7-Hz

intervals for this call part. Part D was a short (1.9 s) call that swept up from 12.9 to 16.5 Hz and followed part B by 0.33 ± 1.0 s. The final part of the call, C, was a FM pulsive call with the strongest component sweeping from 27.9 to 25.5 Hz over 8.4 s. This call had sidebands at ~ 5 -Hz intervals. Time between successive calls was 68.3 ± 2.7 s ($n=21$).

The second variation of this call consisted of four-part, pulsive moans. A FM upswEEP was not associated with this variation. The first part of the call (A) was a 17.7-s long moan with slight frequency modulation from 19.9 to 19.1 Hz (Table II). This part of the call often had upper sidebands at 3.3-Hz intervals. The second part of this call was seen least often in the spectrograms. It was a FM pulsive downswEEP which started at 28.9 and ended at 27.5 Hz. Sidebands occurred at 7-Hz intervals. This part lasted 10.8 s. The third part (C) of this call followed the first by 18.3 ± 6.2 s. The C call was a short 4.0-s segment at 26.0 Hz with no significant frequency modulation. The intercall interval between C and D was 1.8 ± 0.8 s. The D call was also a very pulsive, FM moan with sidebands at ~ 6.4 -Hz intervals. This call part consisted of a 4.5-s unmodulated 26.0-Hz segment that then modulated down to 24.6 Hz during 7 s. Mean time between successive four-part calls was 63.1 ± 2.4 s ($n=15$). Often, only parts C and D were visible while scrolling through the data.

These call types were recorded on all six hydrophones but very rarely above the equator, and primarily at the hydrophones at 8° S 95° W and 8° S 110° W (Table I). The calls were seasonal, with the loudest calls detected from May until the end of August, with calls again detected from late January through May [Fig. 2(e)].

Western Pacific blue whale calls

Calls similar to those recorded for western Pacific blue whales [Fig. 2(f), Thompson and Friedl, 1982; Stafford and Fox, 1998] were detected a total of only five times, once each at 8° N 110° W and 0° 110° W and three times at 8° S 95° W. Although this call type is clearly rare in the study regions, it is included here to illustrate the potential extent of blue whale distribution in the Pacific.

Nine successive calls were loud enough to be digitized for information on time and frequency content. These calls were recorded on the 8° S 95° W hydrophone. Mean duration

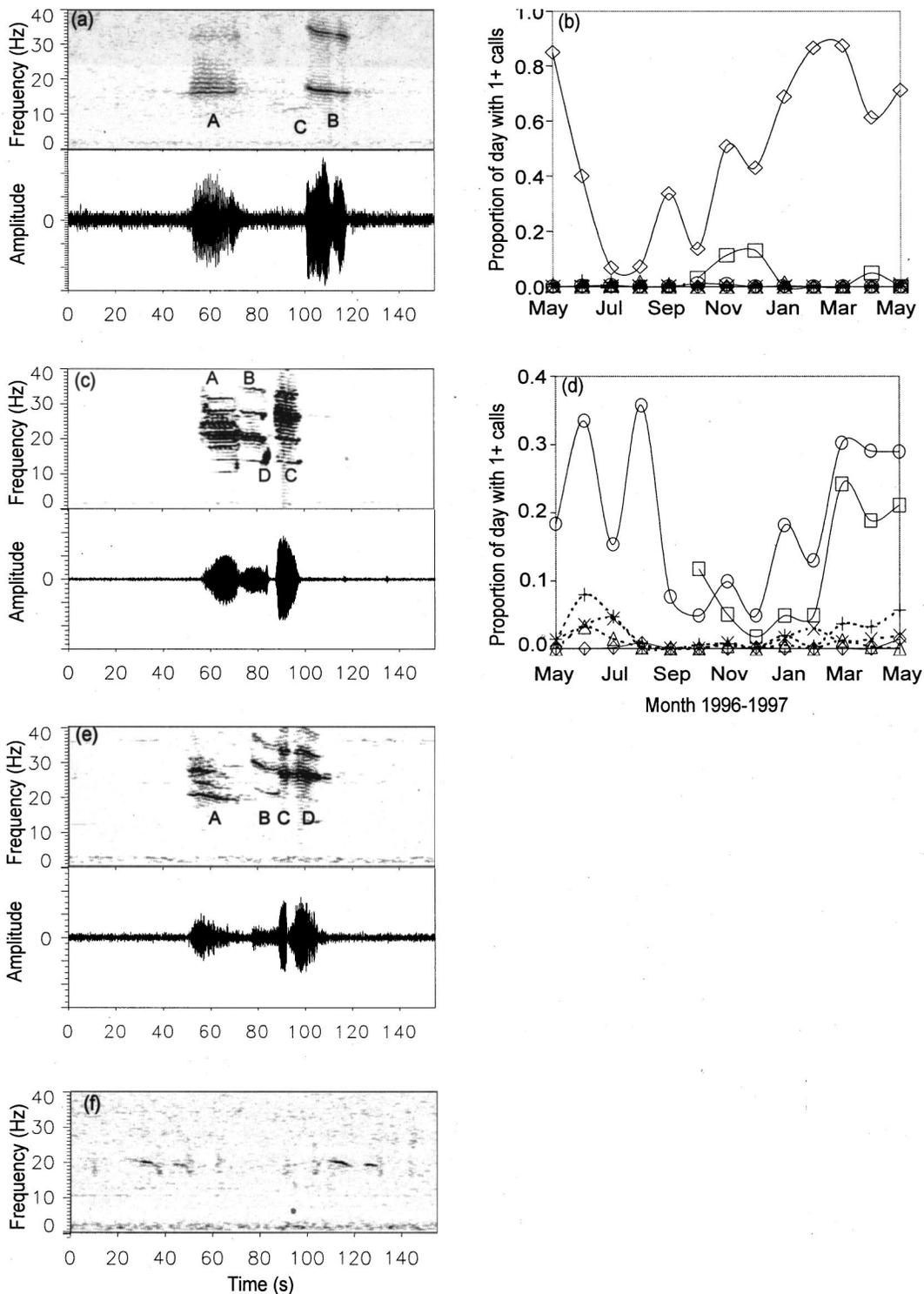


FIG. 2. Spectrograms (FFT 512 points, 94% overlap, analysis bandwidth 0.4 Hz, Hanning window), time series, and seasonal patterns of “known” calls recorded in the ETP. For the seasonality plots, lines represent the mean proportion of each day within a month where there were at least one or more calls recorded for the hydrophones located at 8° N 95° W (\diamond), 0° 95° W (\square), 8° S 95° W (\circ), 8° N 110° W (\triangle), 0° 110° W (\times), and 8° S 110° W ($+$). (a) Typical northeastern Pacific blue whale AB call. (b) Seasonal pattern of NE blue whale AB calls. These calls were primarily recorded on the hydrophone at 8° N 95° W. (c), (d) Spectrogram and corresponding time series of two types of southern blue whale calls recorded in the ETP. (e) Seasonal pattern of all southern blue whale calls. Both variations were recorded on all six hydrophones but primarily at or below the equator, and particularly at the phones at 8° S 95° W and 8° S 110° W. (f) Western Pacific blue whale call. Very few of these calls were recorded; therefore seasonal information is not provided.

of the calls was 21.5 ± 2.2 s. The first part of the calls consisted of a 10-s call that was not very frequency modulated (20 to 19.8 Hz) during the first 6.7 s but that swept down to 18.5 Hz during the last 3 s (Table II). A 5.6 ± 1.2 -s gap sepa-

rated the first from the second half of the call. The second half of the call began at 19 Hz and ended at 18.8 Hz and did not have significant frequency modulation throughout the 5.6-s duration [Fig. 2(f)].

TABLE II. Duration and frequency characteristics of moan-type calls recorded in the ETP, including standard deviations and sample sizes. Frequencies are given in Hertz (Hz) and durations (Dur) are given in seconds.

Call type	Part A			Part B			Part C			Part D		
	Start freq.	End freq.	Dur	Start freq.	End freq.	Dur	Start freq.	End freq.	Dur	Start freq.	End freq.	Dur
NEP blue whale	16.1±0.3 ^a		20.6±4.0	17.7±0.5 ^b	16.1±0.3	18.7±3.0	11.31±0.2 ^b	11.9±0.2	7.7±1.9	
	(n=163)			(n=318)			(n=83)					
“Southern” blue whale variation 1	21.0±0.6 ^a		13.4±1.1	20.2±0.3 ^a		9.1±1.2	27.9±0.8 ^b	25.5±0.6	8.4±0.6	12.9±0.3 ^b	16.5±0.9	1.9±0.5
	(n=25)			(n=25)			(n=25)			(n=25)		
“Southern” blue whale variation 2	19.9±0.2 ^b	19.1±0.4	17.7±3.8	28.9±1.6 ^b	27.5±0.4	10.8±0.6	26.0±0.3		4.0±0.5	26.0±0.3 ^b	24.6±0.4	10.8±1.5
	(n=23)			(n=9)			(n=23)			(n=23)		
Western blue whale	20.1±0.1 ^b	18.5±0.2	9.7±1.0	19.2±0.3	18.8±0.1	5.6±0.3
	(n=9)			(n=9)								
28-Hz moan	28.3±0.2 ^a		9.3±2.2	20.6±1.4 ^b	19.2±1.5	10.7±2.2
	(n=547)			(n=547)								

^aAM.

^bFM.

Unknown sounds

A number of calls were observed that are presumed to be of biological, and very likely cetacean, origin. These calls were loud, low-frequency calls with repetition rates and/or frequency modulations that visually and audibly resemble whale calls. They were divided into two groups: moan-type calls and pulse-type calls. In the moan-type call group were 28-Hz moans. The pulse-type calls included three types designated “pulse series,” “pulse trains,” and “short pulses.” An additional sound type, called “upsweep,” is probably not of biologic origin. This sound was recorded on all six hydrophones and is presented here due to its prevalence in the data set.

28-Hz moans

The 28-Hz moan was the one unidentified moan-type call. It was recorded chiefly on the southern hydrophones [Fig. 3(a)]. These sounds were not recorded very often (about 1% of the time on 8° S 95° W), but were distinctive enough to be identified as a unique call type in the data set. These calls consisted of two overlapping parts. The first part was a pulsive call at 28.3 Hz which lasted on average 9.3 s (Table II). This part was followed by a FM moan from 20.6 to 19.2 Hz over 10.7 s. The overlap between the two was 1.5±2.5 s (n=157). The repetition rate from the end of one call to the beginning of the next was 45.0±10.2 s. These calls were only recorded with regularity in June and July on the hydrophones moored at 8° S 95° W and 8° S 110° W [Fig. 3(b)].

Pulse series

The second most common call type recorded was a series of low-frequency pulses which were characterized by very regular interpulse intervals and very constant bandwidth [Fig. 3(c)]. The pulses occurred in groups of 1–14 pulses per group with interpulse intervals of 11.1 s (Table III). Time between groups of pulses was 26.4 s and the mean long interval periods between bouts of pulse groups was 217.8

±158.4 s (n=46). Average highest frequencies were 27.5 Hz, lowest 14.8 Hz, for a total bandwidth of 12.7±2.1 Hz. Each pulse lasted about 1.4 s.

These calls were recorded most often on the hydrophone at 8° S 95° W but were recorded on all of the hydrophones. The seasonality of calls recorded on all hydrophones was similar; calls were detected from about June through October with peak occurrence in July–September [Fig. 3(d)]. These calls were recorded most often on the hydrophones along 95° W, although the occurrence of this call type dropped off more quickly on the hydrophone at 8° N 95° W than at 8° S. This call type was not recorded from December 1996 through May 1997 on any hydrophone save the one at 8° N 95° W.

Pulse trains

Fifty-three pulse trains were recorded on the hydrophone at 8° N 95° W, and a single pulse train was recorded on the 0° 95° W hydrophone [Fig. 3(e)]. These calls were comprised of 6–25 downsweeps spaced 4.8 s apart (Table III), each lasting 3.1 s. On average, the sounds swept from 38.4 to 25.2 Hz; however, higher frequencies may have existed above the 40-Hz cutoff. Although there were relatively fewer detections of these pulse trains, they appear to be somewhat seasonal. They were occasionally detected (<1%) in the winter months from mid-November 1996 through late January 1997, and then in April but in no other months [Fig. 3(f)].

Short pulses

Short, narrow-band pulses were detected on all of the hydrophones but only a few times on the western hydrophones. These calls were predominantly recorded on two hydrophones, those at 8° S 95° W and 0° 95° W. These pulses were seldom seen in temporally close groupings. Usually only one or two per hour were recorded. Occasionally they were very loud [Fig. 3(g)]. The average bandwidth of these

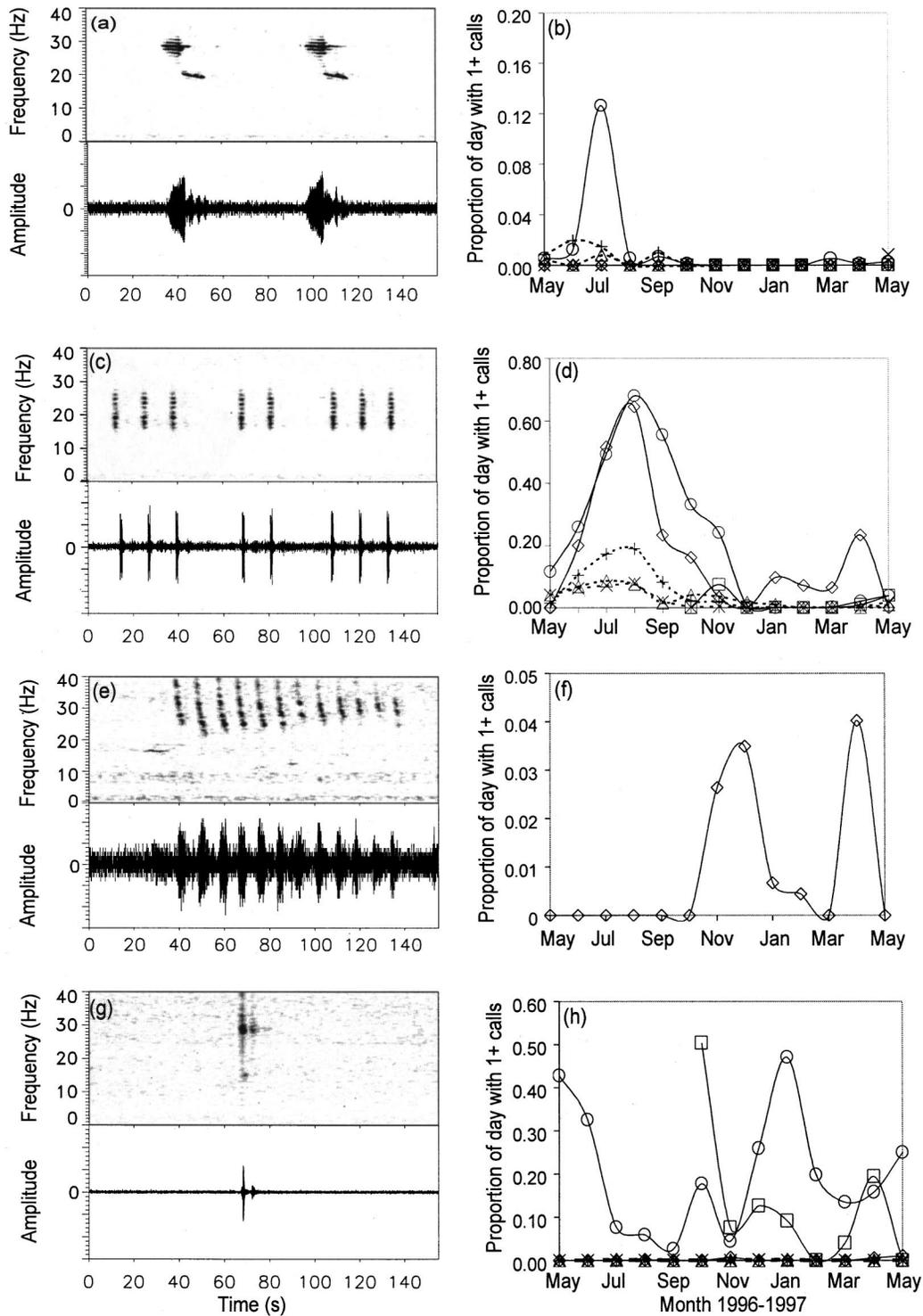


FIG. 3. Spectrograms (FFT 512 points, 94% overlap, analysis bandwidth 0.4 Hz, Hanning window), time series and seasonal patterns of “unknown” calls recorded in the ETP. For the seasonality plots, lines represent the mean proportion of each day within a month where there were at least one or more calls recorded for the hydrophones located at 8° N 95° W (\diamond), 0° 95° W (\square), 8° S 95° W (\circ), 8° N 110° W (\triangle), 0° 110° W (\times), and 8° S 110° W ($+$). (a) Two 28-Hz moans. (b) Seasonal pattern of the 28-Hz moan. This call was recorded primarily on the 8° S 95° W hydrophone. (c) Three groups of pulses of a pulse series. (d) Seasonal pattern of pulse series calls. These calls were recorded most often on the hydrophone at 8° S 95° W but were seen on all of the hydrophones. Calls were detected from about June through October with peaks in occurrence July–September. (e) One pulse train. (f) Seasonal pattern of pulse trains. Virtually all pulse trains were recorded on the 8° N 95° W hydrophone. (g) Short pulse. These short (~1 s), narrowband (~3 Hz) pulses were seldom seen in temporally close groupings; usually only one or two per hour were recorded. (h) Seasonal pattern of short pulses. These calls were detected on all 6 hydrophones but were primarily heard during October on the 0° 95° W hydrophone and the months of December–January and May on the 8° S 95° W hydrophone.

TABLE III. Time and frequency characteristics of pulse-type calls recorded in the ETP, including standard deviations and sample sizes. Frequencies are given in Hertz (Hz), duration and intervals in seconds (s). IPI = Interpulse interval, IGI = Intergroup interval, PPG = pulse per group.

Call type	Maximum frequency	Minimum frequency	Duration	IPI	IGI	PPG
Pulse series	27.5±1.5 (n = 2985)	14.8±0.9 (n = 2985)	1.4±0.3 (n = 2985)	11.1±0.8 (n = 2464)	26.4±3.8 (n = 895)	3.5±1.6 (n = 950)
Pulse trains	38.4±0.8 (n = 569)	25.2±2.5 (n = 569)	3.1±0.6 (n = 569)	4.8±0.8 (n = 530)	...	14.6±4.1 (n = 39)
Short pulses	30.4±1.1 (n = 149)	26.8±0.9 (n = 149)	1.3±0.3 (n = 149)	1

pulses was 3.5 ± 1.8 between 26 and 31 Hz (Table III). Pulse duration was 1.3 s. These calls were heard most often during October, trailing off from November–January, and then were heard again in April on the 95° W hydrophone. They were most prevalent in the months of December–May, peaking in January, on the 8° S 95° W hydrophone [Fig. 3(h)].

Upsweeps

One of the more common sounds recorded on all of the hydrophones was a long, low-frequency upsweep (Dziak *et al.*, 1997) [Fig. 4(a)]. It is unlikely that these sounds are of

biologic origin. Because they present a considerable source of low-frequency noise in the Pacific, and particularly in this data set, they are included here. Beginning and ending frequencies were measured for 1112 upswept sounds. Mean low frequency was 22.6 ± 3.9 Hz (range 11.5 to 33.2 Hz) and mean upper frequency was 32.0 ± 4.2 Hz (range 15.8 to 39.6 Hz). On average, these sounds swept up over $9.4 \text{ Hz} \pm 3.6$ during 8 ± 5.2 s ($n = 1112$). Upsweeps were heard on all hydrophones and a rough source location ($54^\circ \text{ S } 140^\circ \text{ W}$) was determined by use of seismic methods. This location corresponds to a site of underwater volcanic activity (Talandier and Okal, 1996). The occurrence pattern of upsweeps [Fig. 4(b)] was similar among the hydrophones, and showed some seasonality, which may be due to long-range propagation effects (Urlick, 1983).

DISCUSSION

It is likely that marine mammals make sounds for a variety of reasons, including social signaling, echolocation, communication, sexual display, and conveying reproductive status. Many baleen whales produce loud, lower frequency (<1000 Hz) sounds capable of traveling long distances in their underwater environment (Richardson *et al.*, 1995). Clark (1990) classified mysticete call types into three general categories: simple calls (narrow-band, low-frequency, modulated calls), complex calls (broadband, amplitude and/or frequency modulated, often pulsive calls), and short-duration calls (clicks, pulses, knocks, and grunts that may be less than 0.1-s long and are not frequency modulated). The ETP recordings include calls that fall into each of these categories. Three of these calls closely resemble calls well-documented by other researchers and are considered “known calls.” Comparison of the remaining unknown calls with the literature yielded less definitive results.

Known calls

Northeastern Pacific blue whale calls

The calls recorded most often on the ETP data were blue whale calls, which have been recorded throughout the northeast Pacific (Stafford *et al.*, 1999). Northeast Pacific blue whale calls are the best-known blue whale calls to date. These calls have been recorded along the west coast of North America from the Gulf of California, Mexico, to off Vancouver Island, Canada (reviewed in Rivers, 1997), and recently in the ETP (Stafford *et al.*, 1999). The occurrence of these calls showed a seasonal pattern with most recorded from

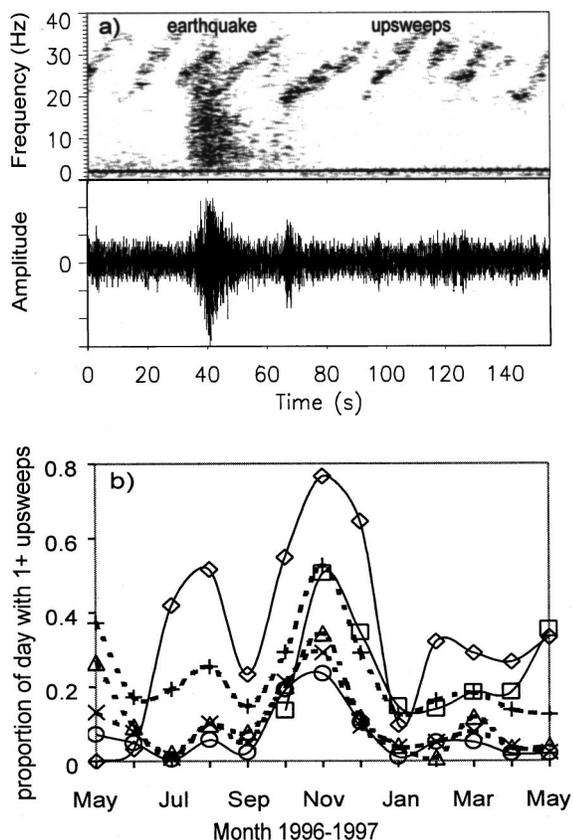


FIG. 4. (a) The spectrogram and corresponding time series of the upsweeps recorded in the ETP (FFT 512 points, 94% overlap, analysis bandwidth 0.4 Hz, Hanning window). Upsweeps were one of the more prevalent sounds recorded on all six of the moored hydrophones. It is unlikely that these sounds are of biologic origin. (b) Seasonal pattern of upsweeps. Lines represent the mean proportion of each day within a month where there were at least one or more upsweeps recorded for the hydrophones located at 8° N 95° W (\diamond), 0° 95° W (\square), 8° S 95° W (\circ), 8° N 110° W (Δ), 0° 110° W (\times), and 8° S 110° W (+).

TABLE IV. Comparison of temporal characteristics of calls resembling southern blue whale calls. Measurements for whale I and whale IV were taken from Cummings and Thompson, 1971, while measurements for variation I and variation II are from this experiment. Parts of calls are labeled A, B, C, and D after Cummings and Thompson, 1971, and the durations of these parts were measured in seconds. AM=amplitude modulated (pulses/s), gap=time (s) between parts of a cell, total duration (total dur) is the duration (s) of the entire call, and intercall duration (intercall dur) was measured from the end of one call to the beginning of the next call.

Call type	Part A (s)	AM A	Gap A-B	Part B (s)	AM B	Gap B-C	Part C (s)	AM C	Total dur.	Intercall dur.
Whale I	15.3	3.8	0	10.3	7.7	2.5	9.0	7.7	36.9	69.1
Whale IV	13.4	3.8	1.6	9.6	7.7	3.7	7.9	7.7	36.2	63.8
Variation I	13.4	6.9	2.3	9.1	6.9	6.1	8.4	none	38.8	68.3
Variation II	17.7	3.5	18.3	4	none	1.8	11	6.4	53.5	63.1

November through May; calls were recorded less often from May through October. However, A-B calls were detected in this area to some degree year-round. These data support Reiley and Thayer's (1990) hypothesis that blue whales are in the ETP all year. Furthermore, instead of the suggestion that northern and southern hemisphere populations of blue whales share this area alternately, it appears that a part of one population (NE Pacific) is in residence year-round.

Southern blue whale calls

One of the call variations attributed here to "southern" blue whales [Fig. 2(c)] very closely resembles the calls recorded in the presence of two blue whales off Chile (Cummings and Thompson, 1971). The primary difference between the two is the upswEEP reported here. The total duration of the calls was similar, the amplitude modulation of the three pulsive moan parts was similar for parts B and C, and the duration and spacing of the three parts were similar (Table IV).

The second variation of this call also consisted of three pulsive-moan parts [Fig. 2(d)]. The interpart intervals for these calls were different from the first variation reported here in that the gap between parts A and B for two different animals was 0–1.6 s, and the time between parts B and C was only 2.5–3.7 s. Durations of each call part also differed between the two animals recorded off Chile and the calls reported here.

All of the calls discussed here are similar in frequency content, however. It was difficult to compare the exact frequency contents of the two call types, as Cummings and Thompson (1971) stated only that the loudest call components were centered at 20, 25, and 31.5 Hz. Both of the variations reported here had most of their energy in bands between 20 and 26 Hz. An additional similarity is in the amplitude modulation of the call parts. Cummings and Thompson (1971) report a modulation rate of part A of 3.85 pulses/s and part B of 7.7 pulses/s. The calls from 8° S 95° W had amplitude modulation from 3.5 to 6.9 pulses/s.

These calls were recorded on all six hydrophones, but primarily on the hydrophones at or below the equator, and particularly at the hydrophones at 8° S 95° W and 8° S 110° W. A seasonal pattern is evident, with the loudest calls detected from May until the end of August, and then calls again detected from late January through May [Fig. 2(e)]. This seasonality might be interpreted as a southern

stock, which spends the southern hemisphere summer in Antarctica and migrates toward the equator during the southern winter.

The acoustic data presented here support a northern hemisphere/southern hemisphere separation between blue whale populations. The calls that have been recorded in the presence of blue whales along the west coast of North America were also recorded in the ETP during what might be considered the "breeding" season for a population that spends the summer engaged in feeding activities farther north (Lockyer, 1984). Although A-B calls were detected on the hydrophones south of the equator, they were detected much less often there than at the equator and north of the equator. This geographic segregation suggests that the A-B call type and the population of blue whales that make this call may be restricted to the northeastern Pacific. The "southern" blue whale call data complement the northern data. These calls were recorded most often in March–August, which comprises the austral summer/fall, and are primarily recorded south of the equator. The similarity of the two variations observed with that recorded by Cummings and Thompson (1971) off of Chile in May 1970 suggests a migratory pattern up the west coast of South America.

Western pacific blue whale calls

Western Pacific blue whale calls have been recorded primarily in the northwestern Pacific, but also north of Hawaii and in the offshore northeast Pacific (Thompson and Friedl, 1982; Stafford and Fox, 1996). These calls were detected only during five hours of the total of 47 682 hours monitored. Therefore they are not considered a common call in the ETP. Nevertheless, they are included here to exhibit the extent of this call type in the North Pacific.

Unknown calls

The origin of the other four biological sounds recorded in the ETP is presently unknown. Other balaenopterid species that have been identified in the area covered by the hydrophone array are the Bryde's, humpback, minke, fin, and sei whales (Ramirez, 1988; Wade and Gerrodette, 1993). Because only the 0–40-Hz frequency band was recorded on this data set, some sounds may represent the low-frequency components of calls reported at higher frequencies, or they may be calls that have not been reported previously.

28-Hz moan

The 28-Hz moan resembles the northeastern Pacific (NEP) blue whale call in its structure and components in that it consists of a first part, which is an amplitude-modulated pulsive call, followed by a FM moan. However, the pulsive part is higher in frequency than the moan that follows and slower in modulation than the NEP blue whale A call (1 vs 1.5 pulse/s). Additionally, the two parts overlap by up to 4 s and the individual parts are about half as long as those made by NEP blue whales. Calls whose spectrograms look similar to the 28-Hz moan were recorded in the presence of blue whales off Madagascar in December 1996 (Ljungblad *et al.*, 1998). "Unit A" of the Madagascar calls begins with a short amplitude modulation (AM) which is similar to the first part of the 28-Hz moan. This is followed by a 38-Hz tone (Ljungblad *et al.*, 1998) which is higher in frequency than the 28-Hz moan discussed here. However, the structure of the two calls is similar. No "unit B" was detected in the ETP data. At present, blue whales may be the best candidate for this relatively rare call type.

Pulse series

The seasonality of these pulse series was similar on all six hydrophones; calls were detected from mid-May through November, and calls were recorded primarily on the 8° S 95° W hydrophone. The predominant occurrence of this call type on all hydrophones was mid-June to early December. This distinct seasonal pattern could be interpreted as a southern hemisphere whale stock migrating equatorward in the austral winter.

Fin whales may have produced these pulse series. The 20-Hz pulses of fin whales are perhaps the best documented low-frequency whale calls. These pulses have been recorded in the Atlantic and Pacific Oceans, and in the Gulf of California (Watkins, 1981; Watkins *et al.*, 1987; Edds, 1998; Thompson *et al.*, 1992). Common characteristics of the 20-Hz pulses were frequency modulation from 25–44 to 16–20 Hz, short (0.5–1 s) durations, and regular interpulse spacing (Thompson *et al.*, 1992). The pulse series described here share all of these characteristics. The production of pulses in groups, with the groups regularly spaced, has not been reported for this species and may represent an example of a regional call difference (e.g., Thompson *et al.*, 1992). Pulse series recorded in the north Atlantic were similar to those shown here and have been attributed to fin Whales (Clark and Charif, 1998).

Historically, fin whales were regularly taken in coastal waters off Peru, south of the array, from 3° 30' S to 8° S to 200 nm offshore (Ramirez, 1988), and a large group was killed off Ecuador in October–November 1926 (reported in Clarke, 1962). However, fin whales were seldom identified during Southwest Fisheries Science Center (SWFSC) dolphin surveys of the ETP between July and December from 1986–1990 (Wade and Gerrodette, 1993) and not at all during an October–November 1959 survey of the Galapagos region (Clarke, 1962), which contradicts fin whale origin of the pulse series, unless that species is more abundant in the ETP than previously reported.

Pulse trains

The recorded pulse trains somewhat resemble those reported for minke whales. However, the duration of the individual pulses reported here is much longer than other reports (Winn and Perkins, 1976; Nishimura and Conlon, 1994). These calls were recorded infrequently, and virtually all were recorded on the hydrophone at 8° N 95° W. Very few minke whales have been identified in the ETP during visual surveys (Wade and Gerrodette, 1993; Aguayo *et al.*, 1998), and direct comparison of these data with the published literature is difficult as there is no very low-frequency information from on-site recordings of minke whale vocalizations. Winn and Perkins (1976) made recordings in the presence of minke whales in the Atlantic. The lower-end response of their recording systems varied from 20 to 300 Hz. They recorded grunt-, pulse train-, ratchet- and pinglike sounds during three encounters with minke whales. The frequency of recorded sounds ranged from about 70–14 000 Hz. Duration of the pulses recorded was on the order of milliseconds. Schevill and Watkins (1972) made recordings in the presence of minke whales along the ice edge in Antarctica. Their recording response was flat from 30–30 000 Hz. Calls that swept in frequency from about 120–60 Hz and lasted 0.2–0.3 s were recorded. The interval between successive sweeps ranged from 8–97 s. The pulse train recordings described here were lower in frequency (center frequency ~30 Hz) and individual pulses were much longer in duration (~3 s) than other studies describing minke whale sounds. Therefore, it is not prudent to attribute these sounds to minke whales. The relative rarity of both the known presence of minke whales in the ETP and the occurrence of pulse trains recorded in the ETP is nevertheless intriguing.

The source of the short pulses is a mystery. These calls usually occurred once or twice an hour and did not seem affiliated with any other low-frequency sounds. It is possible that they may be a low-frequency component of a higher-frequency call.

Calls from large whale species known to occur in the ETP

Bryde's, humpback, and sei whales have been sighted and identified in the area of the hydrophone array. Humpback whales excepted, the acoustic repertoire of these species is almost unknown. A comparison of what is known of the sounds made by these species and the unknown sounds recorded on the array discussed here is not fruitful because of the lack of knowledge of species-specific repertoires and of the function of those sounds, and because of the 40-Hz cutoff imposed here by the geophysical experiment. For example, Bryde's whales are by far the most abundant baleen whale in the ETP. Their estimated population numbers exceed those of blue whales by an order of magnitude (Wade and Friedrichsen, 1979; Reilly and Thayer, 1990; Wade and Gerrodette, 1993). In some parts of the ETP they are found year-round (Ramirez, 1988). Although relatively little is known about their acoustic repertoire, if they make very low-frequency sounds, they should be detected more often than blue whale sounds. Should there be ten times as many calls from Bryde's as from blue whales, because visual surveys

detect approximately ten times as many individuals? Only if the sounds they make serve the same purpose at the same place and time of the year. It is possible that some of the sounds recorded in the ETP were made by this species. But because this experiment was constrained by the sampling requirements of a geophysical experiment, sense cannot be made of the above arguments if indeed Bryde's whales (and other whales) call at frequencies above 40 Hz. Based on reports in the literature, this calling/recording discrepancy seems to be the case.

Only two published reports address calls attributed to Bryde's whales. Edds *et al.* (1993) recorded calls from both captive (Atlantic) and free-ranging (Gulf of California) Bryde's whales. Both pulsive and moan-type calls were recorded from the captive animal with call frequencies between 100 and 900 Hz (minimum recording system response was 100 Hz) and the sounds lasted up to 51 s. Frequency content was similar for calls recorded in the Gulf of California but the duration of all call types was less than 1 s. In contrast, Cummings *et al.* (1986), recorded lower-frequency moans in the presence of Bryde's whales in the Gulf of California. These calls had most energy around 120 Hz (70–245 Hz range) and lasted about 0.4 s. Most of these sounds exhibited some frequency modulation. The response of the recording system went to 25 Hz so lower-frequency calls, were they made, should have been detected.

The distribution of sei whales in the ETP is not clear. Sightings of this species were not differentiated from those of Bryde's whales during the SWFSC surveys (Wade and Gerrodette, 1993). Peruvian whaling reports began to distinguish between the two species in 1974 and noted that sei whales were caught most often off the coast of Peru from August–October. These animals were caught nearer to shore (82 W–80° W) but in the same latitude (4 S to 7° S) as the hydrophone array discussed here. The number of Bryde's whales caught in the same fishery outnumbered sei whales by 8:1 (Ramirez, 1988). Little is known about sei whale sounds. Knowlton *et al.* (1991) and Thompson *et al.* (1979) recorded short, 1.5–3.5 kHz sounds in the presence of sei whales off the coast of Nova Scotia. In the latter study, the frequency response of the hydrophone was a flat 50–7500 Hz, and the frequency response of the Knowlton *et al.* (1991) hydrophone is unknown.

Most of the energy in humpback whale calls is well above our cutoff threshold of 40 Hz. Previous studies have recorded moans, grunts, pulse trains, and songs that included components as low as 20 Hz (Thompson *et al.*, 1986; Richardson *et al.*, 1995). Although some of the sounds recorded during our study may be humpback calls or parts of humpback songs, identification of a humpback call without the important upper-frequency components is difficult.

There were many instances of clicklike sounds that appeared to have higher-frequency components which could have been sperm whale clicks. Sperm whales are the most abundant large cetacean in the ETP (Wade and Friedrichsen, 1979). Sperm whales may produce clicks with some energy below 100 Hz but the duration of individual clicks (2–30 ms) is much shorter than any of the sounds reported here (Backus and Schevill, 1966; Watkins, 1980).

Call distribution

In this study, more calls and more types of calls were recorded on the hydrophones along longitude 95° W than along 110° W. Productivity differences between the two areas probably influence cetacean distribution. Productivity is generally higher east of 98° W. The 8° N 95° W hydrophone is southwest of the Costa Rica Dome, an area of high productivity and known cetacean habitat, and the 0° 95° W and 8° S 95° W hydrophones are close to the Galapagos Islands and Peru current, two other productive areas of known cetacean abundance (Wade and Gerrodette, 1993). In addition, north–south differences in call were recorded by the array. The largest number of calls was recorded at 8° N 95° W, an area where blue whales concentrate (Reilly and Thayer, 1990), and the majority of these calls were northeastern Pacific blue whale calls. The hydrophone at 8° S 95° W recorded the highest variety of calls (six of the seven defined call types). Volkov and Moroz (1977) described the distribution of ‘baleen whales’ (with no reference to any species) in the ETP from January–July 1975 as alternating between north–south zones of higher and lower abundance. Zones of high abundance included 10°–7° N and 2°–1° N to 2°–3° S and appeared to be associated with more productive areas. If number of calls reflects whale abundance, then the region east of the EPR was an area of higher occurrence of whales.

CONCLUSIONS

A tremendous amount of low-frequency biological (and very probably cetacean) sound is present in the eastern tropical Pacific. By using moored hydrophones, animal sounds can be monitored continuously for long periods at low cost relative to visual surveys. Acoustic observations may be used to establish presence or absence of species, to determine patterns of seasonal distribution and relative abundance, and to identify species and possibly regional stocks. However, current ignorance of the complete vocal repertoire of most species of large whales makes it difficult to rely upon acoustic surveys alone in determining presence or absence of a particular species. Many calls were recorded on the hydrophone array during this experiment that cannot be attributed to specific species, or even genera. This lack of knowledge underlines the need for more recordings in the presence of large whales. A better understanding of movements and winter-time distributions of these animals might also allow us to rule out some species and better describe others.

Although the hydrophone array described here was deployed for seismic studies, it has proved very useful for looking at low-frequency whale calls at six locations in the ETP. These data represent the first long-term recordings made in the ETP to be examined for whale calls. Most acoustic surveys of the nature presented here have been opportunistic and consequently, the focus of these studies has not included frequencies above 100 Hz. Conversely, until recently, most directed marine mammal studies have used higher-frequency equipment with little regard for lower frequencies. The result is little overlap in the data sets, and difficulty comparing the two.

To remedy the discrepancy among data sets, more recordings made in the presence of baleen whales, with particular attention to low frequencies, are needed to classify whale calls by species. In the autonomous arrays, a tradeoff between bandwidth and experiment duration has been imposed both by data storage and battery life limitations. Because marine seismologists are interested in only the low-frequency components of submarine earthquakes, sampling has been limited to those frequencies. However, the instruments described here are being upgraded to 32 Gb of storage without an increase in instrument size, and this trend toward more capacity promises to continue. This will allow for either longer deployments (up to one year or more) or higher sampling frequencies (up to 1000 Hz). Given adequate funding, marine bioacousticians might develop experiments that are not ‘piggy-backed’ on naval and geophysical experiments. These experiments would ensure a sampling protocol designed to gather the most useful information on vocalizing cetaceans, and could permit hydrophone placement in areas that may not be seismically active but that are important for cetaceans. Arrays with hydrophones in closer proximity could target specific areas and would allow localization of calling whales. Deploying long-term moored hydrophones in regions of the ocean frequented by whales is a useful means of surveying the sounds that occur in that region and of directing future work.

ACKNOWLEDGMENTS

This work was supported in part by the NOAA VENTS program and the National Ocean Partnership Program through the Ocean Acoustic Observatory Federation. Dr. Haru Matsumoto designed the moored hydrophones, without which this work would not have been possible. Dr. David Mellinger, Sallie Beavers, and an anonymous reviewer greatly improved the final draft of this paper. PMEL contribution No. 2091.

- Aguayo, A., Bernal, R., Olavarria, C., Vallejos, V., and Hucke, R. (1998). “Observaciones de cetaceos realizadas entre Valparaiso e Isla de Pascua, Chile, durante los inviernos de 1993, 1994 y 1995,” *Revista de Biología Marina y Oceanografía* **33**, 101–123.
- Backus, R. H., and Schevill, W. E. (1966). “Physeter clicks” in *Whales, Dolphins and Porpoises*, edited by K. S. Norris (University of California Press, Berkeley, CA), pp. 510–527.
- Clark, C. W. (1990). “Acoustic behavior of mysticete whales,” in *Sensory Abilities of Cetaceans*, edited by J. Thomas and R. Kastelein (Plenum, New York), pp. 571–583.
- Clark, C. W., and Charif, R. A. (1998). “Acoustic monitoring of large whales to the west of Britain and Ireland using bottom-mounted hydrophone arrays, October 1996–September 1997,” Joint Nature Conservation Committee, Report No. 281.
- Clark, C. W., Charif, R. A., 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 Pt. Barrow, Alaska,” *Reports of the International Whaling Commission* **46**, 541–554.
- 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,” *Reports of the International Whaling Commission* **39**, 297–303.
- Clark, C. W., and Fristrup, K. (1997). “Whales ’95: A combined visual and acoustic survey of blue and fin whales off southern California,” *Reports of the International Whaling Commission* **47**, 583–600.
- Clarke, R. (1962). “Whale observation and whale marking off the coast of Chile in 1958 and from Ecuador towards and beyond the Galápagos Islands in 1959,” *Norsk Hvalfangst-Tidende* **51**(7), 265–287.
- Cummings, W. C., and Thompson, P. O. (1971). “Underwater sounds from the blue whale, *Balaenoptera musculus*,” *J. Acoust. Soc. Am.* **50**, 1193–1198.
- Cummings, W. C., Thompson, P. O., and Ha, S. J. (1986). “Sounds from Bryde, *Balaenoptera edeni*, and finback, *B. physalus*, whales in the Gulf of California,” *Fish. Bull.* **84**, 359–370.
- Dziak, R. P., Fox, C. G., Matsumoto, H., and Schreiner, A. E. (1997). “The April 1992 Cape Mendocino earthquake sequence: Seismo-acoustic analysis utilizing fixed hydrophone arrays,” *Mar. Geophys. Res.* **19**, 137–162.
- Edds, P. L. (1988). “Characteristics of finback, *Balaenoptera physalus*, vocalizations in the St. Lawrence Estuary,” *Bioacoustics* **1**, 131–149.
- Edds, P. L., Odell, D. K., and Tershy, B. R. (1993). “Vocalizations of a captive juvenile and free-ranging adult-calf pairs of Bryde’s whales, *Balaenoptera edeni*,” *Marine Mammal Science* **9**, 269–284.
- Fiedler, P. C., Philbrick, V., and Chavez, F. P. (1991). “Oceanic upwelling and productivity in the eastern tropical Pacific,” *Limnology and Oceanography* **36**, 834–850.
- Knowlton, A. R., Clark, C. W., and Kraus, S. D. (1991). “Sounds recorded in the presence of sei whales, *Balaenoptera borealis*,” *The Ninth Biennial Conference on the Biology of Marine Mammals*, Chicago, Illinois.
- Lockyer, C. (1984). “Review of baleen whale (*Mysticeti*) reproduction and implications for management,” *Reports of the International Whaling Commission Special Issue* **6**, 27–50.
- Ljungblad, D. K., Clark, C. W., and Shimada, H. (1998). “A comparison of sounds attributed to pygmy blue whales (*Balaenoptera musculus brevicauda*) recorded south of the Madagascar Plateau and those attributed to ‘true’ blue whales (*Balaenoptera musculus*) recorded off Antarctica,” *Reports of the International Whaling Commission* **48**, 439–442.
- Miyashita, T., Kato, H., and Kasuya, T. (1995). “Worldwide map of cetacean distribution based on Japanese sighting data,” *National Research Institute of Far Seas Fisheries, Shizuoka, Japan*.
- Moore, S. E., Stafford, K. M., Dahlheim, M. E., Fox, C. G., Braham, H. W., Polovina, J., and Bain, D. (1998). “Seasonal variation in reception of fin whale calls at five geographic areas in the north Pacific,” *Marine Mammal Sci.* **14**, 617–627.
- Nishimura, C. E., and Conlon, D. M. (1994). “IUSS dual use: Monitoring whales and earthquakes using SOSUS,” *Marine Tech. Soc. J.* **27**(4), 13–21.
- Perrin, W. F., Scott, M. D., Walker, G. J., and Cass, V. L. (1985). “Review of geographical stocks of tropical dolphins (*Stenella* spp. and *Delphinus delphis*) in the eastern tropical Pacific,” *NOAA Technical Report NMFS* **28**.
- Polacheck, T. (1987). “Relative abundance, distribution and inter-specific relationship of cetacean schools in the eastern tropical Pacific,” *Marine Mammal Sci.* **3**, 54–77.
- Ramirez, P. (1988). “Capturas de ballena sei (*Balaenoptera borealis*) frenta a Paita Peru,” edited by H. Salzwedel and A. Lauda. *Recursos y dinamica del ecosistema de afloramiento Peruano. Memorias del 2do Congreso Latinoamericano sobre Ciencias del Mar (COLAMAR)*, 17–21 Agosto de 1987, Lima, Peru. Tomo 1. Instituto del Mar de Peru. Boletín Volumen Extraordinario. Lima, Peru, pp. 341–343.
- Ramirez, P. (1989). “Captura de cetáceos mayores desde las estaciones costeras del Perú: 1951–1985,” *Boletín de Lima* **64**, 91–95.
- Reilly, S. B. (1990). “Seasonal changes in distribution and habitat differences among dolphins in the eastern tropical Pacific,” *Marine Ecology Progress Series* **66**, 1–11.
- Reilly, S. B., and Fiedler, P. C. (1994). “Interannual variability of dolphin habitats in the eastern tropical Pacific I: Research vessel surveys, 1986–1990,” *Fish. Bull.* **92**, 434–450.
- Reilly, S. B., and Thayer, V. G. (1990). “Blue whale (*Balaenoptera musculus*) distribution in the eastern tropical Pacific,” *Marine Mammal Science* **6**, 265–277.
- Richardson, W. J., Greene, Jr., C. R., Malme, C. I., and Thomson, D. H. (1995). *Marine Mammals and Noise* (Academic, New York).
- Rivers, J. A. (1997). “Blue whale, *Balaenoptera musculus*, vocalizations from the waters off central California,” *Marine Mammal Science* **13**, 186–195.
- Schevill, W. E., and Watkins, W. A. (1972). “Intense low-frequency sounds from an Antarctic minke whale, *Balaenoptera acutorostrata*,” *Breviora* **388**, 1–8.
- Stafford, K. M., and Fox, C. G. (1996). “Occurrence of blue and fin whale

- calls in the north Pacific as monitored by U.S. Navy SOSUS arrays," J. Acoust. Soc. Am. **100**, 2611(A).
- Stafford, K. M., and Fox, C. G. (1998). "A comparison of the acoustic signals produced by blue whales (*Balaenoptera musculus*) in the north Pacific and their use in stock definition/differentiation," World Marine Mammal Science Conference, Monaco.
- Stafford, K. M., Fox, C. G., and Clark, D. (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., Nieukirk, S. L., and Fox, C. G. (1999). "An acoustic link between blue whales in the eastern tropical Pacific and the northeast Pacific," Marine Mammal Science **15**, 1258–1268.
- Talandier, J., and Okal, E. A. (1996). "Monochromatic *T* waves from underwater volcanoes in the Pacific Ocean: Ringing witnesses to geyser processes?" Bull. Seismol. Soc. Am. **86**, 1529–1544.
- Thompson, P. O., Cummings, W. C., and Ha, S. J. (1986). "Sounds, source levels, and associated behavior of humpback whales, Southeast Alaska," J. Acoust. Soc. Am. **80**, 735–740.
- Thompson, P. O., Findley, L. T., and Vidal, O. (1992). "20 Hz pulses and other vocalizations of fin whales, *Balaenoptera physalus*, in the Gulf of California, Mexico," J. Acoust. Soc. Am. **92**, 3051–3057.
- Thompson, P., Findley, L., Vidal, O., and Cummings, W. (1996). "Underwater sounds of blue whales, *Balaenoptera musculus*, in the Gulf of California, Mexico," Marine Mammal Science **12**, 288–292.
- Thompson, P. O., and Friedl, W. A. (1982). "A long term study of low frequency sound from several species of whales off Oahu, Hawaii," Cetology **45**, 1–19.
- Thompson, T. J., Winn, H. E., and Perkins, P. J. (1979). "Mysticete sounds," in *Behavior of Marine Animals*, edited by H. E. Winn and B. L. Olla (Plenum, New York), Vol. 3, pp. 403–431.
- Urick, R. J. (1983). *Principles of Underwater Sound* (McGraw-Hill, New York).
- Volkov, A. F., and Moroz, I. F. (1977). "Oceanological conditions of the distribution of Cetacea in the eastern tropical part of the Pacific Ocean," Reports of the International Whaling Commission **27**, 186–188.
- Wade, L. S., and Friedrichsen, G. L. (1979). "Recent sightings of the blue whale, *Balaenoptera musculus*, in the northeastern tropical Pacific," Fish. Bull. **76**, 915–919.
- Wade, P. R., and Gerrodette, T. (1993). "Estimates of cetacean abundance and distribution in the eastern tropical Pacific," Reports of the International Whaling Commission **43**, 477–493.
- Watkins, W. A. (1980). "Acoustics and the behavior of sperm whales," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 283–290.
- Watkins, W. A. (1981). "Activities and underwater sounds of fin whales," Scientific Reports of the Whales Research Institute **33**, 83–117.
- 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.
- Whitehead, H., and Arnbom, T. (1987). "Social organization of sperm whales off the Galapagos Islands, February–April, 1985," Canadian J. Zoology **65**, 913–919.
- Whitehead, H., Weilgart, L., and Waters, S. (1989). "Seasonality of sperm whales off the Galapagos Islands, Ecuador," Reports of the International Whaling Commission **39**, 207–210.
- Winn, H. E., and Perkins, P. L. (1976). "Distribution and sound of the minke whale, with a review of mysticete sounds," Cetology **19**, 1–12.
- Wyrтки, K. (1964). "Upwelling in the Costa Rica Dome," Fish. Bull. **63**, 355–372.
- Wyrтки, K. (1966). "Oceanography of the eastern equatorial Pacific Ocean," Oceanography and Marine Biology Annual Review **4**, 33–68.

Transmission beam pattern and echolocation signals of a harbor porpoise (*Phocoena phocoena*)

Whitlow W. L. Au

Hawaii Institute of Marine Biology, University of Hawaii, P.O. Box 1106, Kailua, Hawaii 96734

Ronald A. Kastelein, Tineke Rippe, and Nicole M. Schooneman

Harderwijk Marine Mammal Park, Strandboulevard Oost 1, 3841 AB Harderwijk, The Netherlands

(Received 21 November 1998; revised 27 July 1999; accepted 4 August 1999)

The transmission beam pattern of an echolocating harbor porpoise (*Phocoena phocoena*) was measured in both the vertical and horizontal planes. An array of seven Brüel and Kjaer 8103 hydrophones connected to an amplifier-line driver module was used to measure the beam patterns. The porpoise was trained to station in a hoop and echolocate a cylindrical target located at a range between 7 and 9 m while the array was located 2 m in front of the hoop. The 3-dB beamwidth in both the vertical and horizontal planes was the same at approximately 16 degrees and the beam was pointed toward the forward direction. The individual hydrophones in both the vertical and horizontal arrays measured signal waveforms that were similar throughout the 40-degree span of the array. The porpoise emitted signals with intervals that were 20 to 35 ms longer than the round trip travel time between the animal and the target. The average source level, peak frequency, and bandwidth were 157 dB, 128 kHz, and 16 kHz, respectively. © 1999 Acoustical Society of America. [S0001-4966(99)01812-3]

PACS numbers: 43.80.Ka, 43.80.Lb [FD]

INTRODUCTION

The harbor porpoise (*Phocoena phocoena*) is one of the species of odontocetes that are subject to a high annual rate of bycatches in gillnets. Annually, many hundreds of harbor porpoises are entangled in bottom set gillnets or sink-gillnets along the Northwest coastal regions of North America and along the northern coastal regions on both sides of the Atlantic ocean (Jefferson and Curry, 1994; Read, 1994; Perrin *et al.* 1994). Gilbert and Wynne (1988) identified the bycatch of harbor porpoises in bottom gillnets as the greatest potential conflict between marine mammals and New England gillnet fisheries. Because of the seriousness of the bycatch problem, there is a pressing need to better understand the biology, ecology, physiology, acoustics, and other facets of harbor porpoises so that we may better manage and conserve this species.

The echolocation signals emitted by odontocetes tend to fall into two broad categories. Dolphins that typically emit whistle signals also emit brief broadband echolocation signals having between four and eight cycles and duration of 40–70 μ s (Au, 1993). Most odontocetes fall into this class. Odontocetes that do not emit whistle signals, emit narrow-band echolocation signals having at a minimum of about 12 cycles with duration generally greater than 100 μ s (Au, 1993). Among odontocetes that emit signals in this category are the harbor porpoise (Møhl and Andersen, 1973; Kamminga and Wiersma, 1981; Hatakeyama and Soeda, 1990; Goodson *et al.*, 1995), finless porpoise, *Neophocaena phocaenoides* (Kamminga, 1988); Commerson's dolphin, *Cephalorhynchus commersonii* (Kamminga and Wiersma, 1982; Evans *et al.*, 1988); Hector's dolphin, *Cephalorhynchus hectori* (Dawson, 1988); Dall's porpoise, *Phocoenoides*

dalli (Awbrey *et al.*, 1979; Hatakeyama and Soeda, 1990); and the pygmy sperm whale, *Kogia sp* (Carder *et al.*, 1995). Whether riverine dolphins whistle or not is still an open question. Riverine dolphins emit echolocation signals that would fall in the short duration, broadband category.

The transmission beam pattern of six odontocetes species has been measured: the rough-tooth porpoise, *Steno bredanensis* (Norris and Evans, 1967), Black Sea bottlenose dolphin, *Tursiops sp.* (Zaslavkiy, 1971 in Bel'kovich and Dubrovskiy, 1976), common dolphin, *Delphinus delphis* (Korolev *et al.*, 1973; in Bel'kovich and Dubrovskiy, 1976), Atlantic bottlenose dolphin, *Tursiops truncatus* (Evans *et al.*, 1973; Au, 1980; Au *et al.*, 1978, 1986), beluga whale, *Delphinapterus leucas* (Au *et al.*, 1987), and false killer whale, *Pseudorca crassidens* (Au *et al.*, 1995). However, only in the studies conducted with the Atlantic bottlenose dolphin, beluga whale, and false killer whale were the echolocation beam patterns measured in both the vertical and horizontal planes. Au *et al.* (1986, 1987) found that the beam of the Atlantic bottlenose dolphin and the beluga whale were directed upwards at an angle of 5 to 10 degrees above the horizontal plane. However, the beam axis for the false killer whale was between 0 to -5 degrees below the horizontal plane (Au *et al.*, 1995). The beam patterns in both planes had similar shapes with the 3-dB beamwidth for *Tursiops* being approximately 10 and 6.5 degree for *Delphinapterus*. The beamwidth for *Pseudorca* was 12 degrees in the vertical plane and 7.5 degrees in the horizontal plane, for the high-frequency signals used by the animal. The beam pattern of a nonwhistling dolphin or porpoise such as *Phocoena phocoena* that project relatively narrow-band echolocation signals has never been reported.

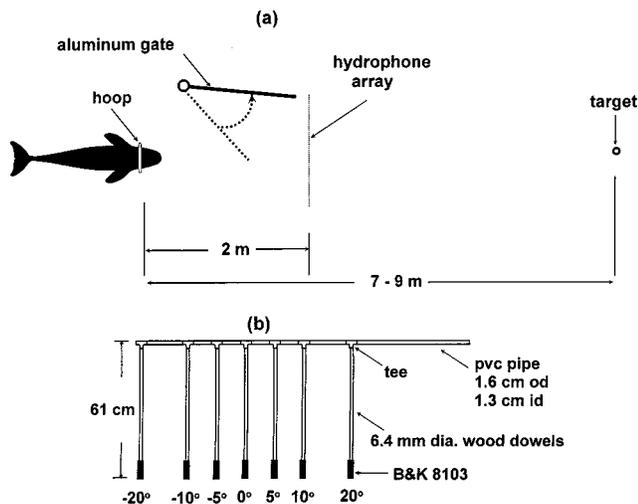


FIG. 1. (a) Geometry of the target detection experiment; (b) schematic of the hydrophone array in the horizontal plane.

I. EXPERIMENTAL CONFIGURATION AND APPROACH

The subject was a stranded male harbor porpoise (code PpSH030) named Marco, who was rehabilitated at the Netherlands Cetacean Research and Rehabilitation Center at the Harderwijk Marine Mammal Park (see Nachtigall *et al.*, 1995; Read *et al.*, 1997) from an approximate age of 5 weeks. At the time of the study, the animal was 3 years old, weighing 30.9 kg and measuring 129.5 cm in length. During this experiment the animal was housed at the Dolphin Rehabilitation and Research Center at Neeltje Jans, The Netherlands, in a floating net pen 20 m in width and 34 m in length with a depth of 3.5 m at low tide and 5 m at high tide (see Kastelein *et al.*, 1999). A small enclosure (2×4 m×1.2 m depth) on one end of the floating structure served as the experimental area. A trial began with the animal swimming into a hoop station of 45 cm diameter. When an aluminum door directly in front of the porpoise was swung open, as depicted in Fig. 1(a), the porpoise began echolocating. The porpoise's task was to report on the presence or absence of a hollow copper cylinder, 2.1 cm in diameter and 15 cm in length located between 7 and 9 m from the hoop. A go/no-go response paradigm was used in which the porpoise backed out of the hoop, turned around, and swam to and touched a float located at the rear of the experimental area to indicate a target-present response. For a target-absent response the animal remained in the hoop until the trainer signaled by blowing a dog-whistle. Eventually the cylindrical target was switched to a waterfilled stainless steel sphere for determining the porpoise target detection range (Kastelein *et al.*, 1999). However, during the beam pattern measurements, the cylindrical target was used.

An array of seven Brüel and Kjaer 8103 hydrophones located 2 m from the hoop station was used to measure the porpoise transmission beam pattern in the vertical and horizontal planes. The hydrophones were attached to 6.4-mm-diam. wooden dowels, 61 cm in length, which were plugged into one of the arms of a polyvinyl chloride (pvc) tee as shown in Fig. 1(b). Appropriate lengths of hollow 1.6-cm-diam. pvc pipes were used to place the hydrophones at

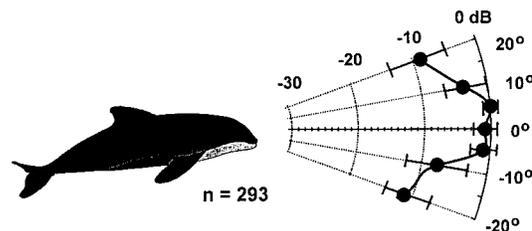


FIG. 2. Beam pattern in the vertical plane. The radial axis represents the relative intensity in dB measured by the various hydrophones. The beam pattern has been normalized so that the angle at which the maximum averaged signal was measured was given a value of 0 dB.

angles of 0, ± 5 , ± 10 , and ± 20 degrees with respect to an imaginary line projected from the center of the hoop and perpendicular to the plane of the hoop. The array orientation shown in Fig. 1(b) was used to measure the transmission beam in the horizontal plane. The array was turned 90 degrees in order to measure the vertical transmission beam.

The hydrophones were connected to an eight-channel variable gain amplifier-line driver module which had a bandwidth of over 1 MHz. Each hydrophone was calibrated before and after the study with a B&K 4223 calibrator to determine the relative sensitivity between hydrophones. The hydrophone signals were recorded with a seven-channel Rascal Store-7 instrumentation recorder operated at tape speed of 76.2 cm/s (30 in./s). At this tape speed the 3-dB bandwidth of the tape recorder was 170 kHz. A 1-V peak-to-peak simulated *tursiops* echolocation signal was recorded on each channel at the beginning of each tape for amplitude calibration purposes. The tape recorder along with the multi-channel amplifier-line driver were housed in a small cabin directly behind the trainer who sat at the edge of the experimental area (see Kastelein *et al.*, 1999).

The tape recorder signals were digitized at a tape speed of 9.52 cm/s (3.745 in./s) using an RC-Electronics multiple-channel 12-bit analog-to-digital (A/D) converter board housed in a personal computer that operated at a sample rate of 125 kHz. Therefore, the effective real-time digitizing rate was 1 MHz per channel. The multiplexer in the A/D unit switched from one channel to another in 1 μ s which represented an effective real-time switching time of 0.125 μ s. The digitized data were stored on the hard disk of a personal computer.

II. EXPERIMENTAL RESULTS

A. Beam pattern results

The results of the beam pattern measurements in the vertical plane are shown in Fig. 2. The beam pattern is plotted in polar coordinates with the relative intensity in dB depicted by the radial axis and the angle in degrees along the circumferential axis. For each signal, the relative intensity in dB for each hydrophone output relative to the hydrophone with the highest output was determined. The mean and standard deviation of the relative intensities in dB at each angle were then calculated for all the appropriate signals. Finally, the statistical results were normalized to the angle with the highest relative mean value. Since the porpoise was free to move his head in the hoop, only signals which produced the

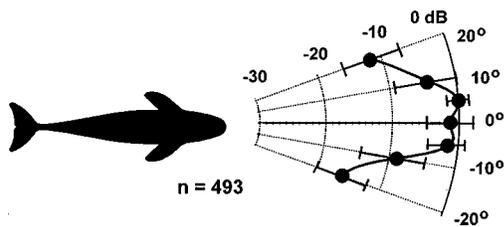


FIG. 3. Beam pattern in the horizontal plane. The radial axis represents the relative intensity in dB measured by the various hydrophones. The beam pattern has been normalized so that the angle at which the maximum averaged signal was measured was given a value of 0 dB.

highest intensity in the hydrophones located at either 0 or ± 5 degrees were used to calculate the beam pattern. A total of 293 signals were used to calculate the beam pattern shown in Fig. 2.

The major axis of the beam in the vertical plane is slightly elevated by about 5 degrees, although the 5-degree spacing of the hydrophones between ± 10 degrees did not allow for a finer determination of the elevation angle. The 293 values used to calculate the mean and standard deviation at both 0 and -5 degrees were arranged in two columns and the difference between the corresponding values of each row was used in a two-tailed *t*-test. The results at 0 and -5 degrees were so similar that we felt it necessary to test if there were statistically different. The results indicated that the mean relative intensities at 0 and -5 degrees were not significantly different. The 3-dB beam width was approximately 16.5 degrees in the vertical plane.

The results of the beam pattern measurement in the horizontal plane are shown in Fig. 3. The same procedure used in deriving the vertical beam pattern was also used to obtain the horizontal beam pattern. The beam in the horizontal plane is directed essentially in the forward direction although the major axis pointed towards $+5$ degrees. In reality, the major axis may be somewhere between 0 and 5 degrees (the 5-degree spacing of the hydrophones between ± 10 degrees did not allow for a finer estimate). As in the vertical beam pattern case, the results at 0 and -5 degrees were so similar that we performed a two-tail *t*-test to determine if they were significantly different. We found that the results at 0 and -5 degrees were again not significantly different. As with the vertical beam pattern determination, only signals that produced the highest intensities between 0 and ± 5 degrees were

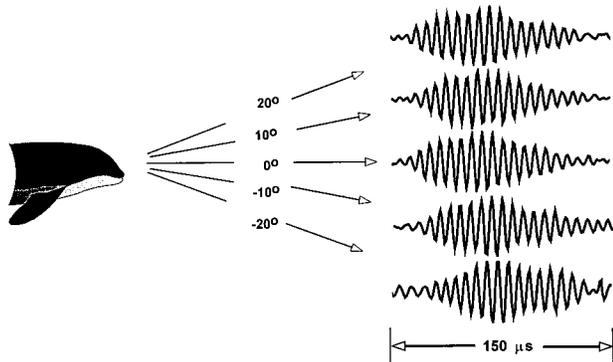


FIG. 4. The signal waveform measured by five of the seven hydrophones in the vertical plane for a single echolocation signal.

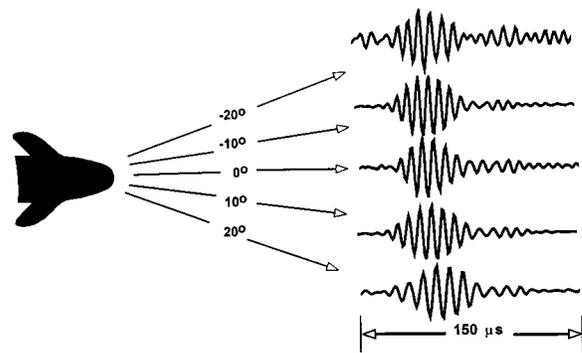


FIG. 5. The signal waveform measured by five of the seven hydrophones in the horizontal plane for a single echolocation signal.

used to compute the horizontal beam pattern. A total of 497 echolocation signals were used to determine the beam pattern of Fig. 3. The 3-dB beamwidth of the horizontal beam was essentially the same as for the vertical beam, approximately 16.5 degrees.

Examples of a single signal measured by five of the hydrophones in the vertical and horizontal planes are shown in Figs. 4 and 5, respectively. These examples indicate that the wave form of the emitted signal remains essentially the same for angles up to at least ± 20 degrees. The signal in Fig. 5 was chosen to illustrate a common variation in signal waveform involving differences in signal duration and amplitude envelope.

B. Signal characteristics

The mean and standard deviation of the interclick intervals for ten consecutive trials are shown in Fig. 6 when the target was approximately 7.5 m from the hoop station. The interclick intervals were always 20 to 35 ms longer than the two-way transit time for a signal to travel from the animal to the target and back. The interclick interval was greater than the two-way transit time for both target present and target absent trials. The interclick intervals associated with the absent trials suggest that the animal had an expectation of the target range from previous trials and conducted its echoloca-

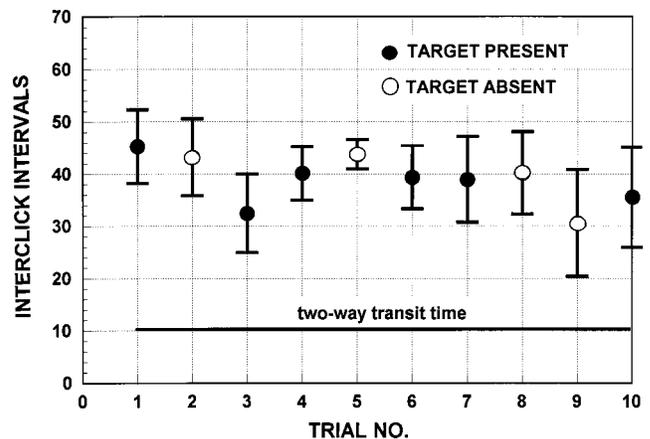


FIG. 6. Interclick interval for ten consecutive trials. The closed circles represent the target-present trials and the open circles represent the target-absent trials.

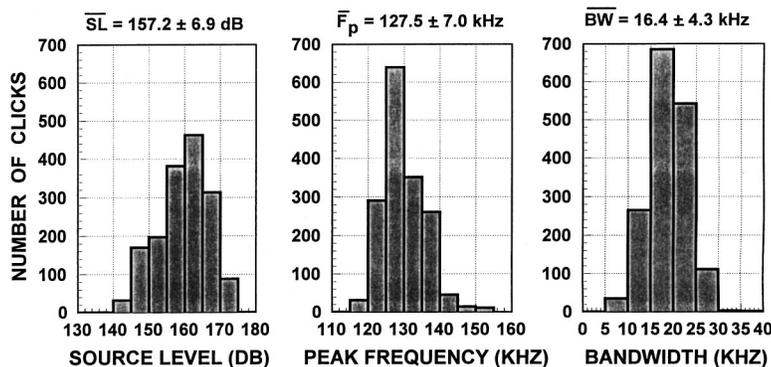


FIG. 7. Histograms of peak-to-peak source level in dB *re* 1 μ Pa, peak frequency in kHz, and 3-dB bandwidth in kHz. The average values are denoted above each histogram.

tion accordingly. The lag times between the interclick intervals and two-way transit time are similar to those observed for echolocating dolphins (Au, 1993).

Histograms of the source level (peak-to-peak sound pressure level referenced to 1 m), the peak frequency, and 3-dB bandwidth are depicted in Fig. 7. The mean and standard deviation of the values comprising each histogram are shown above each histogram. The signals from the hydrophone with the highest level were used to construct the histograms. The source level histogram indicates that source levels between 160–165 dB were used most often, with the average source level being 157.2 ± 6.9 dB *re* 1 μ Pa. The highest peak-to-peak source level recorded was 172 dB *re* 1 μ Pa, although most of the signals (71%) were emitted with source levels between 155 and 170 dB.

The peak-frequency histogram indicates that the porpoise emitted signals with peak frequencies between 125–130 kHz most often. The average peak frequency was 127.5 ± 7.0 kHz. Most of the signals had 3-dB bandwidths between 15–25 kHz with an average of 16.4 ± 4.3 kHz. The Q of a signal is defined as the peak or center frequency divided by the bandwidth. Using the average values for the peak frequency and 3-dB bandwidth will result in a Q of 7.8. The signals used by whistling dolphins have Q's typically between 3–4 (Au, 1993).

III. DISCUSSION AND CONCLUSIONS

It is difficult to access the relative accuracy of the beam shape in Figs. 2 and 3 for angles between ± 10 degrees in both planes because the animal was not required to station on a bite plate as in the other beam pattern measurements (Au, 1993; Au *et al.*, 1995). The front portion of the beams in both planes is relatively flat since the values at 0 and -5 degrees are not significantly different and the values between 0 and 5 degrees are similar. The porpoise could move relatively freely in the hoop and often seemed to direct its beam towards the hydrophones at the extreme ends of the array rather than on the target directly in front of it. Nevertheless, given the inherent limitations in the positional accuracy of the porpoise from trial to trial, and movements of the animal within the hoop, the beam pattern results indicate that the transmission beam is essentially pointed in the forward direction. The slight indentation in both beams at 0 degrees was not caused by improper accounting of the gain and sensitivity of each channel. The reference signal recorded on the

front portion of each tape ensured that the recorder channels were properly calibrated. The calibration of the hydrophones before and after the study ensured that the hydrophones' sensitivity did not change.

The signal waveforms measured simultaneously by all the hydrophones in both the vertical and horizontal arrays indicate that the waveform is similar for angles between ± 20 degrees in both planes. In contrast, signal waveforms measured at different angles in both the vertical and horizontal planes for *Tursiops truncatus* (Au, 1980, 1993), *Delphinapterus leucas* (Au *et al.*, 1987; Au, 1993), and *Pseudorca crassidens* (Au *et al.*, 1995) are very dissimilar. The signals of *Tursiops truncatus*, *Delphinapterus leucas*, and *Pseudorca crassidens* measured at angles of ± 10 degrees and greater from the beam axis are extremely distorted in comparison to signals measured along the beam axis. Similar distortion to the signals of *Phocoena* were not observed because of the relatively narrow bandwidth of the signals. For a planar transducer, the width of the beam will change with frequency becoming narrower as frequency increases. Broadband signals emitted by dolphins like *Tursiops* cover a wide enough frequency range that the beam pattern at the low end of the frequency band of the signal is significantly different from the beam pattern at the high end. Nonwhistling odontocetes like *Phocoena* emit relatively narrow bandwidth signals and the beam patterns hardly change across the signals' frequency bandwidth.

Phocoena phocoena has a broader transmission beam than other odontocetes such as *Tursiops truncatus*, *Delphinapterus leucas* (Au, 1993), and *Pseudorca crassidens* (Au *et al.*, 1995). The directivity of a beam is directly proportional to the size of a transducer; being greater for a larger radiator and vice versa (Urlick, 1983). Conversely, the beamwidth is inversely proportion to the size of a transducer (Urlick, 1983). Since the *Phocoena* has a smaller head than other odontocetes, it is not surprising for this species to have a broader transmission beam.

Directivity index is a measure of the directional property of a transmission or reception beam. From Urlick (1983), the directivity index is given by the expression

$$DI = 10 \log \frac{4\pi}{\int_0^{2\pi} \int_{-\pi/2}^{\pi/2} [P(\theta, \varphi)/P_0]^2 \sin \theta d\theta d\varphi}, \quad (1)$$

where θ is the angle in the vertical plane, φ is the angle in the

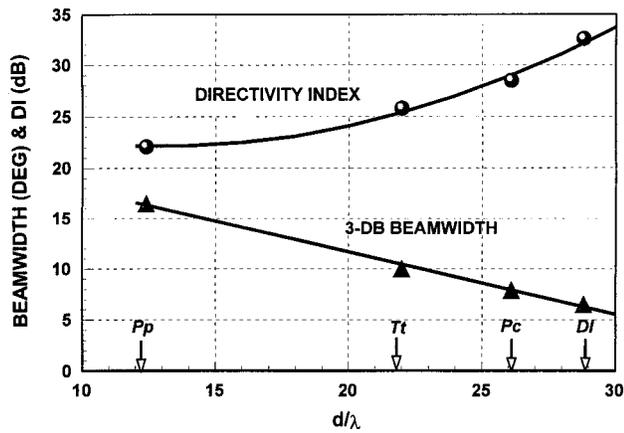


FIG. 8. Transmission directivity index and 3-dB beamwidth for four odontocetes. The directivity index and beamwidth for *Tursiops truncatus* and *Delphinapterus leucas* came from Au (1993) and for *Pseudorca crassidens* from Au *et al.* (1995). The wavelength λ corresponds to the average peak frequency of the animals' echolocation signal. The directivity index is fitted with a second-order polynomial curve and the beamwidth is fitted with a linear curve.

horizontal plane, and $[p(\theta, \varphi)/p_0]^2$ is the three-dimensional beam pattern of a radiator.

The beam patterns of Figs. 2 and 3 were inserted into Eq. (1) and the integral evaluated numerically using Simpson's $\frac{1}{3}$ rule for double integrals. The directivity index and 3-dB bandwidth of four odontocete species are plotted in Fig. 8 as a function the head diameter (measured at the blowhole) over the wavelength corresponding to the peak frequency. The directivity index for *Tursiops* and *Delphinapterus* was calculated by Au (1993) and for *Pseudorca* by Au *et al.* (1995). The horizontal and vertical beamwidth were essentially the same for *Phocoena*, *Tursiops* (Au, 1993), and *Delphinapterus* (Au, 1993). However, the same cannot be said for *Pseudorca* (Au *et al.*, 1995) since the horizontal beam is narrower than the vertical beam. The beamwidth for *Pseudorca* used in Fig. 8 was an averaged of the mean horizontal and vertical beamwidth. Some of the pertinent values associated with the physical dimension of the animals and with their echolocation signals used to generating Fig. 8 are given the Table I.

The directivity index was fitted with a second-order polynomial and there was a 99.8% correlation between the fitted curve and the directivity index values. The equation for the directivity index as a function of d/λ is given by

$$DI = 28.89 - 1.04 \left(\frac{d}{\lambda} \right) + 0.04 \left(\frac{d}{\lambda} \right)^2. \quad (2)$$

The 3-dB beamwidth was fitted with a linear curve with a 98.5% correlation between the fitted curve and the data points. The equation for the beamwidth as a function of d/λ is given by

$$BW = 23.9 - 0.6 \left(\frac{d}{\lambda} \right). \quad (3)$$

From Eqs. (2) and (3), an estimate of the directivity index and 3-dB beamwidth of other odontocetes can be obtained if the head size of the animal and the peak frequency of the echolocation signals are known.

The characteristics of Marco's signals are similar to those reported in previous studies (Kamminga, 1988; 1994; Hatakayama and Soeda, 1990; Goodson *et al.*, 1995). However, one of the parameters not discussed previously for the harbor porpoise is the interclick interval. The interclick intervals in this study were always greater than the two-way transit time. Therefore, an echo was received before the next pulse was transmitted when a target was present. Even in the target-absent trials, the porpoise seemed to have anticipated the arrival time of echoes and therefore controlled his interclick interval appropriately. The typical lag time of 20–35 ms between the reception of an echo and the production of the next signal are similar to those of *Tursiops* (Au, 1993).

The source levels measured in this study are about 20 dB less than signals used by *Tursiops* in tanks (Au, 1993) and about 50–65 dB less than signals used by *Tursiops* in open waters (Au, 1993). The much lower source levels of *Phocoena* is probably related to the much smaller size of the animal. Adult *Tursiops truncatus* weigh about 180–230 kg compared to about 40–60 kg for *Phocoena phocoena*. Unlike *Tursiops*, the *Phocoena* used in this study had source levels in the open waters of Neeltje Jans that were similar to those reported by Goodson *et al.* (1995) in a pool at for same animal in a pool at Harderwijk Marine Mammal Park. Signals of *Tursiops* and other dolphins measured in tanks tend to have source levels of about 170–180 dB whereas in open waters the source levels can increase to about 210–225 dB (Au, 1993). Two possible reasons for the use of lower level signals in tanks may be associated with the animals not wanting to receive high-amplitude reflections from tank walls or that the ranges involved in tanks are relatively small

TABLE I. Pertinent values used to generate the information given in Fig. 8. The measurements were obtained from records taken during medical examinations. The information pertaining to the beam pattern were obtained from Au (1993) for *Tursiops* and *Delphinapterus*, and Au *et al.* (1995) for *Pseudorca*. The values used for *Pseudorca* are for the type IV signals (Au *et al.*, 1995).

	<i>Phocoena phocoena</i>	<i>Tursiops truncatus</i>	<i>Pseudorca crassidens</i>	<i>Delphinapterus leucas</i>
Circumference (cm)	46.5	89.7	120.0	125.0
Diameter (cm)	14.8	28.6	38.2	39.8
Peak frequency (kHz)	127.5	117	104	110
Directivity index (dB)	22.1	25.8	28.5	32.6
d/λ	12.4	22.0	26.1	28.8
Vertical beamwidth (degrees)	16.5	10.2	9.7	6.5
Horizontal beamwidth (degrees)	16.5	9.7	6.2	6.5

so that high-amplitude signals are not necessary. It seems that the first reason is probably closer to the truth since even at distances that are comparable to tank dimensions, dolphins have been observed to emit peak-to-peak source levels in excess of 200 dB in open waters (Au, 1993) while no one has ever reported source levels greater than 190 dB for dolphins in tanks. The reason for the similarity in the amplitude of Marco's echolocation signals measured in a pool and in open waters may be related to the inherent low levels of the animal's echolocation signals. Reflection of *Phocoena* signals from a pool wall will not be at a very high level.

Approximately 76% of the signals had a peak frequency between 125 and 140 kHz, with an average of 127.5 kHz. Goodson *et al.* (1995) reported a much higher peak frequency of 145 kHz for the same animal. However, in their study, statistical information of peak frequency was not given so it is difficult to ascertain the range of peak frequencies that were emitted by Marco. Furthermore, the measurements reported here were conducted several years after the measurements of Goodson *et al.* (1995). There may be a possibility that juvenile *Phocoena* may emit higher frequency signals than adults. Møhl and Andersen (1973) and Kamminga (1994) reported peak frequencies of about 140–145 kHz for juvenile harbor porpoises which were higher than the 120–150 kHz reported for an adult by Kamminga and Wiersma (1981). However, these studies did not present statistical information and the peak frequencies reported by the investigators were considered "typical" of the signals measured.

Approximately 75% of Marco's echolocation signals had 3-dB bandwidth between 15 and 25 kHz with an average value of 16.4 ± 4.3 kHz. These values are slightly higher than the 12.8 kHz reported by Goodson *et al.* (1995) for a single *Phocoena* signal. The bandwidths measured in this study are much smaller than the 35–60 kHz for *Tursiops*. The bandwidth of an oscillatory signal of relatively constant period will be inversely proportional to the length of the signal. Since the duration of the harbor porpoise signal is two to three times as long as for the bottlenose dolphin, the bandwidth is narrower by a factor of 2 to 3.

The results of this study should contribute to a deeper understanding of the acoustics of harbor porpoises and that could be helpful in the search for solutions to the bycatch problem. The width of the echolocation beam is one of the most fundamental property of an echolocation system and knowledge of the beam pattern can be helpful in understanding how *Phocoena phocoena* uses its echolocation in the wild. Designers of acoustic transponders that are activated by echolocation signals can be assured that signals radiating from a porpoise head in different directions should have relatively similar shape. The width of the porpoise beam can also be used to calculate the distance passive acoustic reflectors and active acoustic transponders should be spaced to be effective at specific ranges. Finally, knowledge of typical average source levels is also important in designing active transponders and passive reflectors.

ACKNOWLEDGMENTS

The authors thank Michiel Schotten for digitizing the tape recorded signals. We also thank Dr. Paul Nachtigall and Jeffrey Pawloski of the Hawaii Institute of Marine Biology, University of Hawaii, for their advice during the planning phase of this study. Funding for this project was obtained from Harderwijk Marine Mammal Park, The North Sea Directorate and The Netherlands National Institute for Coastal and Marine Management (Contract No. 22971983), the Department of Nature Management of the Netherlands Ministry of Agriculture, Nature Management and Fisheries (Contract Nos. 9785-1997-2204-11110039 and 3003009274). We also thank Dr. Harold Hawkins of the Office of Naval Research in the United States for his support (Grant No. N00014-95-1-0462). This is Hawaii Institute of Biology Contribution No. 1050.

- Au, W. W. L. (1980). "Echolocation Signals of the Atlantic Bottlenose Dolphin (*Tursiops truncatus*) in Open Waters," in *Animal Sonar Systems*, edited by R. G. Busnel and J. F. Fish (Plenum, New York), pp. 251–282.
- Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer-Verlag, New York).
- Au, W. W. L., Floyd, R. W., and Haun, J. E. (1978). "Propagation of Atlantic bottlenose dolphin echolocation signals," *J. Acoust. Soc. Am.* **64**, 411–422.
- Au, W. W. L., Moore, P. W. B., and Pawloski, D. (1986). "Echolocation transmitting beam of the Atlantic bottlenose dolphin," *J. Acoust. Soc. Am.* **80**, 688–691.
- Au, W. W. L., Penner, R. H., and Turl, C. W. (1987). "Propagation of Beluga echolocation echolocation signals," *J. Acoust. Soc. Am.* **82**, 807–813.
- Au, W. W. L., Pawloski, J. L., Nachtigall, P. E., Blonz, M., and Gisiner, R. C. (1995). "Echolocation signals and transmission beam pattern of a false killer whale (*Pseudorca crassidens*)," *J. Acoust. Soc. Am.* **98**, 51–59.
- Awbrey, F. T., Norris, J. C., Hubbard, A. B., and Evans, W. E. (1979). "The Bioacoustics of the Dall's Porpoise-Salmon Drift Net Interaction," H/SWRI Technical Report, pp. 79–120.
- Bel'kovich, V. M., and Dubrovskiy, N. A. (1976). *Sensory Bases of Cetacean Orientation* (Izdatel'stvo Nauka, Moscow).
- Carder, D., Ridgway, S., Whitaker, B., and Geraci, J. (1995). "Hearing and Echolocation in a pygmy sperm whale *Kogia*," Eleventh Biennial Conference on the Biol. of Mar. Mamm. 14–18 December, Orlando, FL(A).
- Dawson, S. M. (1988). "The High Frequency Sounds of Free-Ranging Hector's Dolphin, *Cephalorhynchus hectori*," *Rep. Int. Whal. Commission* (Special Issue 9), 339–341.
- Evans, W. E. (1973). "Echolocation by marine delphinids and one species of fresh-water dolphin," *J. Acoust. Soc. Am.* **54**, 193–199.
- Evans, W. E., Awbrey, F. T., and Hackbarth, H. (1988). "High Frequency Pulse Produced by Free Ranging Commerson's Dolphin (*Cephalorhynchus commersonii*) Compared to those of Phocoenids," *Rep. Int. Whal. Commission* (Special issue 9), 173–181.
- Gilbert, J. R., and Wynne, K. M. (1988). "Marine mammal interactions with New England gillnet fisheries," Final report to the National Marine Fisheries Service, Northeast Fisheries Center, Woods Hole, MA. Contract No. NA-84-EA-00070.
- Goodson, A. D., Kastelein, R. A., and Sturtivant, C. R. (1995). "Source levels and echolocation signal characteristics of juvenile harbour porpoises (*Phocoena phocoena*)," in *Harbour Porpoise: Laboratory Studies to Reduce Bycatch*, edited by P. E. Nachtigall, J. Lien, W. W. L. Au, and A. J. Read (De Spil, Woerden, The Netherlands), pp. 41–53.
- Hatakeyama, Y., and Soeda, H. (1990). "Studies on Echolocation of Porpoises Taken in Salmon Gillnet Fisheries," in *Sensory Abilities of Cetaceans*, edited by J. A. Thomas and R. A. Kastelein (Plenum, New York), pp. 269–281.
- Jefferson, T. A., and Curry, B. E. (1994). "A global review of porpoise (Cetacea: Phocoenidae) mortality in gillnets," *Bio. Conserv.* **67**, 167–183.
- Kamminga, C. (1988). "Echolocation Signal Types of Odontocetes," in *Animal Sonar: Processes and Performance*, edited by P. E. Nachtigall and P. W. B. Moore (Plenum, New York), pp. 9–22.

- Kamminga, C. (1994). "Research on Dolphin Sounds," Doctoral Thesis, Technical University Delft, The Netherlands.
- Kamminga, C., and Wiersma, H. (1981). "Investigations on Cetacean Sonar II. Acoustical Similarities and Differences in Odontocete Sonar Signals," *Aquatic Mammal*, **8**, 41–62.
- Kamminga, C., and Wiersma, H. (1982). "Investigations on Cetacean Sonar V. The True Nature of the Sonar Sound of *Cephalorhynchus commersonii*," *Aquatic Mammal*, **9**, 95–104.
- Kastelein, R. A., Au, W. W. L., Rippe, H. T., and Schooneman, N. M. (1999). "Target detection by an echolocating harbor porpoise (*Phocoena phocoena*)," *J. Acoust. Soc. Am.* **105**, 2493–2498.
- Korolev, L. D., Lipatov, N. V., Rezvov, R. N., Savel'yev, M. A., and Flenov, A. B. (1973). "Investigation of Delphinid Locating System Capabilities in Passive Ranging," in *Ref. Dokl. 8-y Vses. Akust. Konf., Moscow*, Vol. 1, pp. 125–126.
- Møhl, B., and Andersen, S. (1973). "Echolocation: High-frequency component in the click of the harbor porpoise (*Phocoena ph. L.*)," *J. Acoust. Soc. Am.* **54**, 1368–1372.
- Nachtigall, P. E., Lien, J., Au, W. W. L., and Read, A. J. (Eds.) (1995). *Harbour Porpoises: Laboratory Studies to Reduce Bycatch* (De Spiel, Woerden, The Netherlands).
- Norris, K. S., and Evans, W. E. (1967). "Directionality of Echolocation Clicks in the Rough-Tooth Porpoise, *Steno bredanensis* (Lesson)," in *Marine BioAcoustics*, edited by W. N. Tavolga (Pergamon, New York), Vol. 2, pp. 305–316.
- Perrin, W. F., Donovan, G. P., and Barlow, J. (Eds.) (1994). *Gillnets and Cetaceans* (International Whaling Commission, Cambridge, UK).
- Read, A. J. (1994). "Interaction between cetaceans and gillnet and trap fisheries in the Northwest Atlantic," in *Gillnets and Cetaceans*, edited by W. F. Perrin, P. Donovan, and J. Barlow (International Whaling Commission, Cambridge, UK), pp. 133–147.
- Read, A. J., Wiepkema, P. R., and Nachtigall, P. E. (Eds.) (1997). *The Biology of the Harbour Porpoise* (De Spiel, Woerden, The Netherlands).
- Urick, R. J. (1983). *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York).
- Zaslavskiy, G. L. (1971). "Directionality of Sound Emission in the Black Sea Bottlenose Dolphin," in *Tr. Akust. Inst., Moscow*, Vol. 17, pp. 60–70.

Modeling vibration and sound production in insects with nonresonant stridulatory organs

Gunther Tschuch^{a)}

Martin-Luther-Universität, Institut für Zoologie, Entwicklungsbiologie, Domplatz 4, D-06099 Halle, Germany

Denis J. Brothers^{b)}

University of Natal Pietermaritzburg, School of Botany and Zoology, Private Bag X01, Scottsville, 3209 South Africa

(Received 4 May 1999; revised 23 August 1999; accepted 3 September 1999)

Using a simple model for a nonresonant stridulatory organ, the intensities and shapes of the vibrations produced and the radiated sound were calculated. To test the results, the stridulation of femal mutillid wasps (Mutillidae, Hymenoptera, Insecta) was investigated by the help of sound-analytical methods and with laser vibrometry. The model was shown to give good estimations for the sound spectrum and for the absolute power of the second to fifth harmonics. The wasps investigated, females of *Dasylabris maura* and “*Smicromyrme*” *praedatrix* with body lengths of about 10 mm, are able to make audible sounds with a power level of about 40 dB at a distance of 0.1 m. Stridulatory organs like those in mutillid wasps occur in several species of insects (e.g., beetles, wasps, and ants) and in other arthropods. © 1999 Acoustical Society of America. [S0001-4966(99)05812-9]

PACS numbers: 43.80.Ka [WA]

INTRODUCTION

Stridulation is defined as sound emission by friction of two specialized surfaces against one another.¹ Stridulatory organs are widespread in insects and other arthropods. In most cases there is a scraper (plectrum, strigilator) on one part moving over a more or less regular structure (stridulatory file, stridulitrum, pars stridens, strigil) with periodically recurring elevations on another part. The structure and size of the stridulatory file are extremely variable; it may be made up of bristles, tubercles, teeth, or, frequently, of very broad ripples. Stridulatory organs have been found on virtually all regions of the body. According to the review of Dumortier,¹ there are many file-plectrum combinations in insects, virtually anywhere where two parts of the exoskeleton may touch each other.

Resonant structures may be associated with or attached to the stridulatory organ, especially if the sound is optimized for long-range intraspecific communication at distances of about 10 to 1000 m.² The benefits of using resonant structures are: limited bandwidth of the signal, optical sound radiation, and high energy produced with low energy per stroked bristle, tubercle, or tooth. Therefore, the stridulatory file is narrow if the resonator has a high quality Q . The width-to-period ratio of the stridulatory file is usually below 10:1.^{1,3} Resonant structures are, for example, specialized areas of the forewings as in crickets^{4,5} or air-filled cavities in the abdomen.¹

For short-range intra- and interspecific communication, no resonant structures are necessary. Sometimes radiation far from the optimal sound impedance match is better, for ex-

ample, to avoid the attraction of predators or to avoid the jamming of communication inside insect colonies. Stridulation without resonance is frequently used in distress situations, when the sound is presumably an aposematic signal aimed at discouraging predators from attacking the toxic, armed, or otherwise unpalatable sender. Some insects present acoustical signals together with optical (aposematic coloration) and/or chemical signals (allomones). A stridulatory file without associated resonant structures looks like a washboard and the width to ripple-period ratio of the file is above 10:1, usually in the order of 100:1. This paper deals with such stridulatory organs with broad files. They occur in different genera of many beetle families, like the cranioprothoracic organs, e.g., in Chrysomelidae, metathoracic organs in Cerambycidae, and abdominoelytral organs, e.g., in Scarabaeidae.¹ Abdominal organs occur in Hymenoptera in some genera of the wasp families Bradynobaenidae, Rhopalosomatidae, and Sphecidae, in almost all genera of Mutillidae,⁶ and in several ant species.⁷

I. MATERIAL AND METHODS

A. Experimental animals

Measurements of stridulating female mutillid wasps were used to test the results obtained from computations based on the theoretical model (see below). In mutillids, the stridulatory organ occurs in both sexes on the abdomen dorsally between the second and third metasomal segments. The measurements were performed with three females of *Dasylabris maura*⁸ (subfamily Sphaerophthalminae) from Germany and three females of “*Smicromyrme*” *praedatrix*⁹ (subfamily Mutillinae) from South Africa. The mutillid wasps were kept in a terrarium where they had free access to water and a 50% solution of honey in water.

^{a)}Electronic mail: tschuch@zoologie.uni-halle.de

^{b)}Electronic mail: brothers@zoology.unp.ac.za

B. Anatomical investigations

Dead dry specimens were rehydrated and softened in a pepsin solution (24 h, 40 °C). The expanding took place in distilled water. During washing in water and acetone (15' each, 40 °C) the animals were treated with ultrasound to remove dust particles from the cuticle surface. After air-drying, the specimens were coated with gold (Balzers SCD 004 sputter coater) and examined in a Hitachi S-2400 scanning electron microscope (SEM). For all measurements the SEM was calibrated by using a Planotec silicon test object (Plano).

C. Sound and vibration recordings

Sound signals were recorded digitally with a computer sound board Pro Audio Spectrum 16 (Media Vision, settings: sampling rate 44.1 kHz; resolution 8 bit mono), and analyzed by the software Avisoft Sonograph Pro 2.7 (Specht). A 1-in. condenser microphone MK 102 was mounted at a distance of about 0.1 m. The sound levels were measured with a precision pulse levelmeter 00 017 (Messelektronik "Otto Schön") which also served as a microphone amplifier.

The vibrations were measured with a laser vibrometer instrument containing a sensor head OFV300 (Polytec) and the vibrometer controller OFV2100 (Polytec, with enlarged frequency range, settings: displacement 1.62 $\mu\text{m}/\text{V}$, velocity 5 mm/s/V, 50-kHz low pass filter). Single small (diameter approximately 20 μm) half-mirrored spheres made from Scotchlite 7610 foil (3M) served as reflectors for the sensor head at a distance of 0.6 m. Displacement and velocity were digitized parallel to the sound signal with the help of a digital storage oscilloscope DS-8601A (Iwatsu, settings: sampling rates 75 and 300 kHz). The digitized signals containing 12 000 words each were accessed through a RS232-C interface SX-0111/0114 (Iwatsu) and converted using a PASCAL program to WAV format computer files.

II. RESULTS

A. Anatomy

In both species, the stridulatory file of the female is a shield-shaped area of approximately 160 very regular ripples with ripple distances of 2.4 to 4.8 μm . In the middle part of the file the ripple distance is 3.3 μm in *D. maura* (Fig. 1) and 2.6 μm in "*S.*" *praedatrix*, with depths of approximately 2 μm . The file is approximately 0.5 mm \times 0.5 mm in size, and the plectrum width is 0.20 mm in *D. maura* and about 0.15 mm in "*S.*" *praedatrix*.

B. Model

Figure 1 shows the movement of the plectrum relative to the file. If the file moves outward the plectrum goes from position 1 to position 2 in a short time down into the valley between the two ripples. The time taken to reach a position as in 1 on the next ripple (position 3) is much longer, because of the oblique angle of the plectrum. The up-and-down displacement of the plectrum is therefore not symmetrical and can be described using a sawtooth function over time as a first approximation. If the direction of movement is reversed, the sawtooth function is also reversed [Fig. 2(A) and

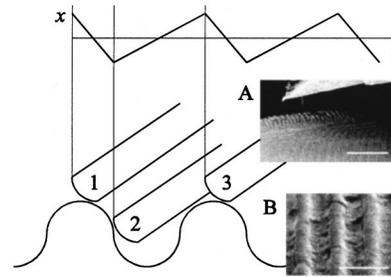


FIG. 1. The model for plectrum vibrations (right anterior, left posterior of the body). 1, 2, and 3 are the positions of the plectrum (above) relative to the stridulatory file while stroking one ripple during the movement of the file outwards. Above the plectrum, the displacement is plotted in simplified form. The time between 1 and 2 is t_b , and the period between 1 and 3 is T . Insets SEM pictures of *Dasylabris maura*: (A) Cross section of the plectrum (white bar: 100 μm). The plectrum is not in contact with the file below (artifact of preparation). (B) Three ripples of the file (white bar: 5 μm).

(B)]. The differentiation of x with respect to t gives the velocity u of the vibration. It is a periodic step function.

C. Vibration

Real displacement-time and velocity-time plots measured by laser vibrometry (Fig. 3) show waveforms very similar to the model. But the width of the high-speed part (between 1 and 2 in Fig. 1) is much shorter than in the model (Fig. 2). This can be explained by faster slipping of the plectrum into (or out of) the spaces between ripples during outward (or inward) movement of the file. This is possible because the plectrum is flexible and the forces causing inward and outward movements of the file with respect to the plectrum are weak compared with the forces causing the pressure of the file against the plectrum. During inward movement the duration t_b of the highest velocity remains relatively constant at a mean value of 0.16 ms. It is relatively independent of the observed period T , which ranges from below 1 ms to more than 2 ms in both species. Whereas T can easily be modified by the stridulating animal, the time t_b depends more on the mechanical properties of the stridulatory organ and to a lesser extent on the forces applied. The mean velocity u_0 usually reaches about 10 mm/s (Fig. 3). At that velocity the plectrum moves 1.6 μm in 0.16 ms, which is in the range of the ripple height of 2 μm observed anatomically and used in the model in Fig. 1.

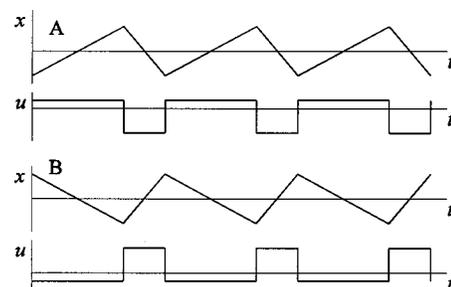


FIG. 2. Displacement x and velocity u of the plectrum during movement of the stridulatory file outward (A) and inward (B).

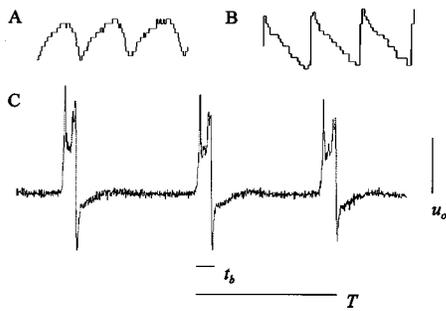


FIG. 3. Displacement x (A, B) and velocity u (C) of the vibrating plectrum of a female of “*S.*” *praedatrix* as a function of time. The plots show three ripples of the stridulatory file stroked by the plectrum during the movement of the third metasomal segment outward (A) and inward (B, C). See text.

D. Sound spectrum

Because of the small dimensions and the low frequencies produced by these insects, there is no optimal power transfer from the vibrating source to the surrounding air.¹⁰ This causes an attenuation of the lower frequencies. The radiated sound power P may be expressed as the product of the squared effective velocity u_{eff} of the oscillating area and the real part of the acoustical radiation impedance, the resistance w_r

$$P = u_{\text{eff}}^2 w_r. \quad (1)$$

The calculation of u_{eff} is easily possible if $u(t)$ is expanded into a Fourier series. A shift of the step function along the time axis allows the sine terms to be ignored. Because of the absence of any influence of constant terms on the sound radiation it is possible to write

$$u(t) = \sum_{n=1}^{\infty} u_n(t) = \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{2\pi n}{T} t\right) \right]. \quad (2)$$

If u is equal to u_0 between $-t_b/2$ and $t_b/2$ (the duration of highest velocity) and 0 in the other ranges in the period from $-T/2$ to $T/2$, the a_n can be calculated as follows

$$a_n = u_0 \frac{2}{T} \int_0^{t_b/2} \cos\left(\frac{2\pi n}{T} t\right) dt, \quad (3)$$

$$a_n = \frac{2u_0}{\pi n} \sin\left(\frac{\pi n}{T} t_b\right). \quad (4)$$

According to Eq. (2) the expression for $u_n(t)$ is given by

$$u_n(t) = \frac{2u_0}{\pi n} \sin\left(\frac{\pi n}{T} t_b\right) \cos\left(\frac{2\pi n}{T} t\right). \quad (5)$$

For the calculation of the power P it is necessary to have the square of the effective velocity. This is the integral for all times over the square of $u(t)$. Because of the periodicity it is sufficient to integrate over one period T . The mean value of the squared cosine function is 0.5 and the term in the right-hand brackets does not depend on time. Therefore the result of the integral is

$$u_n^2_{\text{eff}} = \frac{2u_0^2}{\pi^2 n^2} \sin^2\left(\frac{\pi n}{T} t_b\right). \quad (6)$$

It is more complicated to obtain a good estimate of the acoustical radiation resistance w_r . The laser vibrometry shows that the main area of vibration comprises about 0.03 mm² posteriorly on the plectrum. Furthermore, the females stridulate with the abdomen less than 1 mm above the ground. Therefore, as a first approximation, the far-field radiation may be that of an oscillating piston in a baffle,¹¹

$$w_r = c \varrho \left[1 - \frac{J_1(2kR)}{kR} \right] A, \quad (7)$$

where c and ϱ are the speed of sound and the density of air, A and R are area and radius of the piston, k is the wave number

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{c} = \frac{2\pi n}{cT}, \quad (8)$$

and J_1 is the Bessel function of the first kind and of order one. It can be expressed in a series expansion in the following manner:

$$J_1(x) = \frac{x}{2} - \frac{x^3}{16} + \frac{x^5}{384} - \dots \quad (9)$$

For frequencies up to 50 kHz the wave number k is less than 1 mm⁻¹ and R is about 0.1 mm. Thus $kR \ll 1$ is always true. Therefore Eq. (9) can be reduced to the first two terms of the series. Under these conditions Eq. (7) can be written as follows:

$$w_r \approx \frac{c \varrho k^2 A^2}{2\pi}, \quad (10)$$

or, according to Eq. (8):

$$w_r \approx 2\pi \frac{\varrho n^2}{cT^2} A^2. \quad (11)$$

Inserting Eq. (6) and Eq. (11) in Eq. (1) yields a sound power of line n in the line spectrum as follows:

$$P_n \approx \frac{4}{\pi} \frac{\varrho u_0^2 A^2}{cT^2} \sin^2\left(\frac{\pi n}{T} t_b\right) \quad (12)$$

or

$$P_n \approx \frac{2}{\pi} \frac{\varrho u_0^2 A^2}{cT^2} \left[1 - \cos\left(\frac{2\pi n}{T} t_b\right) \right]. \quad (13)$$

Because the expression in front of the square brackets depends only on the observed period T , the entire power spectrum depends only on T and t_b , the duration of the highest velocity. Figure 4 shows the variation of the logarithmic spectrum with T , assuming a constant t_b of 0.16 ms. Whereas T influences the line density in the spectrum and the power of all lines together, t_b causes attenuations only around $1/t_b$ (6 kHz) and integer multiples of that.

To compare a real sound pulse with an artificially made one and to estimate the maximal sound power, it is necessary to measure the period T . Because the fundamental frequency is of low power, cepstrum analysis was used to get the fundamental frequency f exactly. Figure 5 shows the minimal frequency usually found in the middle part of each sound

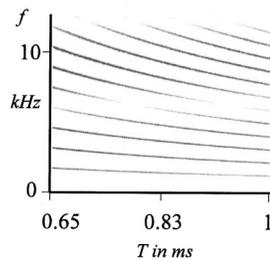


FIG. 4. Variation of the calculated acoustical spectrum with period T of the vibration. The period T is varied from 0.65 to 1 ms whereas the width t_b remains constant at 0.16 ms. The logarithmic sound power P is coded as a gray scale; the frequency scale is linear.

pulse. With the minimal value for $T=0.65$ ms and with $u_0 = 0.01$ m/s and $A=0.03$ mm², Eq. (13) yields

$$P_n \approx 0.2nW \left[1 - \cos\left(\frac{2\pi n}{T} t_b\right) \right]. \quad (14)$$

Thus the first maximum of P can be found near $n=2$ (3 kHz) with a value of 0.4 nW. If T goes up to 1 ms the first maximum will be reached at $n=3$ (3 kHz) with only 0.2 nW because the factor in front of the square brackets goes down to 0.1 nW. The sound will be radiated in a hemisphere above the ground. At a distance of 0.1 m (approximately the wavelength of 3 kHz) the hemisphere has a surface area of 0.06 m². Thus the maximal sound intensity at this distance is approximately 7 nW/m², yielding a sound power level of 38 dB relative to 1 pW/m². The two species investigated produced sound power levels between 30 and 40 dB. Thus the estimation of sound intensity made above is quite good.

If the ripple period is constant, the time course of the fundamental frequency f represents the velocity of the plectrum relative to the stridulatory file. In a real sound pulse there are slow parts at the beginning and at the end (Fig. 6 left). The fastest movement occurs in the middle part. Using a self-written PASCAL program, it was possible to construct a synthetic sound pulse according to the model used (Fig. 6 right). Some characteristics are the same in both sonograms. Typically, but not in all actual sound recordings, it is evident that in the middle of the sound pulse the $4f$ harmonic, and at the beginning and the end of the pulse the $5f$ harmonic, are attenuated near 6 kHz. This is because the frequency of the attenuation is relatively constant and depends mainly on the mechanical properties of the stridulatory organ. The most conspicuous differences in the real sound pulse compared with the synthetic one relate to the greater attenuation of the

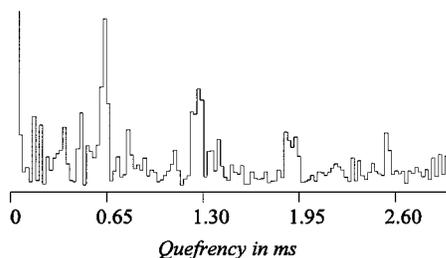


FIG. 5. Cepstrum of the middle part of one sound pulse of a female *D. maura*. The quefrequency of 0.65 ms corresponds to the period T . The fundamental frequency is the reciprocal quefrequency: $1/0.65$ ms = 1.54 kHz.

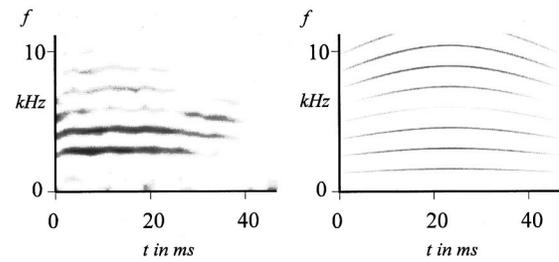


FIG. 6. Sonogram of an actual pulse produced by a stridulating female of *D. maura* (left) and sonogram computed according to the model (right). The frequency scale is linear. For further explanation see text.

frequencies at f and above $4f$. The inaccurate power calculation at the fundamental frequency results from ignoring the reversed velocity course following the step function (Fig. 3). Further, because of limited acceleration of the plectrum, the time function of the velocity is not really a step function. Thus the model is inaccurate at higher harmonics. On the other hand, the model shows good agreement with the frequency of maximal sound intensity, and it allows an explanation of the broad frequency spectrum of the real sound ranging from approximately 1 to 10 kHz.

E. Oscillogram

The distress call contains several syllables, each made up of two sound pulses generated during the movement of the abdominal segment inward and outward (Fig. 7). The two pulses are usually readily distinguishable in the time function because of their asymmetry relative to the time axis (Fig. 8). Although the time function of the radiated sound pressure has a very different shape compared with the vibration velocity (Fig. 3), the asymmetry caused by the highly damped plectrum vibration is evident here also. The findings were confirmed both by the model and by actual recordings.

Whereas the syllable period and the pulse width remains fairly constant in the calls of female “S.” *praedatrix*, female *D. maura* are able to switch to faster syllable repetition with shorter sound pulses depending on the stress intensity (Fig. 7). Because the velocity of movement between plectrum and file does not change in the faster syllables, the fundamental frequency remains approximately the same. Thus the plectrum strokes a lesser number of file ripples in this case, but the produced spectra and time functions of the sound do not differ in the faster and slower parts.

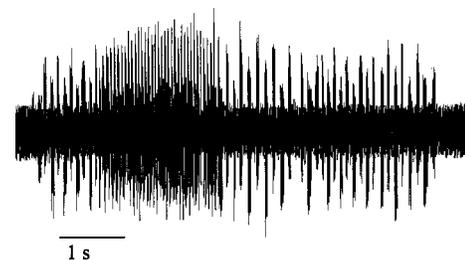


FIG. 7. Oscillogram of the distress call of a female of *D. maura*. Note the discontinuous changes in the syllable period. One syllable consists of two sound pulses corresponding to the movement of the stridulatory file inward and outward.

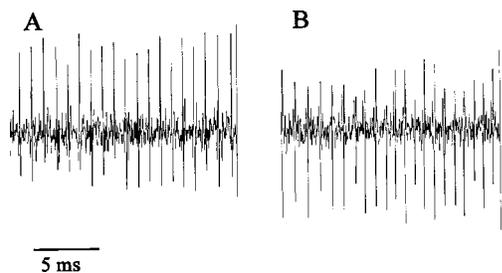


FIG. 8. Magnified detail of Fig. 7. Each ripple of the stridulatory file produces one spike in the sound pressure. The spikes are up and down corresponding to the direction of movement in the stridulatory organ (A) inward, (B) outward.

III. DISCUSSION AND CONCLUSIONS

The signals investigated in this paper have a broad spectrum with many harmonics. It may be that aposematic acoustical signals should have a broad frequency spectrum, for example, because the potential predators (e.g., lizards, mammals) are phylogenetically very different, or because the signals may have evolved in a Müllerian mimicry to hissing by vertebrates. But there is an unsolved problem. According to the model explored here, the sound power level of the stridulating insect depends on A^2/T^2 , in addition to u_0^2 . Assuming that a scaling down of the body size will influence A (the square of the length) more than T and u_0 , the sound power level would be very low in small insects. While the investigated species have a body length of about 10 mm, there are many mutillid wasps and other insects with similar stridulatory organs only half the length or less. Half radius yields quarter radiating area and a reduced specific acoustic resistance, thus the sound power level will be reduced by 12 dB or more. The resulting sound, less than 26 dB, would scarcely be perceptible to many of the potential predators. But the sound power level is also not high enough to be used for intraspecific communication over a distance. This may provide a biophysical explanation of why some small insects may use the vibration signals for short-range intraspecific communication; some findings of Markl¹² suggest this. But there are many other possibilities for the use of stridulatory organs. In leaf-cutting ants the vibrations serve to reduce energy consumption during cutting,¹³ and in mutillid wasps vibrations may aid soil manipulation.¹⁴ However, there is still no good explanation why small insects also have stridulatory organs.

The model in this paper is a first approach to simulation of spectra and time functions of vibration and radiated sound. It has been shown to give good agreement with data taken from the investigated specimens of mutillid wasps and with the data of Spangler and Manley.¹⁵ The model will probably be useful for other arthropods with nonresonant stridulatory organs too, especially where they have similar file structures and approximately the same values of ripple-strike period T .^{16,17} Furthermore, the results allow the synthesis of artificial vibrations and sounds. Therefore, the findings could be helpful for future investigations dealing with the effects of such signals in inter- or intraspecific communication.

ACKNOWLEDGMENTS

The work was financially supported by the Deutsche Forschungsgemeinschaft (DFG) Grant Nos. Ts53/1-1 and Ts53/1-2 (to GT) and by the University of Natal Research Committee (to DJB). The authors wish to thank N. Elsner (Göttingen) for the use of the laser vibrometer in his laboratory. Furthermore we thank the two anonymous reviewers for their valuable advice.

¹B. Dumortier, "Sound emission apparatus in arthropoda," in *Acoustic Behavior of Animals*, edited by R.-G. Busnel (Elsevier, Amsterdam, 1963), Chap. 11, pp. 277–345.

²H. C. Bennet-Clark, "Insect sound production: transduction mechanisms and impedance matching," in *Biological Fluid Dynamics*, edited by C. P. Ellington and T. J. Pedley (Company of Biologists, Cambridge, 1995), pp. 199–218.

³K.-G. Heller, *Bioakustik der Europäischen Laubheuschrecken* (Margraf, Weikersheim, 1988).

⁴H. C. Bennet-Clark, *J. Exp. Biol.* **52**, 619–652 (1970).

⁵H. Nocke, *Z. Vergleichende Physiol.* **74**, 272–314 (1971).

⁶D. J. Brothers, *Univ. Kans. Sci. Bull.* **50**, 483–648 (1975).

⁷B. Hölldobler and E. O. Wilson, *The Ants* (Springer, Berlin, 1990), Chap. 7, pp. 255–257.

⁸C. Linnaeus, *Caroli Linnaei Systema Naturae* T.1 (Salvins, Holmiae, 1758).

⁹F. Smith, *Description of new Species of Hymenoptera in the Collection of the British Museum* (British Museum, London, 1879).

¹⁰H. C. Bennet-Clark, *Nature (London)* **234**, 255–259 (1971).

¹¹P. M. Morse, *Vibration and Sound* (McGraw-Hill, New York, 1948) [reprinted: Acoustical Society of America, Woodbury, 1981].

¹²H. Markl, *Z. Vergleichende Physiol.* **57**, 299–330 (1967).

¹³J. Tautz, F. Roces, and B. Hölldobler, *Science* **267**, 84–87 (1995).

¹⁴H. G. Spangler, *J. Kansas Entomological Soc.* **46**, 157–160 (1973).

¹⁵H. G. Spangler and D. G. Manley, *Ann. Entomol. Soc. Am.* **71**, 389–392 (1978).

¹⁶W. M. Masters, *J. Comp. Physiol. A* **135**, 259–268 (1980).

¹⁷W. M. Masters, J. Tautz, N. H. Fletcher, and H. Markl, *J. Comp. Physiol. A* **150**, 239–249 (1983).

Receiver operating characteristics and temporal integration in an insect auditory receptor cell

Jakob Tougaard^{a)}

Department of Zoophysiology, Center for Sound Communication, University of Aarhus, DK-8000 Aarhus C, Denmark

(Received 23 March 1999; revised 1 September 1999; accepted 7 September 1999)

Receiver operating characteristics (ROC) analysis was applied to the problem of temporal integration in the A1 auditory receptor of the noctuid moth *Noctua pronuba*. The index of detectability (d') was calculated either from ROC curves or from a procedure equivalent to a two-alternatives forced-choice paradigm of psychophysics. Stimuli of six different durations (1–40 ms) were used. The detectability increased exponentially with intensity for all signals and with the rate of increase independent of duration. This allowed for an evaluation of the influence of stimulus duration on detectability (duration/intensity trade-off). A model of statistical threshold improvement (the “multiple looks” model) could be applied and the influence of statistics on thresholds could be separated from the influence of intensity integration. For stimulus durations significantly shorter than the time constant of the receptor (3.4 ms for *N. pronuba*), intensity integration was the main factor determining thresholds and thus an energy detector model was applicable to the data. For durations significantly longer than the time constant, the multiple look model alone could account for changes in thresholds with duration. For durations in-between, a mixture of both models was applicable, making specific predictions difficult. © 1999 Acoustical Society of America. [S0001-4966(99)06112-3]

PACS numbers: 43.80.Lb, 43.64.Tk [WA]

INTRODUCTION

The threshold is a fundamental concept of sensory physiology, defined as the smallest stimulus which gives rise to a response. Thresholds are, however, due to their stochastic nature, difficult to measure unambiguously. In a physiological preparation it is generally always possible to identify a range of stimulus intensities where the receptor responds only to a fraction of the stimulus presentations, even in the absence of spontaneous activity in the receptor. Since it is often desired to express the sensitivity of the sensory system in question by a single figure, a threshold criterion must be defined. In an electrophysiological experiment this criterion is selected by the experimenter. The exact nature of the threshold criterion will vary, depending to a large degree on the specific type of experiment and stimuli used. Sometimes the choice is based on a well-developed theory, but equally often it is dictated primarily from practical and parsimonious considerations.

Since a wide range of threshold criteria are possible, the thresholds derived in a physiological experiment will depend on the particular choice made by the experimenter. This may or may not be a problem. At least two factors are to be considered when choosing a criterion. The first is comparability—the more similar the criteria of different experiments, the stronger the conclusions possible in a comparison. The second factor is the possibility that certain criteria affect the threshold measurements in a biased way because the criterion itself favors detection of certain stimuli

over others. An example of this is the determination of neural thresholds for stimuli of different durations. If thresholds are determined on the basis of a spike count criterion, they are likely to decrease with increasing stimulus duration. This is a simple consequence of the longer period of depolarization of the cell, which gives more time for spikes to accumulate and thus an increased probability that the threshold criterion will be exceeded. This is a purely statistical effect and is independent of the sensitivity of the receptor in question.

Receiver operating characteristics (ROC) analysis, derived from the theory of signal detection (Green and Swets, 1966), provides a valuable supplement to traditional threshold measurements. Whereas ROC analysis has been used widely in vertebrate psychophysics, the applications to electrophysiology have however, been limited, centered on vertebrate vision (e.g., FitzHugh, 1957; Barlow *et al.*, 1971; Cohn *et al.*, 1975; Lee *et al.*, 1993) and physiology of the 8th nerve (e.g., Relkin and Pelli, 1987; Viemeister *et al.*, 1992; and Fay and Coombs, 1992). The power of ROC-based measurements compared to more traditional methods lies in the ability to separate the *sensitivity* of the receptor from the *threshold criterion* used. The sensitivity of the sensory system, as defined by Green and Swets (1966), then becomes a measure of how well a given signal can be discriminated from background noise.

The ear of noctuid moths (Lepidoptera: Noctuidae) provides an excellent preparation for the study of sound reception at the receptor cell level. The noctuid ear is located on the animal's last thoracic segment and consists of a thin cuticular membrane (tympanum), backed by an air-filled cavity. Attached to the membrane is the receptor organ, which

^{a)}Present address: Center for Sound Communication, Institute of Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark; electronic mail: jakob.t@biology.ou.dk

contains only two receptor cells (*A1* and *A2*, see Ghiradella, 1971 for details). The two receptor cells are broadly tuned to ultrasonic frequencies. Their main function is to detect echolocating bats and thus help the moth evade predation (see, e.g., Miller and Surlykke, in press). The two cells have similar tuning. The only significant difference between the cells is that the *A1* cells is about 20 dB more sensitive than the *A2* (Roeder, 1974; Surlykke and Miller, 1982). This implies that at low stimulus levels, below the threshold of the *A2* cell, the activity measured in the auditory nerve is almost exclusively due to the *A1* cell.

It is well known that the threshold of the *A1* receptor decreases for increasing stimulus durations. This has previously been ascribed solely to a temporal integration of stimulus intensities (energy detection, Adams, 1971; Surlykke *et al.*, 1988; Tougaard, 1996). The presence of a statistical component in the analysis was demonstrated by Tougaard (1998) and the aim of the present work is to investigate the nature of this duration-dependent statistical threshold improvement.

By assessing the detectability for stimuli of different durations, rather than determining thresholds by a traditional spike count criterion, it is possible to evaluate the influence of stimulus duration on detection probabilities directly. It then becomes possible to investigate the relative importance of the statistical threshold improvement over the temporal integration of stimulus intensities (energy detection). This provides important information as to whether threshold changes with stimulus duration can correctly be modeled by an energy detector model or alternatively by a statistical model, if it is intended that the model reflect the mechanism responsible for the observed threshold changes.

I. METHODS

The preparation, stimulus generation, and data collection have been described in detail elsewhere (Tougaard, 1998). Briefly, recordings of spike activity were obtained from the auditory nerve of the noctuid moth, *Noctua pronuba* (Lepidoptera: Noctuidae) using a glass suction electrode. The *A1* receptor is a primary sensory cell and no synapses are thus located between the receptor cell body and the recording site. The nerve signal was fed into a computer, where a 12-bit A/D-converter recorded 102 ms of receptor activity, beginning 51 ms before signal onset.

Stimuli were 23 kHz pure tone signals of different durations (1–40 ms), generated by a Wavetek 112 generator, controlled from the computer. The stimulus frequency corresponds roughly to the frequency of best hearing of this species of moth.

The experiments were conducted as a series of sessions. Each session consisted of 900 stimulus presentations with a signal of constant duration, distributed with 50 presentations at each of 18 different intensities. Intensities were distributed in steps of 1 dB around an initial threshold estimate.

ROC analysis requires large data sets to achieve the desired accuracy in d' calculations. A sufficiently large data set was obtained from one individual. This individual showed an unusual stability of responses over a 7-hour period (no significant changes in thresholds, average number of spikes

elicited per stimulation, or spontaneous activity) and thus provided recordings of responses to around 30 000 single stimulus presentations. Such a data set poses unique opportunities for ROC analysis. By collecting the entire data set from a single individual, interindividual variance is eliminated and only the variance within the receptor cell itself is left. This is the only variance assumed to be relevant to the moth and also the variance of interest with respect to indices of detectability. Strictly speaking, the conclusions drawn in the following can only be said to be valid for the particular receptor cell investigated. As a check for consistency with other receptor cells, a partial analysis was performed on a pooled set of recordings from individuals of a second noctuid, *Spodoptera littoralis*. These data support the notion that the present results can be considered representative for noctuid *A1* cells in general.

A. Evaluation of data—receiver operating characteristics

The generation of receiver operating characteristics (ROC) curves follows the method described in detail by Cohn *et al.* (1975), who analyzed single-cell data from the frog optic nerve.

For each stimulus presentation, two observation intervals were chosen. One began 50 ms prior to stimulus onset and represents the signal-absent, or noise-alone, condition. This is termed the *N*-interval. The other interval began at stimulus onset (arrival time at the tympanal membrane), lasted 50 ms and represents the signal present condition (signal+noise), and is termed the *SN*-interval. For each recording (i.e., one stimulus presentation), the number of spikes in the two observation intervals, *N* and *SN*, was counted. Pooling the data from a large number of identical stimulations, two spike count histograms could be constructed, one for signal absent and one for signal present (Fig. 1).

The two observation intervals from the recordings were classified on the basis of a spike count criterion, *c*. Six different values of the criterion were used (1 through 6 spikes). Depending on the number of spikes in each interval, the type of interval (*N* or *SN*), and the criterion, *c*, the recordings were divided into four groups. *SN*-intervals, in which *c* spikes or more were found, were grouped as hits, and those that contained fewer than *c* spikes as misses. In the same way, *N*-intervals were grouped into correct rejections (fewer than *c* spikes) and false alarms (*c* spikes or more). The terms hit, miss, false alarm, and correct rejection are technical descriptions (Green and Swets, 1966), used here solely in order to retain a consistent terminology and are not intended to imply that any kind of decision was made by the receptor.

The hit rate associated with a particular criterion *c* is the ratio of hits to the total number of *SN*-intervals, and likewise is the false-alarm rate of the ratio of false alarms to the total number of *N*-intervals. Corresponding hit and false-alarm rates are plotted in a ROC plot (Fig. 2).

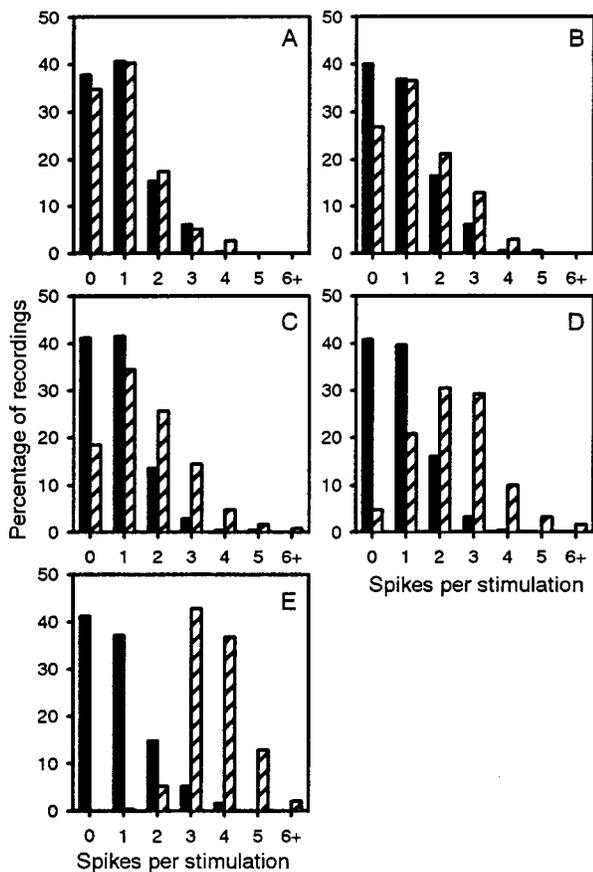


FIG. 1. Spike count distributions in 50-ms noise only (N) interval (solid bars) and 50-ms signal+noise (SN) interval (hatched bars). Stimulus duration 5 ms, five different sound-pressure levels: 28 (a), 32 (b), 34 (c), 36 (d), and 40 dB (e). Each histogram is based on pooled data from five sessions, equal to 250 stimulus presentations.

B. Index of detectability

For each signal duration and intensity, d' was estimated by a maximum likelihood fit to the experimental ROC data. This estimate is termed d'_{ROC} .

A second estimate of d' was obtained by evaluating the data with a procedure equivalent to the two-alternatives forced-choice (2AFC) paradigm of psychophysics (Relkin and Pelli, 1987). For each recording, the computer algorithm was “forced” to choose which of the two intervals (N or SN) was the most likely to contain a response to the signal, ignoring the fact that the signal always occurred in the last interval. This choice was based on a comparison of the spike counts in the two intervals. An observation, following a single stimulus presentation, was classified as a correct response if more spikes were counted in the SN -interval than in the N -interval and classified as wrong if the opposite was the case. In situations where there was an equal number of spikes in the two intervals, the observation was classified as correct in 50% of the cases. The proportion of correct classifications was thus calculated as

$$\text{Percent correct} = \frac{m_{SN>N} + m_{SN=N}/2}{m_{\text{total}}} \times 100, \quad (1)$$

where $m_{SN>N}$ is the number of observations where more spikes were found in the signal present interval, $m_{SN=N}$ is

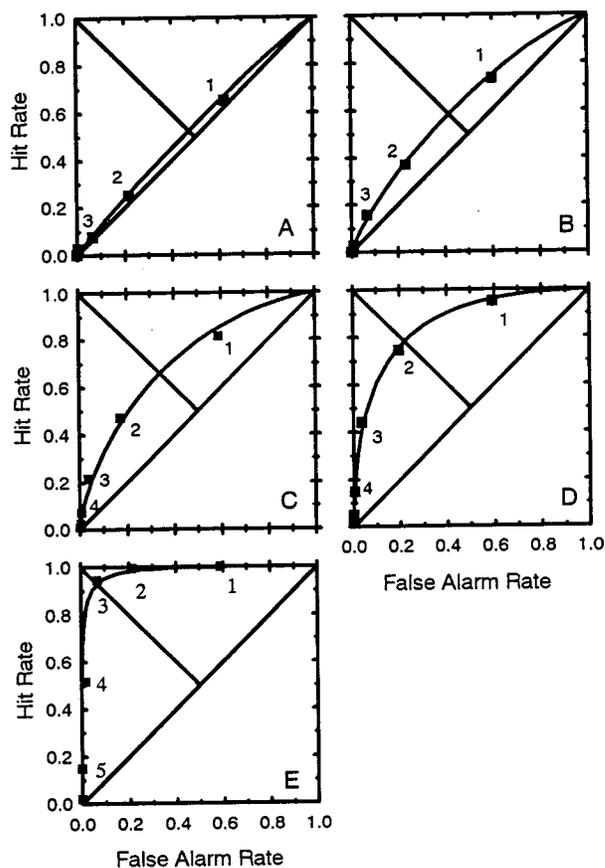


FIG. 2. Receiver operating characteristics of the A1 receptor cell. Stimulus duration 5 ms, five different sound-pressure levels, (a) 28 dB, $d' = 0.13$, (b) 32 dB, $d' = 0.4$, (c), 34 dB, $d' = 0.8$, (d) 36 dB, $d' = 1.6$, and (e) 40 dB, $d' = 3.0$. Solid lines are best-fitting ROC curves (Gaussian assumption, equal variance). The response criterion (number of spikes, c), corresponding to each data point, is indicated. For clarity, data points in the lower left corners are not marked.

the number of observations where there was an equal number of spikes in the two intervals, and m_{total} is the total number of observations. The 2AFC percent correct is monotonic with d' (Green and Swets, 1966) and could be converted to d' values by means of a table (Elliott, 1964).

In the calculation of d'_{ROC} , all spikes in the 50-ms intervals were included. To assess how the size of the observation interval influenced the calculated detectability, d' was calculated from the 2AFC evaluation for different subintervals (windows) ranging from 1 to 50 ms. Thus for a 15-ms evaluation window size, the number of spikes found between onset and 15 ms after onset (SN -interval) was compared to the number found in a similar interval prior to stimulus onset (N -interval). For each window size, a percent correct was calculated as above, and a corresponding d' could be found. For each signal duration and intensity there was an optimal size of the evaluation window, which resulted in the maximal detectability. This maximal d' was used in the evaluation of the influence of stimulus duration on detectability and was termed d'_{2AFC} . Thus two detectability indices were calculated for each stimulus—one by use of a 50-ms evaluation window, irrespective of stimulus duration (d'_{ROC}), and one where the window size was optimally adjusted for each duration (d'_{2AFC}).

C. Influence of signal duration on detectability

The detectability is a monotonic function of signal-to-noise ratio. The noise in the present experiments was most likely determined by the spontaneous activity of the receptor, which is indistinguishable from, though probably not caused by, external, acoustic noise. The noise is considered constant and independent of the signal and d' will therefore be a monotonic increasing function of signal intensity. The relationship between d' and stimulus intensity (expressed in dB SPL for convenience) was described by fitting exponential functions to the data (Egan *et al.*, 1969),

$$d'(SL) = ba^{SL}, \quad (2)$$

where SL is the sound pressure level in dB SPL and a and b are constants. The constant a determines the steepness of the increase in detectability and equals the slope of the detectability function, when plotted on a logarithmic axis. The constant b is an offset parameter, equal to the detectability index of a 0-dB SPL signal.

For each signal duration, Eq. (2) was fitted to d'_{2AFC} . Only data points for $d' < 3$ were included. At detectability indices higher than 3, corresponding to more than 98% correct, an unattainable high number of samples per data point is required in order to estimate d' reliably.

II. RESULTS

The mean number of spikes counted in the SN -interval increased with intensity of the signal, whereas the mean number of spikes in the N -interval remained constant. In Fig. 1 is shown, for a 5-ms signal at five different intensities, the occurrence of spikes in the N - and the SN -intervals, separated in bins from none to six spikes per stimulation. The histograms of spikes in the N -interval show that about 80% of these intervals contained either none or only one spike. The last 20% contained between two and five spikes. The mean number of spikes in the N -interval remained the same at all signal intensities (Kruskall-Wallis test: $H = 17.08$, 19 df, $P = 0.59$). The histograms for the spikes in the SN -intervals, on the other hand, showed marked changes as the signal intensity was increased. At the lowest intensity shown [28 dB SPL, Fig. 1(a)], the mean number of spikes counted was almost identical to the counts from the N -interval (paired t -test on difference between mean number of spikes in N - and SN -intervals: $P = 0.07$, $n = 5$ sessions of 50 stimulations). As the intensity was increased, the mean number of spikes in the SN -interval increased, which is observed as a right shift in the histograms. At the highest intensity shown in the figure [40 dB SPL, Fig. 1(e)] there was little overlap between the spike counts in the N - and SN -intervals.

A. Receiver operating characteristics

The ROC curves derived from the data in Fig. 1 are shown in Fig. 2. Each curve was constructed by plotting the hit and false-alarm rates associated with incremented values of the criterion, c . The index of detectability (d'_{ROC}) corresponding to the best-fitting ROC curve (Gaussian assumption, equal variance) is given in each figure.

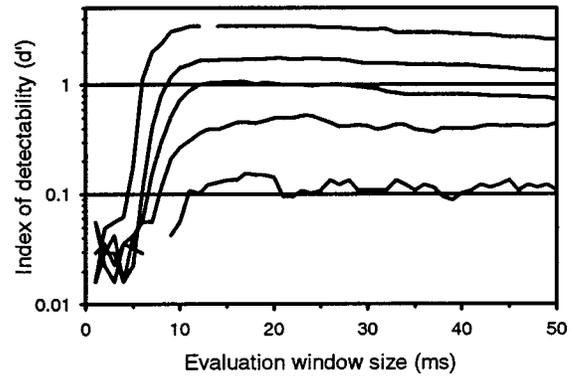


FIG. 3. The index of detectability, d' , as a function of the size of the window used in the two-alternatives forced-choice (2AFC) evaluation, as described in the text, at five different sound-pressure levels. From bottom and up: 28, 32, 34, 36, and 40 dB. The gap in the 40-dB SPL curve is due to a calculated value of d' of infinity and the gap in the 28-dB curve due to d' values below 0.01.

At low signal intensities [e.g., 28 dB SPL, Fig. 2(a)], the points of the ROC curve are close to the lower left–upper right diagonal, corresponding to a low index of detectability. With increasing signal intensity, the curve shifts toward the upper left corner, indicating an increase in d' .

In the above analysis, the size of both observation windows (N - and SN -intervals) was fixed at 50 ms. The influence of the size of the observation window on the detectability is shown in Fig. 3. The index of detectability was calculated from a 2AFC evaluation of the data as described in the methods section. Results from five different signal intensities (same as in Figs. 1 and 2) are shown. With a window size smaller than 5 ms, d' was close to zero at all intensities, reflecting the latency of the response. An increase in interval size to between 10 and 15 ms resulted in an increase in d' , with a steeper increase the more intense the signal. The increase also occurred earlier at higher intensities, indicating a decrease in latency. Increasing the size of the observation window beyond 20–25 ms had little effect at low intensities but resulted in a decrease of the detectability at higher signal intensities, more pronounced the higher the intensity.

If the underlying distributions can be considered Gaussian, with identical variance, then d' calculated from a ROC curve should be identical to d' calculated from the percent correct in a 2AFC paradigm (Green and Swets, 1966). No significant difference was found between d' , calculated from the 2AFC paradigm with a window size of 50 ms and d'_{ROC} (t -test on difference in slopes of linear regressions, $P = 0.88$; t -test on difference in elevation, $P = 0.91$; $n = 16$ stimulus intensities, 250 stimulus presentations at each intensity, 5-ms duration).

The detectability index d'_{2AFC} was found from the curves in Fig. 3 and corresponding curves for other intensities and durations as the maximum value of each curve. The two estimates of the detectability of a 5-ms signal at different intensities are shown in Fig. 4.

For each of the six different signal durations and for a range of intensities, separated in 1-dB steps, d'_{2AFC} was calculated. Results are shown in Fig. 5.

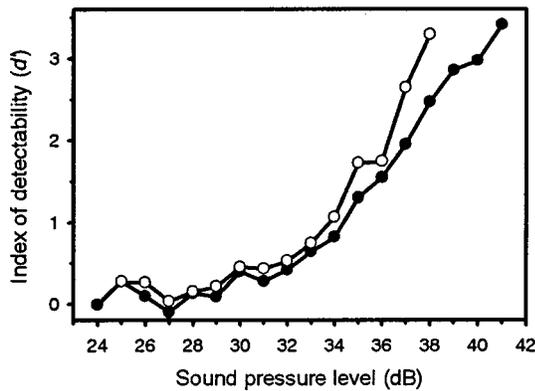


FIG. 4. Index of detectability, d' , for a 5-ms stimulus at different stimulus intensities. The detectability was calculated both from a maximum likelihood fit to the data points in the corresponding ROC curves (d'_{ROC} , solid circles), and from a two-alternatives forced-choice analysis of the data (d'_{2AFC} , open circles). Each data point is based on 250 stimulus presentations, except at 24 and 41 dB SPL, which are based on 150 and 200 stimulus presentations, respectively.

The detectability increased with signal intensity for all six durations, approximately following a straight line when plotted on log axes, and with a larger value of b [offset parameter in Eq. (2)] the larger the duration (seen as a shift of the curves to the left). Equation (2) was fitted to the data of each session and the results are shown in Table I(a). No significant differences were found between the slope parameter a across different stimulus durations (ANOVA, $P = 0.61$, $n = 31$ sessions with 900 stimulus presentations in each, six different stimulus durations). This allowed for the second fit of Eq. (2) to the data, with the common slope equal to the overall mean value (1.31 , $\text{s.e.} = 0.002 \text{ dB}^{-1}$). The results are shown in Table I(b) and the corresponding detectability functions are shown as straight lines Fig. 5.

By interpolation, based on Eq. (2), signal intensities corresponding to $d' = 1$ were calculated for each session (consisting of 50 stimulus presentations at 18 different signal intensities). The mean was calculated for each stimulus duration and plotted in Fig. 6(a), which shows the duration/intensity trade-off of the receptor. The longer the duration, the lower sound pressure required to obtain a detectability corresponding to $d' = 1$. This decrease was 1.5 dB per doubling of stimulus duration (linear regression: $r^2 = 0.98$, regression line not shown). Included in Fig. 6(a) are the pre-

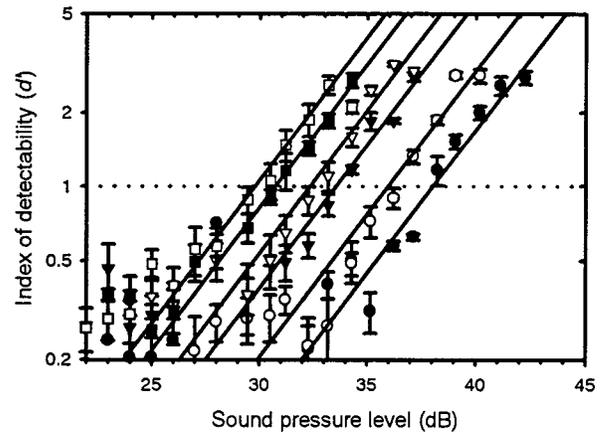


FIG. 5. Index of detectability, d'_{2AFC} , for different signal intensities and six different signal durations (from left to right: 40, 20, 10, 5, 2, and 1 ms). Solid lines are fitted exponential functions, fitted from Eq. (2) as described in the text. The dotted horizontal line indicates the standard level corresponding to the data points in Fig. 6(a). Error bars indicate standard errors.

dicted thresholds of the multiple look model (Viemeister and Wakefield, 1991; see also discussion below) and the predictions of an energy detector with a time constant τ of 3.4 ms (Tougaard, 1996), both curves arbitrarily offset to pass through the 5-ms data point.

Data from a limited analysis on a pooled set of recordings from individuals of the noctuid *Spodoptera littoralis* are shown in Fig. 6(b). Thresholds were calculated in the same way as above from a 2AFC analysis of the data, the only difference is that each data point is the mean of a variable number of sessions from five different individuals, explaining the much larger standard errors.

III. DISCUSSION

The following discussion falls into two parts. One is methodological, related to the use of ROC analysis on single-cell data, the other is concerned with the relation between stimulus duration and detectability in the A1 receptor of the moth ear.

A. ROC analysis on single-cell data

From Fig. 1 it is clear that an unambiguous threshold cannot be identified for the noctuid A1 cell. Even at the low-

TABLE I. Results of fitting Eq. (2) to the calculated indices of detectability (d'_{2AFC}) for the six different signal durations. (a) Fit of both slope (a) and offset (b) parameters to the data. (b) Fit of offset parameter (b) to the data with a common value of the slope parameter ($a = 1.31 \text{ dB}^{-1}$).

Duration (ms)	n^a	A		B
		a parameter (s.e.) dB^{-1}	b parameter (s.e.)	b parameter (s.e.)
1	5	1.28 (0.024)	$1.64 \cdot 10^{-4}$ ($7.39 \cdot 10^{-5}$)	$4.00 \cdot 10^{-5}$ ($2.20 \cdot 10^{-6}$)
2	5	1.32 (0.028)	$9.76 \cdot 10^{-5}$ ($4.46 \cdot 10^{-5}$)	$6.93 \cdot 10^{-5}$ ($2.37 \cdot 10^{-6}$)
5	5	1.29 (0.026)	$4.52 \cdot 10^{-4}$ ($2.18 \cdot 10^{-4}$)	$1.32 \cdot 10^{-4}$ ($4.44 \cdot 10^{-6}$)
10	5	1.34 (0.033)	$1.87 \cdot 10^{-4}$ ($8.41 \cdot 10^{-5}$)	$1.82 \cdot 10^{-4}$ ($9.52 \cdot 10^{-6}$)
20	5	1.29 (0.025)	$8.90 \cdot 10^{-4}$ ($4.36 \cdot 10^{-4}$)	$2.80 \cdot 10^{-4}$ ($1.24 \cdot 10^{-5}$)
40	6	1.31 (0.020)	$5.37 \cdot 10^{-4}$ ($2.22 \cdot 10^{-4}$)	$3.60 \cdot 10^{-4}$ ($4.37 \cdot 10^{-5}$)

^a n equals the number of sessions, each consisting of 900 stimulus presentations, 50 at each of 18 stimulus levels.

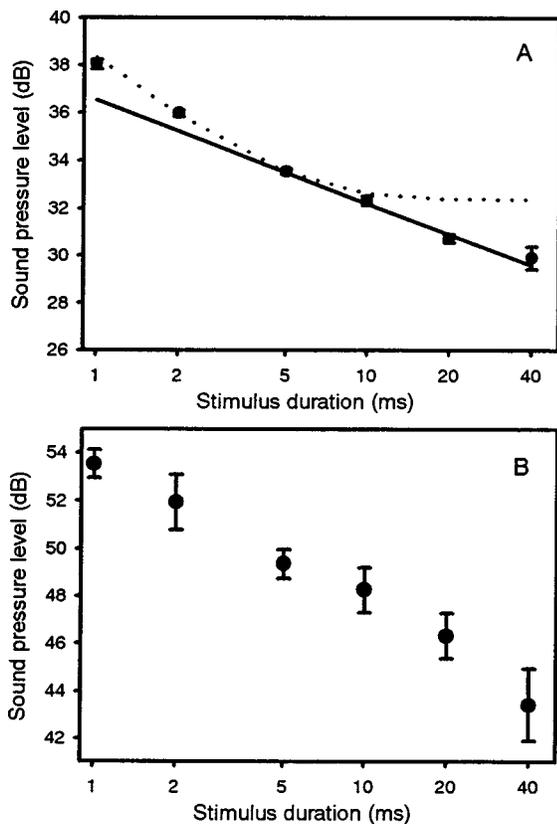


FIG. 6. Duration/intensity trade-off in the A1 receptor cell. Plotted against stimulus duration are the intensities corresponding to $d'_{2AFC} = 1$. (a) *Noctua pronuba* cell data from Fig. 5. Included are the predicted thresholds of the multiple look model (solid line) and the predicted thresholds of an energy detector with a time constant τ of 3.4 ms (dashed line). Both curves are arbitrarily offset on the sound-pressure level axis to pass through the 5-ms data point. (b) Pooled data from 73 sessions on five individuals of *Spodoptera littoralis*. Error bars in both (a) and (b) indicate standard errors.

est intensity shown [Fig. 1(a)], a difference between means of the two spike count histograms (N - and SN -condition, respectively) was seen, and this difference increased gradually with stimulus intensity. This is also reflected in the ROC curves, where a gradual increase in detectability is observed, seen as an increased curvature of the ROC curves with intensity.

1. The index of detectability

Assuming the underlying distributions of noise and signal+noise are Gaussian with equal variance, ROC curves can be described by an index of detectability, d' . By definition, d' equals the difference of the means of the underlying distributions of noise and signal+noise, with the difference expressed in units of standard deviations. It should be noted that the underlying distributions are not identical with the spike count histograms shown in Fig. 1. As is also the case for psychophysical applications, the nature of the distributed parameter is not entirely clear. In the present preparation, this parameter is likely to be a loosely defined measure of the average receptor cell depolarization, averaged over the entire evaluation window. While there is no way to identify the shape of the underlying distributions based on a ROC curve

alone (Kaernbach, 1991), the symmetry around the negative diagonal observed in the present data is consistent with the Gaussian assumption with equal variance.

The introduction of d' provides a single parameter description of the ROC curve, and thus the detectability of a given signal at a given intensity. This allows for a direct, criterion independent comparison between different experiments, performed under different conditions. Since the percentage correct in a 2AFC paradigm is monotonic with d' , this value is equally useful in this respect.

2. Two alternatives forced choice

Calculating d' from ROC curves is quite laborious and requires a high number of replicates. A simpler alternative is to evaluate the data with a 2AFC paradigm. This approach also has the advantage of providing reliable estimates of d' in situations with low spontaneous activity. The construction of a ROC curve requires the presence of a not too low spontaneous activity, since data points will otherwise cluster in the lower left corner of the ROC plot. The 2AFC paradigm will work even in the almost complete absence of spontaneous activity and still allow a calculation of d' (from percentage correct responses). This is, however, still under the assumption that the underlying distributions are Gaussian and with equal variance, an assumption which must be justified by other means.

3. Size of counting window

When comparing responses to signals of different duration, the question of evaluation window size is always present. One can either choose to compare responses in evaluation windows of similar size (large enough to contain the response to the longest signal), or to reduce the size of the window for the shorter stimuli so as to better match the actual duration of each response. Both approaches are likely to affect the analysis in different ways. With classical threshold measures, such as a fixed number of spikes, it is not straightforward to analyze this question and determine which approach is the appropriate one. By using a detectability based method (such as the 2AFC), however, it becomes possible to study how the size of the window in which spikes are counted affects the results obtained.

The detectability of a 5-ms signal, measured with different window sizes, is shown in Fig. 3. A window of about 20 ms matched the duration of the response quite well, and the largest detectability was observed for windows of approximately this size. Smaller windows did not contain all spikes elicited by the signal, explaining the lower detectability.

The detectability also decreased if the size of the window was increased beyond 25 ms, most pronounced at higher intensities. At least two reasons for this seem plausible, both related to a decrease in signal-to-noise ratio.¹

If the window size extends beyond the duration of the signal response, more noise is added to the analysis, since any additional spikes added are likely to be unrelated to the stimulus. This causes a decrease in d' . This applies not only to extending the window beyond the end of the response, but also if the window begins before response onset. Ideally one

would shift the beginning of the counting window to match the beginning of the response to compensate for neural latency. Due to the low spontaneous activity of the receptor this was not done in the present experiments and the influence on d' is considered to be marginal.

The second factor likely to cause a decrease in detectability for window sizes larger than the duration of the response, is a post excitatory suppression of the receptor. A post excitatory suppression is clearly present in the A1 receptor (Perez and Coro, 1986), seen as a decrease in spontaneous activity immediately following stimulus offset. The calculation of d' under the 2AFC paradigm is based on the assumption that the noise in the SN -interval is identical to the noise in the N -interval. The latter is then used in the calculations, since it is easily measured in contrast to the noise in the SN -interval. However, if a post excitatory suppression is present, the noise in the SN -interval is suppressed, and the assumption is violated. The detectabilities calculated will be based on a signal-to-noise ratio which is too low, since the higher noise level in the N -interval is used in the calculations. This will result in an underestimation of the detectability. Put this in another way: the post excitatory suppression suppresses some spontaneous spikes, which would otherwise have occurred after stimulus offset and added to the spikes elicited by the signal. This reduces the mean number of spikes observed in the SN -interval, which again reduces the detectability.

Using a simple spike count criterion it seems difficult to devise an appropriate strategy for dealing with stimuli of different durations. With thresholds based on detectabilities the situation is different. Since d' (and 2AFC percent correct) is independent of the number of spikes used as a criterion, it is permitted to reduce the size of the counting window such that it matches the duration of the response closely. This was done with the d'_{2AFC} index.

B. The influence of stimulus duration on d'

The detectability increased not only with signal intensity but also with duration (Figs. 5 and 6). This duration/intensity trade-off is usually attributed to a temporal integration (Adams, 1971; Surlykke *et al.*, 1988; Tougaard, 1996). It has, however, previously been shown (Tougaard, 1998) that the decrease in threshold with increasing stimulus duration in the A1 receptor is caused by two factors. One factor is an energy detector with a short time constant, the other is a statistical factor, sensitive to the choice of threshold criterion. The contribution from the energy detector can be evaluated by other types of experiments [e.g., double click stimuli, Tougaard (1996), or amplitude modulation and gap-detection experiments, Surlykke *et al.* (1988)]. By calculating indices of detectability it is possible to assess the statistical component.

The statistical improvement of detection with duration has been described, among others, by Green and Swets (1966), Viemeister and Wakefield (1991), and Viemeister *et al.* (1992), in the latter two papers under the term “multiple looks.”

If detection of a signal of duration d can be considered equivalent to detecting n single intervals (or “looks”), each

of duration d/n , then the summated detectability of the entire signal is given as

$$d' = \sqrt{\sum_{i=1}^n (d'_i)^2}, \quad (3)$$

where d'_i are the detectabilities of the n individual intervals of duration d/n . Equation (3) is valid under the assumption that the individual “looks” are independent of each other and that the information from all intervals are combined in an optimal way. If we can furthermore assume that all the individual detectabilities are equal, we get

$$d' = d'_i \sqrt{n}. \quad (4)$$

Thus d' increases with the square root of the duration increase (for further details and formal proof, see Green and Swets, 1966). The predicted threshold function of the multiple look model in Fig. 6(a) is based on these assumptions.

Figure 6(a) shows that most of the threshold decrease with increased duration observed in the present experiment can be accounted for by the multiple look model. However, at durations below 5 ms, thresholds change more with duration than predicted by the model. This is due to the energy detector properties of the receptor. This energy detector, which in *N. pronuba* has a time constant of 3.4 ms (Tougaard, 1996, 1998), will cause a decrease in thresholds with up to 3 dB per doubling of duration. This will, however, only be for short durations, up to a few times the time constant. The predictions of the energy detector are shown in Fig. 6(a) as the dotted line.

Combining the two models, desirable as it may be, is, however, not trivial. The premises of the two models are different, the most important difference being that the simple energy detector model does not incorporate noise at all, whereas noise is a central parameter in the stochastic model. Below follows an attempt to combine the two models based on qualitative arguments.

One of the central assumptions of the multiple look model is independence between individual “looks.” This assumption is likely to be fulfilled at long durations, but at short durations—on the order of magnitude of the energy detector time constant and shorter—it is almost certainly violated. Integration and independence are mutually exclusive. The very nature of a temporal integrator is that the output at any given time depends on events in the past.

The extent of the dependence backward in time is reflected in the energy detector time constant. Any two events temporally separated by more than five times the time constant (approx. 15 ms for *N. pronuba*) can be considered practically independent events. Events which occur closer to each other will be increasingly interdependent the closer they are in time.

Thus for duration/intensity trade-off experiments, we should expect thresholds to be divided into three regions, based on the duration of the signal, and not sharply separated. The first region covers thresholds for short duration signals, significantly shorter than the energy detector time constant. For these durations, any two events which occur within the duration of the signal will be interdependent to

such a high degree that the multiple look model cannot be applied. In this, the energy detector region, thresholds are predicted to decrease with 3 dB per doubling of duration.

The second region covers thresholds for signals of durations significantly longer than the time constant. For these durations, the energy detector does not contribute significantly to threshold changes with duration. This means that events occurring within the duration of the signal can be considered independent, and the multiple look model is applicable. In this, the multiple look region, detectability is predicted to increase with $\sqrt{2}$ per doubling of duration [Eq. (4)]. The actual decrease in thresholds is not given directly from Eq. (4), but will depend on the slope of the intensity/detectability curves (as in Fig. 5). A doubling of duration will result in a threshold decrease of $\ln \sqrt{2}/\ln a$, where a is the slope parameter from Eq. (2). This corresponds to a slope of -1.3 dB per doubling of signal duration for a value of a of 1.31, as in the present data. This is the slope of the solid curve in Fig. 6(a).

The middle region of thresholds is for signals of intermediate durations, comparable to the size of the time constant. This is a gray zone where the energy detector on one hand has some influence on thresholds, but with a contribution less than -3 dB per doubling of duration. On the other hand, events which occur within the duration of the signal are sufficiently independent for d' to improve due to the statistics, although with less than $\sqrt{2}$ per doubling of duration. Due to the partial independence between individual looks, it is difficult to calculate specific predictions for this region.

The concept of the three different regions is essential to understand when considering problems related to temporal integration. By knowing which region is pertinent to a given experiment, it is possible to apply the appropriate model (energy detector vs multiple look model). If one can measure the energy detector time constant by other methods (e.g., double click detection), it is also possible to identify the middle region of durations, where neither of the models produces reliable predictions on their own.

The present analysis has thus demonstrated that receiver operating characteristics analysis can be successfully applied to an insect sensory receptor cell and that this analysis can provide means of assessing central questions regarding the temporal integration of the receptor—questions, which would otherwise be difficult or impossible to assess.

ACKNOWLEDGMENTS

This study was supported by the Danish Natural Science Research Council and the Danish National Research Foundation (Center for Sound Communication). Discussions with Bertel Møhl, Friederike Gilbert, Kristian Beedholm, and Jens Astrup played a major role in maturing the ideas presented here. Also the comments of Bertel Møhl, Jens Astrup,

Annemarie Surlykke, and A. Supin significantly improved the final manuscript. The maximum likelihood procedure used to fit d' values to ROC curves was kindly provided by Bernt Guldbrandtsen.

¹In the present preparation, the signal-to-noise ratio is most likely determined by the spontaneous activity of the receptor. Thus noise should be understood as neural noise rather than external, acoustic noise. Although the term signal-to-noise ratio in this case is somewhat loosely defined, it nevertheless still retains its meaningfulness in relation to the detectability.

- Adams, W. B. (1971). "Intensity characteristics of the noctuid acoustic receptor." *J. Gen. Physiol.* **58**, 562–579.
- Barlow, H. B., Levick, W. R., and Yoon, M. (1971). "Responses to single quanta of light in the retinal ganglion cells of the cat," *Vision Res.* **11**, Suppl. 3, 87–101.
- Cohn, T. E., Green, D. G., and Tanner, W. P. J. (1975). "Receiver operating characteristic analysis. Application to the study of quantum fluctuation effects in optic nerve of *Rana pipiens*," *J. Gen. Physiol.* **66**, 583–616.
- Egan, J. P., Lindner, W. A., and McFadden, D. (1969). "Masking-level differences and the form of the psychometric function," *Percept. Psychophys.* **6**, 209–215.
- Elliott, P. B. (1964). "Tables of d' ," in *Signal Detection and Recognition by Human Observers*, edited by J. A. Swets (Wiley, New York), pp. 651–684.
- Fay, R. R., and Coombs, S. L. (1992). "Psychometric functions for level discrimination and the effects of signal duration in the goldfish (*Carassius auratus*): Psychophysics and neurophysiology," *J. Acoust. Soc. Am.* **92**, 189–201.
- FitzHugh, R. (1957). "The statistical detection of threshold signals in the retina," *J. Gen. Physiol.* **40**, 925–948.
- Ghiradella, H. (1971). "Fine structure of the noctuid moth ear," *J. Morphol.* **134**, 21–46.
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).
- Kaernbach, C. (1991). "Poisson signal-detection theory: Link between threshold models and the Gaussian assumption," *Percept. Psychophys.* **50**, 498–506.
- Lee, B. B., Wehrhahn, C., Westheimer, G., and Kremers, J. (1993). "Macaque ganglion cell responses to stimuli that elicit hyperacuity in man: Detection of small displacements," *J. Neurosci.* **13**(3), 1001–1009.
- Miller, L. A., and Surlykke, A. (2000). "How some insects detect and avoid being eaten by bats. The tactics and counter tactics of prey and predator," *BioScience* (in press).
- Pérez, M., and Coro, F. (1986). "Physiological characteristics of the tympanic organ in noctuid moths. III. Post-stimulus phenomena," *J. Comp. Physiol. A* **159**, 115–122.
- Relkin, E. M., and Pelli, D. G. (1987). "Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures," *J. Acoust. Soc. Am.* **82**, 1679–1691.
- Roeder, K. D. (1974). "Responses of the less sensitive acoustic sense cells in the tympanic organs of some noctuid and geometrid moths," *J. Insect Physiol.* **20**, 55–66.
- Surlykke, A., and Miller, L. A. (1982). "Central branchings of three sensory axons from a moth ear (*Agrotis segetum*, Noctuidae)," *J. Insect Physiol.* **28**, 357–364.
- Surlykke, A., Larsen, O. N., and Michelsen, A. (1988). "Temporal coding in the auditory receptor of the moth ear," *J. Comp. Physiol. A* **162**, 367–374.
- Tougaard, J. (1996). "Energy detection and temporal integration in the noctuid A1 auditory receptor," *J. Comp. Physiol. A* **178**, 669–677.
- Tougaard, J. (1998). "Detection of short pure tone stimuli in the noctuid ear: What are temporal integration and integration time all about?," *J. Comp. Physiol. A* **183**, 563–572.
- Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal integration and multiple looks," *J. Acoust. Soc. Am.* **90**, 858–865.
- Viemeister, N. F., Shivapuja, B. G., and Recio, A. (1992). "Physiological correlates of temporal integration," *Adv. Biosci.* **83**, 323–330.

Generation and maintenance of bubbles in small tubes by low-frequency ultrasound

Kwok W. Lam

Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada

James M. Drake

Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada and Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada

Richard S. C. Cobbold^{a)}

Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada

(Received 13 March 1999; accepted for publication 4 August 1999)

A system has been developed to generate and maintain a bubble population in nonscattering fluid media within a small 1.3-mm (i.d.) transparent silicone tube by use of low-frequency ultrasound at 28.8 kHz. The system uses a sequence of three different kinds of acoustic cavitation processes: transient, stable and cyclic. In the final step of the sequence, low-duty cycle (~10%) and low-pressure (~30 kPa) ultrasound is used to maintain a microbubble population whose mean diameter is ~12 μm , which can be used as Doppler ultrasound scatterers. To understand and measure the behavior of bubbles for three different kinds of cavitation processes in the tube, a CCD imaging system and a synchronized short duration flash were used to capture movie frame sequences. Subsequent analysis of the frames enabled the following bubble parameters to be measured: size distributions, fractional volume concentration, flow characteristics, rise velocity, dissolution time, growth rate and various bubble formation processes. In addition, the cavitation threshold of distilled water inside the tube was found to be $\sim 292 \pm 99$ kPa. Techniques were also developed to identify successful bubble initiation and excitation by either detecting the harmonic components of the bubble emission power spectra or by detecting the power level of the Doppler spectrum. © 1999 Acoustical Society of America. [S0001-4966(99)06411-5]

PACS numbers: 43.80.Vj, 43.80.Gx, 43.25.Yw [FD]

INTRODUCTION

The work described in this paper was motivated by a proposed method for measuring flow of a nonscattering fluid in small diameter silicone tubes. Such tubes are routinely implanted as shunts to reduce and control cerebrospinal fluid (CSF) pressure in patients with hydrocephalus. Because these shunts frequently become blocked, a simple means for noninvasive flow measurement is needed. In our previous work, we have shown^{1,2} that the flow velocity of either degassed distilled water or CSF in shunt tubes can be measured by first creating cavitation bubbles and then using these as scatterers for Doppler ultrasound measurements.

Three distinct types of cavitation are often identified, transient (inertial), stable (noninertial) and cyclic.³ Transient cavitation generally involves a high-intensity acoustic field in which bubbles or cavities undergo a major expansion followed by a very rapid collapse accompanied by the release of energy often with the formation of microbubble fragments. Stable cavitation describes the process whereby bubbles in a periodic acoustic field oscillate around their equilibrium size, often in a nonlinear manner. With sufficient dissolved gas in the liquid and an acoustic pressure higher than a certain threshold, bubbles will gradually grow by a process known

as rectified diffusion.⁴ Finally, cyclic cavitation⁵ involves the emission of microbubbles from the surface of a much larger bubble due to instabilities created by surface wave oscillations induced by the acoustic field. Some of these microbubbles are used as cavitation nuclei to support the growth of new bubbles. For acoustic frequencies lower than 30 kHz, the acoustic pressure required for the emission of microbubbles from bubbles smaller than resonant size can be relatively small, e.g., <20 kPa. Generally, acoustic cavitation is highly dependent on the acoustic pressure and frequency, the gas saturation level, the preexisting nuclei concentration and distribution, temperature, ambient pressure, surface tension and fluid viscosity.⁶

For many years Doppler ultrasound has been used for detecting and measuring naturally occurring bubbles either as they flow in tubes^{7,8} or when they are produced in blood vessels as a result of decompression.⁹ In their study of bubbles generated by cavitation, Miller *et al.*¹⁰ used second harmonic emissions from resonantly excited bubbles to measure the number of bubbles created. In this system bubbles were generated in tap water flowing in a 21-mm-diam tube, using frequencies in the range of 0.514–1.61 MHz and at pressure amplitudes of 200 to 400 kPa. Subsequently, the same group reported¹¹ the results of exposing canine left ventricles to high-intensity ultrasound (0.514–1.61 MHz, up to 16 W/cm²) but found no evidence that this could create

^{a)}Electronic mail: cobbold@ecf.utoronto.ca

cavitation bubbles unless it was seeded with bubbles. Interesting studies on inducing cavitation in both water and blood in a plastic arterio-venous shunt using a lithotripter shock wave at a frequency of 1.6 MHz have been reported by Williams *et al.*¹² Although they observed cavitation *in vitro*, they found no evidence that cavitation bubbles could be created *in vivo*. More recently, Ivey *et al.*¹³ detected cavitation *in vivo* using very-high-intensity pulses (e.g., 4.3 kW/cm² at 1.8 MHz, 250 ms) focused on the canine abdominal aorta.

The processes associated with acoustic cavitation in a highly confined region, such as inside a small diameter tube, do not appear to have been investigated. As will be demonstrated, the cavitation process in a small diameter tube usually involves clouds of bubble with a wide distribution of sizes whose dynamical behavior differs from that of a single spherical bubble in an unconfined volume. Specifically, because the volume occupied by the bubble cloud may be comparable to the surrounding fluid volume and the distance between the bubbles may be small, the bubble behavior is likely to differ considerably from that of a free bubble. Moreover, bubbles trapped on the tube wall will generally be nonspherical since the tube surface is unlikely to be perfectly hydrophilic. As a result, theoretical models that assume spherical, unattached bubbles in an infinite medium must be used with great caution.

Recently, we have developed a novel system to generate and maintain bubbles in distilled water inside a small diameter silicone tube (i.d.=1.3 mm) by low-frequency ultrasound (28.8 kHz) and have demonstrated that the generated microbubbles can be used as ultrasound scatterers for Doppler measurements. In this respect they act in a similar manner to commercially available ultrasound contrast agents, e.g., Alunex[®] (Molecular Biosystems, Inc., San Diego), which must be injected. Moreover, using these microbubbles we have shown that velocity measurements down to a mean velocity of ~ 1 mm/s are feasible.² A sequence of three steps used in generating and maintaining these microbubbles involved transient, stable and cyclic cavitation processes. The microbubble population was maintained by using a low-frequency, low-duty cycle and low-amplitude ultrasound.

The use of low-frequency ultrasound for a variety of applications that include emulsification and cleaning is well established and there have been many papers discussing the processes that occur. Much of this work has been reviewed in the excellent books by Young¹⁴ and Leighton.³ However, most of this work involves the processes that occur in liquid volumes that can be considered to be unbounded. As previously noted, there appears to have been no prior studies for highly confined conditions.

The purpose of the investigation described in this paper is to develop an understanding of the behavior of ultrasonically generated bubble dynamics inside a small transparent silicone tube under both static and low Reynolds number steady flow conditions. An imaging system has been used to obtain movie frame sequences showing the bubble behavior inside the tube for various cavitation processes. The information obtained from these sequences has enabled us to characterize the three-step bubble generation scheme and has helped provide plausible explanations for some observations

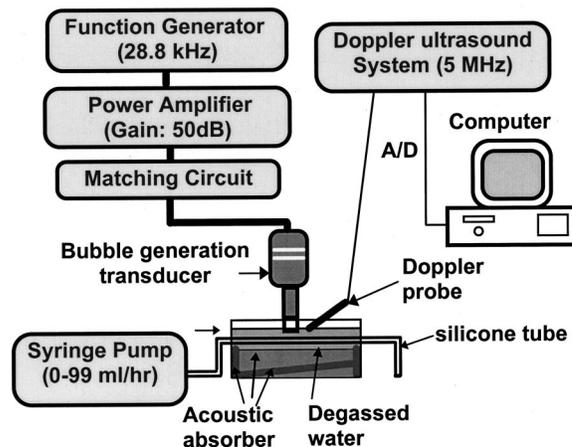


FIG. 1. Experimental system used for generating bubbles inside a silicone tube. The Doppler ultrasound system uses ultrasonically excited microbubbles as Doppler ultrasound scatterers for flow measurement.

made during Doppler ultrasound flow measurements. In addition, the paper will illustrate the techniques that can be used to reliably identify successful transient and stable cavitation inside the tube by either detecting the subharmonic and harmonic components of the emitted spectra or by measuring the backscattered Doppler power level.

I. MATERIALS AND METHODS

A. Bubble generation

The experimental system used for generating the bubbles inside the small tube is illustrated in Fig. 1. The tube was a clear silicone CSF shunt (ACCU-FLO Clear Ventricular Catheter, 82-1203, Codman & Shurtleff, Inc., Randolph, MA) with an internal lumen diameter of 1.3 mm and a wall thickness of 0.6 mm. The liquid used consisted of distilled water that had been vacuum degassed to achieve an oxygen saturation level of around 70%. This saturation level was selected so as to approximately mimic the properties of CSF as discussed previously.² Using a dissolved oxygen meter (Model 810, ATI Orion, Boston, MA) the gas saturation level was measured before filling a 60-ml plastic syringe from which all residual bubbles were expelled. A syringe infusion pump (Model 2720, Harvard Apparatus, South Natick, MA) was used to control the flow rate over the range 0 to 99 ml/h to an accuracy of ± 0.1 ml/h. The Reynolds number corresponding to 100 ml/h is $Re=27$. The liquid was then collected in a measuring cylinder where the O₂ saturation was again determined. For these low Reynolds number conditions and the long inlet length used, Poiseuille flow should be present and, consequently, the flow velocity at the tube center should be twice the mean flow velocity.

The tube was suspended in a Plexiglas[™] tank (width=10 cm, length=19.2 cm and height=19.8 cm) containing degassed distilled water with a gas saturation of less than 40%. Acoustic absorbing sheets (Sorbothane[™]) were used on the tank side walls and bottom (tilted at 10 degrees) to minimize acoustic reflections and standing waves. The presence of these sheets was found to reduce the acoustic pressure due to reflections by over 70%. The bubble generating transducer

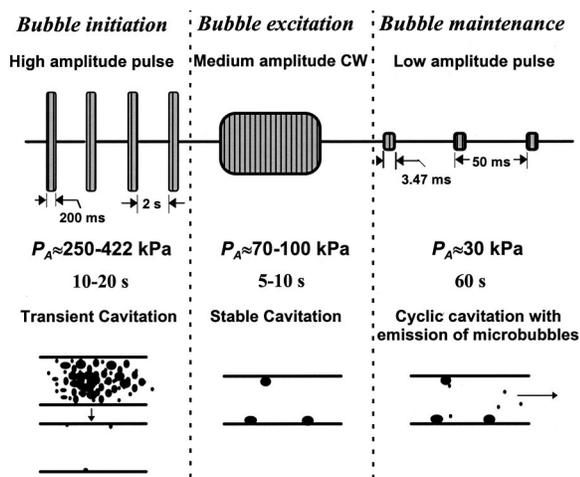


FIG. 2. Bubble generation scheme used for initiating, exciting and maintaining microbubbles inside the tube. The top section indicates the bubble generation scheme, while the bottom section represents the bubble processes.

consisted of a sandwich-type PZT transducer with a resonant frequency of 28.0 kHz coupled to an aluminum velocity transformer in the form of a tapered horn. To improve transmission efficiency into water, a quarter-wavelength Plexiglas™ acoustic matching layer was glued to the horn tip. Details of the construction have been described elsewhere.² The overall resonant frequency was found to be 28.8 ± 0.5 kHz. As previously described,² the transducer pressure field was measured by a calibrated hydrophone (Model 8103, Brüel & Kjaer, Naerum, Denmark).

The silicone tube was placed 5 mm away from the matching layer surface of the excitation transducer. At this location it was found that the acoustic pressure amplitude was linearly related to the transducer excitation voltage by $P_A = (1.42V_{\text{input}} + 7.03)$ kPa, which was used to determine the pressure from a measurement of the excitation voltage. Accurate control of the excitation waveform was achieved with a digital function generator (Rohde and Schwarz AFG, Model No. 377.2100.02, Munich, Germany) whose output was amplified by a broadband rf power amplifier (ENI Model 240L, 50 dB gain, 200 W, Rochester, NY) that was coupled to the transducer by a matching circuit.

Figure 2 illustrates a sequence of steps used for initiating, exciting and maintaining a bubble distribution in the tube suitable for Doppler ultrasound measurements.² All three stages are the subject of this investigation. They consist of (i) bubble initiation using high-amplitude pulses ($T_{\text{on}}=200$ ms on duration; $T_{\text{off}}=2.0$ s off time; typical pressure amplitude $P_A \approx 250\text{--}420$ kPa; application duration 10–20 s), (ii) bubble excitation using a medium amplitude continuous wave (CW, $P_A \approx 70\text{--}100$ kPa, 5–10 s duration), and (iii) bubble maintenance using a low-amplitude pulse [$T_{\text{on}}=3.47$ ms, $T_{\text{off}}=46.53$ ms (duty-cycle <10%) $P_A \approx 30$ kPa, 60 s].

B. Imaging of bubbles

The imaging system illustrated in Fig. 3 was developed to study the dynamical behavior of cavitating bubbles inside

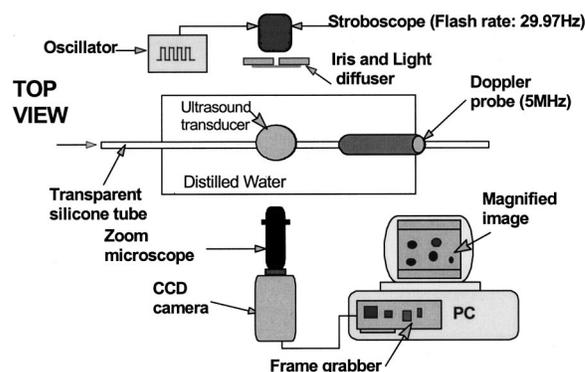


FIG. 3. View from above the imaging system that illustrates the illumination and image capture systems.

the transparent tube. The same system was also used to evaluate the bubble initiation threshold and to investigate the effects of the flow rate and the liquid properties on the success of bubble initiation. It consisted of a variable zoom microscope (VZM Model 450, Edmund Scientific, Toronto, Canada: zoom=28–180X, focal length=95 mm) coupled to a monochrome CCD camera (Cohu 4912-2010, RS-170, San Diego: 29.97 frames/s) whose electronic shutter was disabled. Images were captured and displayed in real-time using a PC with a frame grabber (LG-3, PCI, 8-bit resolution, frame rate=29.97 frames/s, Scion Corporation, Frederick, MD). To provide both the flexibility and accuracy of dynamic focusing on different positions of the tube, both the camera and the microscope were mounted on a precision 3-D translation stage with 0.1-mm step resolution.

To obtain bubble images free from motion blurring and which had good contrast, a digital stroboscope (Nova-Strobe DA 115, Monarch Instrument, Amherst, NH) was used. This illuminated the tube section by delivering a high-intensity (~ 180 mJ) short duration (~ 30 μ s) flash. The stroboscope was triggered by a very stable oscillator (Hewlett Packard, HP3325A) that was set to the CCD frame frequency (29.97 Hz). An iris, used to control the depth of field, and a light diffuser were placed in front of the stroboscope. The diffuser helped reduce image deterioration caused by surface roughness on the silicone tube wall. An image software package (Scion Image, Release Beta 2, Scion Corporation, Frederick, MD) was used to control the frame grabber board and to perform image analysis and processing.

Each movie frame had an 8-bit gray-scale resolution (256 levels) and a maximum image size of 640×480 pixels. The maximum resolution of the overall imaging system corresponded to 2 μ m/pixel, and as a result the diameter of the smallest microbubble resolvable from the image background was ~ 4 μ m. When the imaging system was operating at its maximum and minimum magnifications, the fields of view were 1.30×0.97 mm² and 8.36×6.24 mm², respectively. At these two extremes the depths of field were ~ 300 μ m (@ 90 line-pair/mm) and ~ 4 mm (@ 10 line-pair/mm), respectively. In most of the studies, movie frames were captured with a depth of field encompassing the entire CSF shunt lumen.

For quantitative analysis, bubble sizes were measured manually from individual movie frames. Dissolution and

growth measurements were performed by tracking individual bubbles over a sequence of movie frames. To characterize the velocity of the microbubbles flowing inside the tube, two consecutive movie frames were digitally subtracted. The resulting image showed the same microbubble at times corresponding to the times at which the flash occurred during the first and second frames. The vector joining the two microbubbles together with a knowledge of the time difference between the strobe flashes enabled the 2-D velocity vector in the image plane to be determined. The two components of this vector are the axial and vertical components, the latter arising from the effect of the buoyancy force. Since Poiseuille flow was present in the tube, any radial component should be small and as a result this 2-D velocity vector should be a good approximation to the total velocity vector.

The imaging system was also used to estimate the cavitation threshold and to investigate the effects of the flow rate and fluid properties on bubble initiation. To obtain the cavitation threshold pressure, the initiation scheme described earlier was used but with the pressure amplitude starting at 100 kPa and being incremented by 10 kPa after each group of five pulses. The distilled water had a gas saturation of $\sim 70\%$ and there was no flow. Before each trial, the transparent shunt was inspected to ensure that there were no bubble nuclei visible on the surface of the silicone tube. The threshold was taken to be the pressure at which a distinct bubble cloud created by transient cavitation was first observed. For each new trial, the same procedure was followed. The results of 30 independent successful trials gave the distribution of thresholds.

For a liquid with a uniform nuclei concentration, a higher flow rate should reduce the amount of time required for bubble initiation (more nuclei would pass by the excitation region). To confirm this hypothesis, experiments were performed using either the tap or distilled water at flow rates of 20 ml/h ($Re=5.4$) and 700 ml/h ($Re=189$), respectively. While the acoustic pressure pulse amplitude was fixed at ~ 250 kPa, a total of 120 independent trials (30 trials for each combination of flow rate and liquid) were performed to obtain the percentage of successful bubble initiations.

C. Detection of bubbles

In the presence of a sufficiently high amplitude CW acoustic field, the nonlinear component of bubble oscillation can become sufficiently large so that the emitted radiation contains higher harmonic,^{15,16} i.e., $2f$, $3f$, $4f$, nf , etc. The amplitude of the harmonic component is highly dependent on the type of cavitation as well as the acoustic pressure amplitude. In fact, the appearance of harmonic components in the spectrum is a useful indicator of successful stable cavitation. During transient cavitation, not only will the emitted pressure field contain higher harmonics, but subharmonics; $f/2$, and higher harmonics, $3f/2, 5f/2, \dots$, may also be present. These can be used as indicators for transient cavitation or successful bubble initiation.¹⁵

The experimental arrangement for the acoustic emission study of cavitating bubbles inside the silicone tube is illustrated in Fig. 4. Acoustic emission from the cavitation field inside the tube was recorded by the calibrated low-frequency

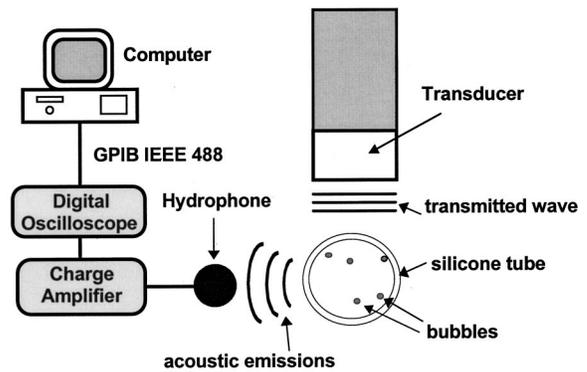


FIG. 4. Experimental system for measuring acoustic emissions from cavitating bubbles inside a tube.

hydrophone (Model 8103, Bruel & Kjaer, Naerum, Denmark) located 5 mm away from the side of the insonified tube section. The hydrophone signal was amplified by a charge-amplifier (Model 2635, Bruel & Kjaer, Naerum, Denmark) and sampled by a digital oscilloscope (Tektronix TDS320, Beaverton, OR). The captured waveform was then transferred to a PC via a GPIB interface (AT-GPIB/TNT, National Instrument, Austin, TX), and then processed and displayed as the acoustic emission spectra. During the signal acquisition, the bubble activity was simultaneously observed by the imaging system.

The Doppler signals were measured with a 5-MHz pulsed Doppler system (Model CP-1 Doppler, D. E. Hokinson, Issaquah, WA) using a 4-mm-diam focused probe and were recorded on a PC by using a 12-bit A/D card (Lab-PC+, National Instruments, Austin, TX). Typically the distance between the probe and shunt was 15–20 mm, and at this distance the sample volume length was ~ 2 mm, which was close to the inner wall to wall path length at the fixed Doppler angle of 30 degrees. By modifying the Doppler high-pass filter to ~ 3 Hz (3 dB), velocities down to 0.5 mm/s (~ 3 ml/h) could be measured at 30 degrees. The sample volume was positioned just downstream from the point of highest acoustic pressure. This enabled microbubbles to be detected before they had either dissolved or become trapped on the wall.

II. RESULTS AND DISCUSSION

Figure 5 shows three sets of movie frames taken during the initiation, excitation and maintenance steps, for a flow rate of 20 ml/h. In the first step, high-amplitude pulses are used to launch seed bubbles. During this initiation step, there can be a mixture of transient and stable cavitation and the images revealed a misty bubble cloud with a wide distribution of bubble sizes. Typically, the axial length of the bubble cloud reaches a maximum just before the pulse terminates. The cloud disappears in less than 50 ms when the high-intensity pulse terminates, leaving some seed bubbles trapped on the tube wall. In the second step a medium amplitude CW wave is used to develop a population of bigger bubbles from these seeds by means of rectified diffusion. Many of the larger bubbles become trapped on the wall. In the third step a microbubble population is perpetuated by

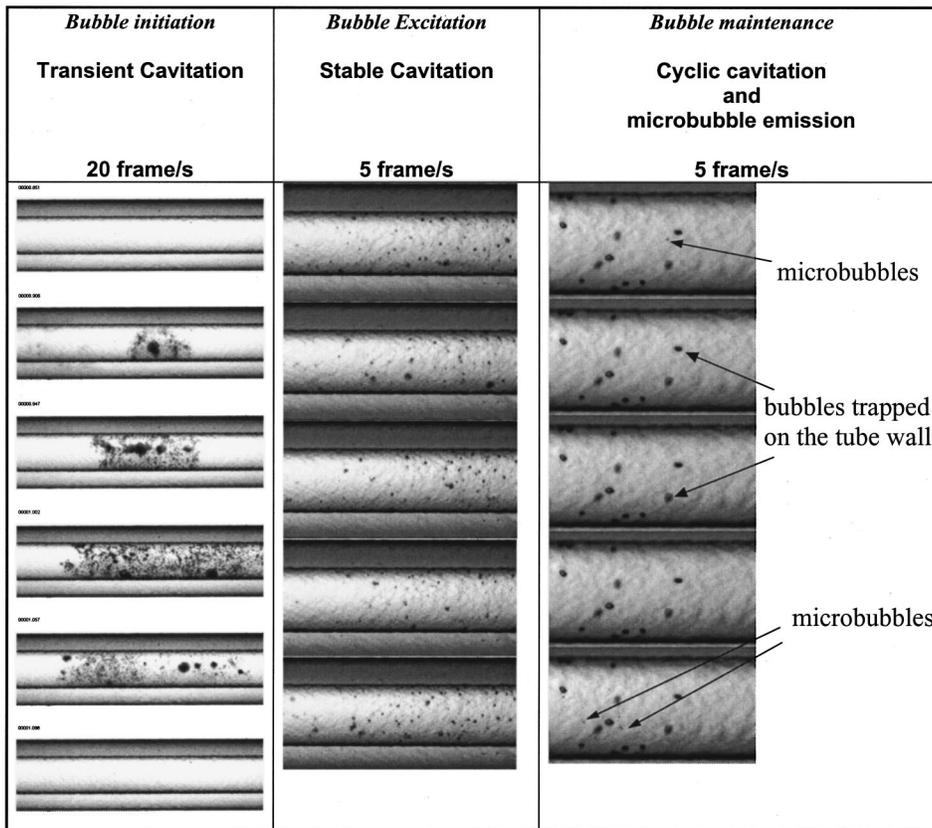


FIG. 5. Three sets of movie frames taken during initiation, excitation and maintenance steps. These were obtained on a 1.3-mm-diam transparent silicone tube at a flow rate of 20 ml/h using a CCD camera system. They provide the experimental support for the physical process involved in the three steps. In the final step, the small microbubbles (5–30- μm diameter) responsible for the Doppler signal are just visible. The depth of focus encompassed the entire tube.

means of the low-amplitude, low-duty-cycle pulse scheme shown in Fig. 2. These microbubbles are created primarily as a result of surface instabilities in larger bubbles.^{5,17} An important advantage of this maintenance scheme is that it reduces the movement of the trapped bubbles and their surface oscillations.

A more detailed examination of the characteristics of these three steps is presented below.

A. Initiation threshold

Figure 6 shows the threshold distribution for successful bubble initiation in distilled water with no flow and the percentage of successful bubble initiations for flows of 0, 40, and 90 ml/h using the maximum available pressure of ~ 422 kPa. As can be seen from Fig. 6(a), in the absence of flow, the average initiation threshold pressure was 292 ± 99 kPa

with a range from 160 to 422 kPa. At ~ 30 kHz, our measured initiation threshold pressure corresponded to that reported by Galloway¹⁸ for distilled water at a gas saturation of 70% (~ 170 kPa). System limitations did not permit the use of much higher pressures. With successful bubble initiation, the threshold for subsequent initiation that occurred within a few tens of seconds of the first was considerably reduced due to the presence of the trapped microbubbles. However, the subsequent initiation threshold was found to return to approximately the original level if sufficient time had elapsed (e.g., 5 min). In the absence of flow the success was $\sim 87\%$, which increased to $\sim 93\%$ for flows of 40 and 90 ml/h. The large variation of the initiation threshold may be partially due to variations in the bubble nuclei population in the small insonified fluid volume, but there are possibly other factors

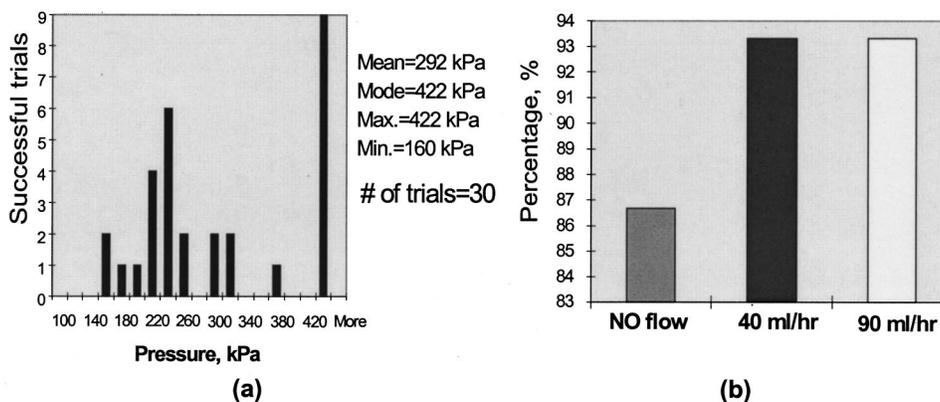


FIG. 6. (a) Distribution of the pressure thresholds for successful bubble initiation in the absence of flow. (b) Percentage of successful bubble initiations at three different flow rates: 0, 40, and 90 ml/h.

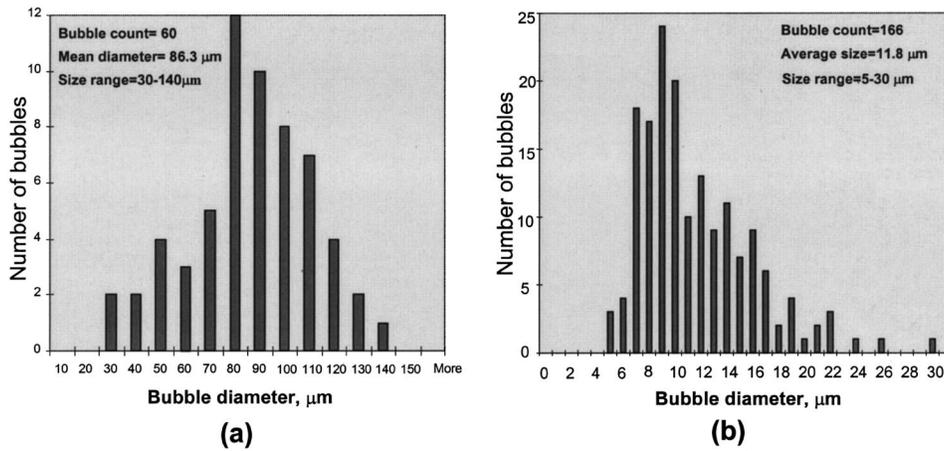


FIG. 7. Bubble size distributions. (a) Distribution of bubbles trapped on the tube wall just after excitation using CW ultrasound at a pressure of ~ 80 kPa. (b) Distribution of flowing microbubbles during the maintenance step.

that are not yet understood. We also performed measurements in both tap and distilled water for flow rates of 20 and 700 ml/h, at a fixed pressure amplitude (250 kPa). As expected, these results clearly showed a considerably higher success rate for tap water, e.g., 100% vs 87% for distilled water at 700 ml/h: it also demonstrated the important influence of flow rate. Moreover, when the tube was immersed in gelatin instead of distilled water, the initiation threshold and success rates were found to be similar.

B. Dynamics of bubbles

1. Bubble size and density distribution

Figure 7(a) shows the size distribution of the bubbles trapped on the tube wall after the excitation step terminates. The average diameter of the bubbles trapped on the tube wall was found to be $86.3 \pm 24.9 \mu\text{m}$ with a range of 30–140 μm . Figure 7(b) shows the size distribution of the flowing microbubbles emitted by the source bubble trapped on the wall during the maintenance step. These microbubbles had an average diameter of $11.8 \pm 4.5 \mu\text{m}$ and a range of 5–30 μm . Using the formula quoted by Young¹⁹ the resonant frequencies of these microbubbles are in the range 0.2–1.6 MHz,

which is much higher than the excitation frequency (28.8 kHz) used for maintenance. Moreover, because the duty-cycle and the acoustic pressure amplitude of the maintenance scheme are very low, the time averaged radiation forces should be insignificant.²⁰

By counting the number of microbubbles for each movie frame, we have also been able to estimate the bubble density and the fractional volume concentration inside the tube. By analyzing 50 movie frames for a depth of field $\sim 500 \mu\text{m}$ over a 1-mm tube length with a flow rate of 40 ml/h (8.4 mm/s) and frame rate=15 frames/s, the average number of microbubbles per frame was 4.6 ± 2.8 with a range of 0 to 16. Assuming a mean diameter of 11.8 μm , the mean fractional microbubble volume concentration was found to be $\sim 6.2 \times 10^{-6}$. At 40 ml/h (8.4 mm/s), there could be less than 150 microbubbles per second passing through a given cross section.

2. Dynamical behavior of bubbles

The various kinds of bubble activity observed in the tube during the excitation and maintenance steps will now be considered. Figure 8(a) illustrates how a microbubble trapped on

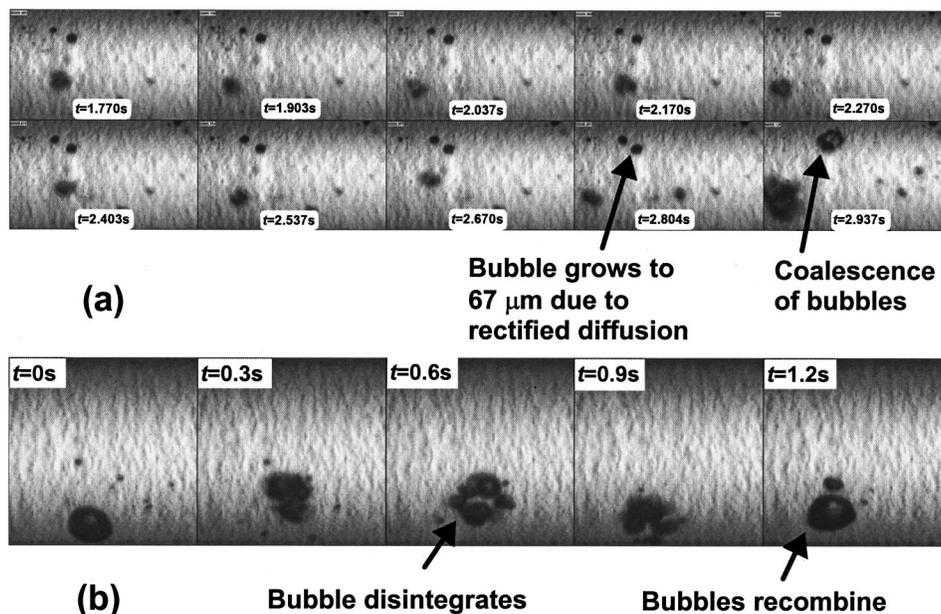


FIG. 8. (a) Later stages of growth by rectified diffusion of a bubble whose initial diameter was $\sim 32 \mu\text{m}$ at $t=0$ s. At $t=2.937$ s, the two bubbles coalesced. (b) Bubble behavior in absence of flow during the excitation step. The relative time between each image is indicated.

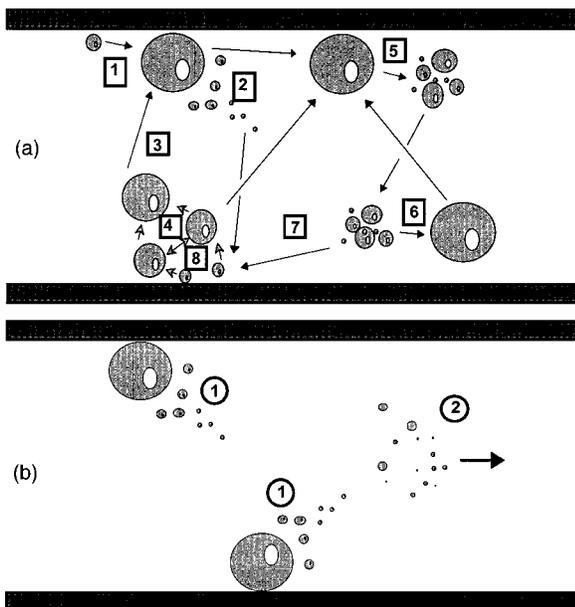


FIG. 9. Schematic illustration of the various bubble formation processes during (a) the excitation step: (1) residual bubble growth, (2) surface emission of microbubbles from the bubble trapped on the wall, (3) medium-sized bubble grows to a larger bubble, (4) coalescence of bubbles, (5) disintegration of bubble, (6) bubble recombination, (7) disintegrated bubbles trapped on the wall, and (8) growth of disintegrated bubbles and (b) the maintenance step: (1) microbubble emission and (2) emitted microbubbles carried away by the flow.

the tube wall grows from an initial diameter of 32 to $67 \mu\text{m}$ in just 2.8 s by means of rectified diffusion when 28.8 kHz CW ultrasound is present for $P_A \approx 90$ kPa. By using the rectified diffusion growth rate equation for a free spherical bubble in an infinite air-saturated water, as given by Crum and Hansen,⁴ the theoretical time required for such a growth would be greater than 27 s. In fact, if the saturation is taken to be 70% , corresponding to the fluid used, the theory predicts bubble dissolution rather than growth. This discrepancy may be due to one or more of the following factors: (1) Since a bubble trapped on the wall will no longer be a spherical bubble, the simple spherical assumption is no longer valid and the presence of the tube wall on part of its surface will undoubtedly influence the rectified diffusion process. (2) The volume inside the small silicone tube is extremely small. For example, a 1 -cm tube section contains only $13 \mu\text{l}$ ($\sim \frac{1}{4}$ drop of water) and the assumption of an infinite liquid with a constant gas-saturation medium is unlikely to be appropriate, especially when other bubbles are also present in the neighborhood with the possibility of significant interbubble gas exchange. (3) Acoustic streaming caused by the bubble oscillation can increase the growth rate by refreshing the liquid immediately surrounding the bubble.²¹

Figure 8(a) also illustrates how bigger bubbles can form through the coalescence of smaller bubbles due to an attractive secondary radiation force.^{3,20} This occurs when the smaller bubbles are close together and have resonant sizes smaller than the acoustic wavelength.

Figure 8(b) shows a short sequence from a movie during the application of 28.8 kHz CW ultrasound with $P_A \approx 90$ kPa

that illustrates disintegration of a bubble into smaller microbubbles and subsequent recombination. Observations of many such sequences showed rapid bubble oscillation under the influence of primary and secondary radiation forces.³ It seemed likely that the recombination was aided by the attractive nature of the secondary radiation forces. The smaller bubbles produced by disintegration had a wide distribution of sizes and sometimes these eventually became attached to the tube and grew again to bigger bubbles by means of rectified diffusion.

3. Summary of bubble excitation and maintenance processes

In Fig. 9 the various processes just described for the excitation and maintenance steps are summarized. For the excitation step the sketch in (a) begins with the assumption that small fragmented bubbles are trapped on the tube wall. These bubbles may grow due to rectified diffusion. Once the bubble has grown to a certain size, the bubble can either disintegrate into smaller bubbles or it can emit microbubbles from the surface due to surface instability. Those smaller bubbles that become trapped on the wall can grow to bigger bubbles and so the entire process repeats itself. In the maintenance step illustrated in (b) the primary bubble activity is the emission of microbubbles due to surface instability. These microbubbles can be carried away by the flow. As noted earlier [see Fig. 7(b)], they are quite uniform in size, typically having diameters in the range 7 to $16 \mu\text{m}$.

4. Bubble flow characteristics

Figure 10 shows a series of the 2-D microbubble flow velocity vectors obtained by digitally subtracting consecutive movie frames at different flow rates. The movie frames were obtained when the microscope was operated at its maximum magnification. During the maintenance step, these microbubbles were emitted from the bubbles trapped on the wall due to the surface instability. The movie series were taken downstream in a region where the ultrasound field was small, so that secondary radiation forces and streaming should be negligible. Since the focus was at the center of the lumen and the image had a depth of field of $\sim 300 \mu\text{m}$, the microbubble velocity vectors that were measured should be near the plane that coincides with the tube axis and which is normal to the viewing direction. As a result, the velocity measurements should closely approximate the flow velocity profile. For example, Fig. 10(b) shows that the velocity of the microbubbles flowing near the tube center is ~ 8 mm/s, which is close to the maximum flow velocity of 8.4 mm/s for parabolic profile at 20 ml/h. It will be noted that the buoyancy force causes the 2-D velocity vectors to point upwards, as expected. The rise angle, θ_{rise} , defined as the angle between the 2-D bubble velocity and the axial flow velocity, is generally higher for lower flow rate.

5. Rise velocity

For small bubble at very small Reynolds numbers, the rise velocity, v_{rise} , is given by^{22,23}

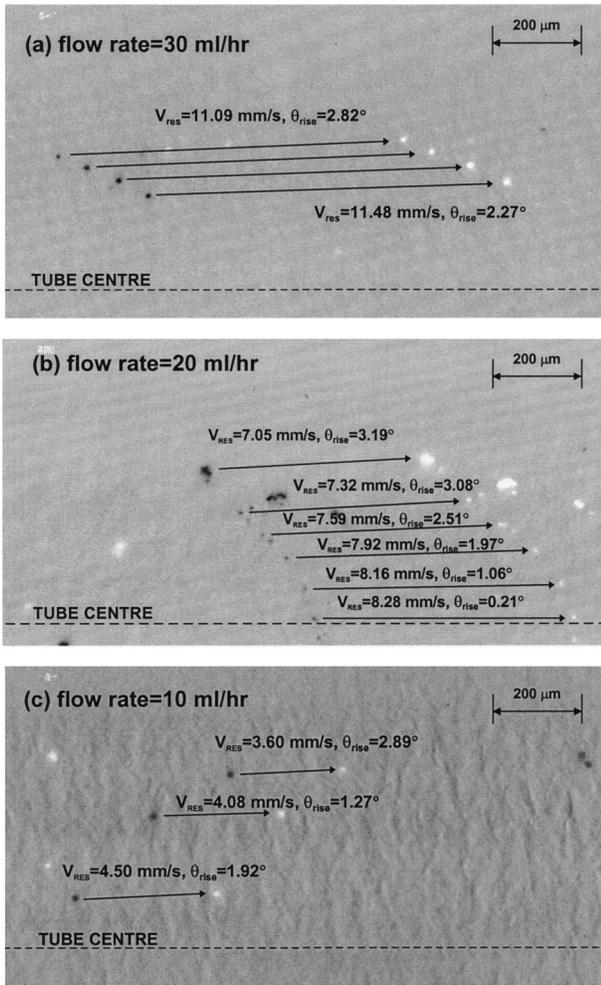


FIG. 10. The 2-D velocity vectors of flowing microbubbles obtained by digital subtraction of consecutive frames (15 frames/s) with flow rates at (a) 30 ml/h, (b) 20 ml/h, and (c) 10 ml/h.

$$V_{\text{rise}} = \frac{R_0^2 \rho g}{3\mu}, \quad (1)$$

where R_0 is the bubble radius, ρ and μ are the density and viscosity of the liquid and g is the acceleration of gravity. Figure 11 plots the measured rise velocities for a range of microbubble sizes together with the theoretical value for water at 20 °C [$\rho=998 \text{ kg/m}^3$ and $\mu=0.001002 \text{ kg/(m}\cdot\text{s)}$], as calculated from Eq. (1). It is evident that good agreement was obtained, which also supports our assertion that radiation forces in the measurement region were negligible. For the smaller microbubbles the accuracy in estimating the bubble diameter becomes rather large, as indicated by the horizontal bar.

6. Dissolution

Unlike contrast agents such as Alunex[®] for which the microbubbles are stabilized by a gas-impermeable shell, the microbubbles generated by ultrasound have no such shell to protect them from dissolving. In the absence of an acoustic field, the free spherical microbubble dissolution time in an undersaturated solution can be estimated from²³

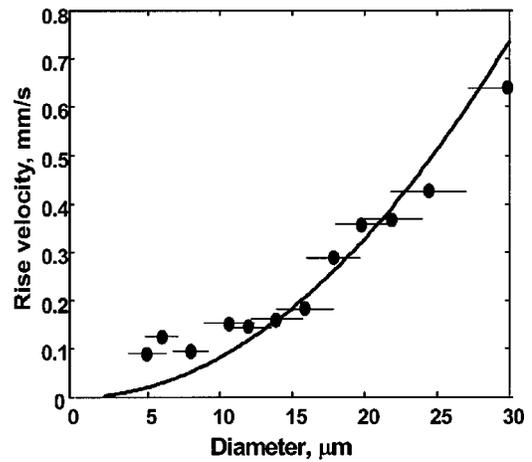


FIG. 11. Relationship between the rise velocity and the microbubble diameter. The solid line represents the theoretical rise velocity and the solid circles are the experimental measurements. Horizontal bars indicate the maximum possible measurement error.

$$t_d = \frac{R_0^2}{2\kappa(C_0/\rho)(1-C_i/C_0)}, \quad (2)$$

where κ is the diffusivity of gas in the liquid, C_i is the concentration of gas dissolved in the liquid and C_0 is the saturation concentration of gas dissolved in the liquid. The dissolution time increases with the dissolved gas concentration and as the square of bubble radius. In the absence of flow and acoustic pressure, the dissolution time for free microbubbles in water at a gas saturation level of 70% was determined by following a movie frame series at 10 frames until the microbubble dissolved completely in the tube. Figure 12 shows the experimental and theoretical dissolution time of free microbubbles. The theoretical curve was calculated from Eq. (2) assuming²³ an air bubble in water at 22 °C, $\kappa=2\times 10^{-9} \text{ m}^2/\text{s}$ and $C_0/\rho=0.02$. For smaller bubbles it can be seen that the experimental measurements agree very well with the theory. Previously reported measurement² of the effective lifetime for a flowing bubble population having

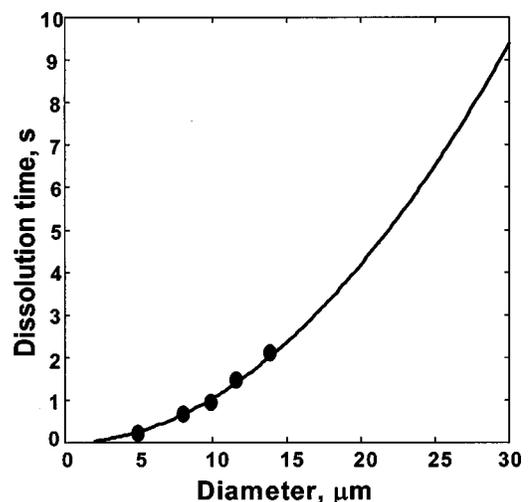


FIG. 12. Dissolution time of the microbubbles. The solid line indicates the theoretical dissolution time as a function of microbubble size and the solid circles are the experimentally measured dissolution time.

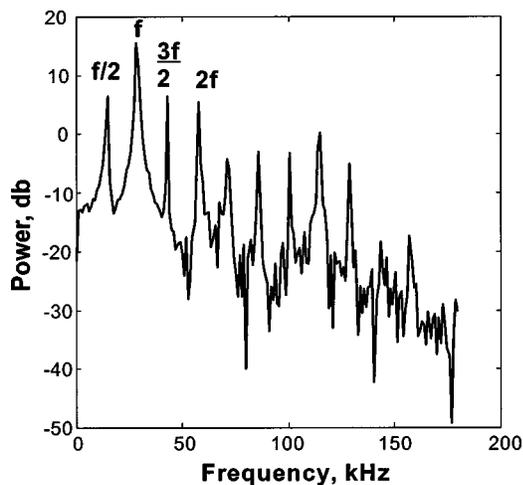


FIG. 13. Power spectrum obtained during the initiation step ($P_A \approx 400$ kPa) showing the sub and higher-harmonics of the pressure wave scattered from the transient cavitation cloud inside the tube.

a mean diameter of $11.8 \mu\text{m}$ using a Doppler-based transit-time method was found to be $\sim 1.6 \pm 0.7$ s, which conforms to that indicated by Fig. 12.

C. Acoustic emission detection

Figure 13 shows the bubble emission power spectra, as measured using the hydrophone system shown in Fig. 4, for $P_A \approx 400$ kPa. During the initiation phase there is a wide distribution of bubble sizes (Fig. 5, left panel) that are created by transient cavitation and the emission spectra contains both sub- and higher-harmonics of both fractional and integer order. These experimental observations are consistent with previous observations,^{15,19} which noted that the emission spectrum changes significantly as soon as transient cavitation starts. The presence of such a spectrum can be used as an indicator of successful bubble initiation.

Once a bubble population has been developed using the excitation step (Fig. 2), the application of CW ultrasound of sufficient intensity will cause the emission spectra to contain harmonic components. Shown in Fig. 14(a) are the power spectra in the absence and presence of bubbles using CW ultrasound (28.8 kHz, ~ 95 kPa). In addition, the dependence

of the harmonic amplitudes on the applied acoustic pressure is given in Fig. 14(b). The second harmonic is dominant and has an amplitude that is consistently 10–20 dB higher than that if the bubbles are absent. Furthermore, as the excitation pressure is increased, the second harmonic power increases at a much higher rate than the fundamental. Because the hydrophone was within the geometric field of the horn, the absence of higher harmonics when bubbles were not present indicates that the incident pressure waveform was very close to sinusoidal and therefore did not contribute to the measured spectrum in the presence of bubbles.

D. Doppler ultrasound measurements

Because there are relatively few microbubbles within the sample volume of our Doppler transducer, the Doppler power is expected to vary in a linear manner with the bubble number density. Each data point in Fig. 15 was obtained by using the imaging system to count the number of bubbles that passed a given cross-sectional plane over a 6-s interval and at the same time determining the total Doppler power over the same interval. It can be seen that the Doppler power is linearly proportional to the number of flowing microbubbles with a correlation coefficient of $r=0.985$. Even though there are relatively few microbubbles within the sample volume at a given instant of time, the signal-noise ratio for the 6-s averaging interval was greater than 20 dB. This is primarily because the backscattering coefficient of an air bubble is in the order of 10^{10} higher than that of a liquid-filled scatterer having the same volume.

For the proposed scheme for measuring CSF flow it would be useful to have a noninvasive method for determining whether bubble initiation and excitation had been successfully achieved. Three methods based on the use of ultrasound are possible. The first makes use of B-mode imaging in which the high bubble backscattering cross section produces a clear visual indication of the presence of bubbles, as has been demonstrated by Fowlkes *et al.*²⁴ for a dog urinary bladder. The second method uses power-mode Doppler imaging to detect the motion of a bolus of bubbles following their generation, as was demonstrated by Ivey *et al.*¹³ for detecting the presence of cavitation bubbles in the *in vivo*

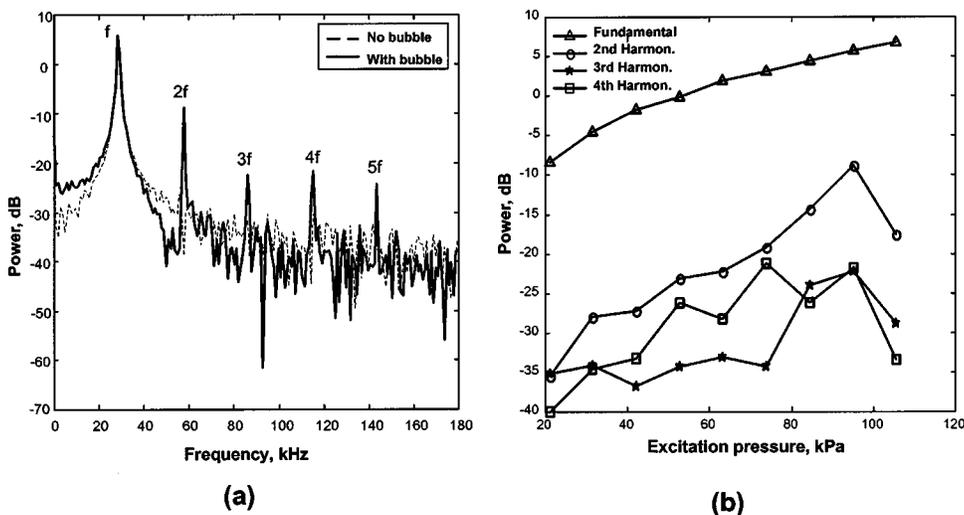


FIG. 14. Power spectra obtained with CW ultrasound (28.8 kHz, $P_A \approx 95$ kPa) following the initiation step. (a) In the absence and presence of bubbles. (b) Dependence of the harmonic amplitudes on the excitation pressure.

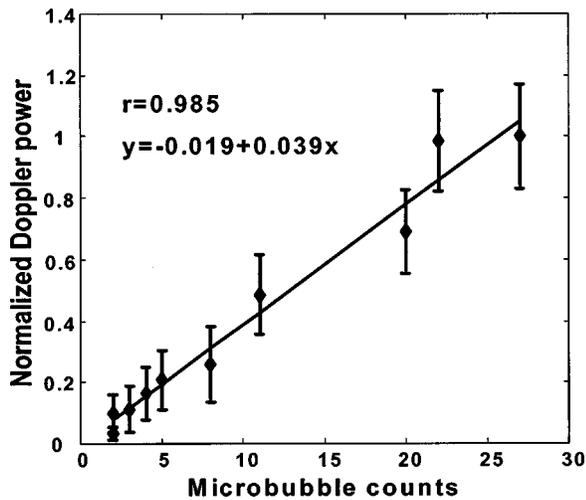


FIG. 15. Relationship between the normalized Doppler power and the number of microbubble counts (correlation coefficient of 0.985). The solid straight line shows the best linear fit.

canine abdominal aorta. The third method uses a measurement of the Doppler ultrasound flow spectrum *during* the bubble generation process. Figure 16(a) shows the acquired Doppler spectra in the absence of flow during the initiation step (28.8 kHz, ~400 kPa). For this measurement six initiation pulses of 200-ms duration were used and these were evenly spaced over a 60-s interval. The Doppler power was averaged over this interval. It is likely that the signal received in the presence of bubbles reflects their violent movement as well as the rapid bubble expansion which can be expected to produce Doppler frequencies up to and beyond the Nyquist frequency (7.5 kHz) of our system, resulting in aliasing. With successful bubble initiation, it is evident that the Doppler power increases significantly (20–40 dB) over a relatively wide frequency range. Figure 16(b) shows that during the excitation step the Doppler power is significantly higher than when the bubbles were absent, though the difference decreases with the Doppler frequency.

III. CONCLUDING DISCUSSION

This paper has described measurements and analysis of the process involved in generating and maintaining bubbles

in a small diameter silicone tube containing a flowing non-scattering fluid medium using 28.8kHz ultrasound. The bubble generation scheme, which corresponded to that previously proposed for Doppler ultrasound CSF shunt flow measurements, had a sequence of three steps, utilizing three different acoustic cavitation processes: transient, stable and cyclic. In the last step, microbubbles were excited from the surface of larger bubbles trapped on the tube inner wall.

The bubble imaging system has allowed us to better understand the physical processes involved in the three steps: these include the cavitation thresholds, size and density distribution of the trapped and free bubbles, growth, dissolution and flow characteristics of the flowing microbubbles. This information provides a useful basis for improving the bubble generation scheme. In addition, we have also been able to develop harmonic detection techniques that allow the presence of cavitation to be identified.

By studying the properties of the flowing microbubbles during the maintenance step, we have been able to explain some experimental results previously reported² in relation to Doppler flow measurements in a small diameter tube. For a Doppler probe angle of 30 degrees, we observed that the Doppler measured mean velocity was 25%–50% higher than the actual mean flow velocity² at ~1 mm/s. We believe that at low flow the effects of buoyancy become more significant, causing a (nonaxial) velocity component that increases the velocity component in the Doppler probe direction. For example, the rise velocity of a 30- μ m-diam bubble was found to be 0.63 mm/s, which was close to the mean flow velocity of 1 mm/s. This would result in a Doppler measured mean velocity ~40% higher than the actual mean flow velocity.

The acquired data for the bubble activity should be useful to improve the current bubble generation scheme and to explain previous observations of the Doppler spectra.² Many aspects of the detailed mechanisms involved in the initiation, excitation and maintenance steps are still poorly understood and require further investigation. Future studies could also consider how these ultrasonically generated microbubbles could serve as contrast agents for B-mode or Doppler ultrasound imaging purpose.

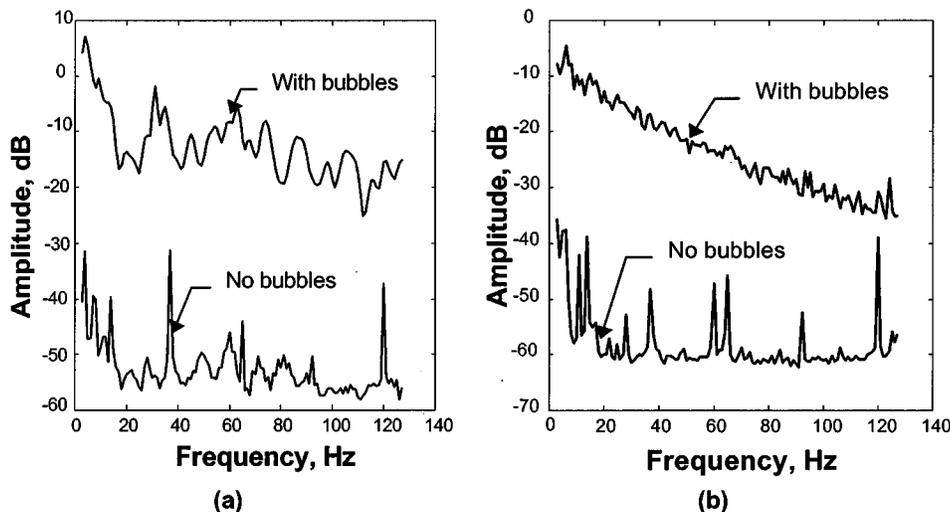


FIG. 16. Acquired Doppler spectra (a) with successful and unsuccessful bubble initiation and (b) with successful and unsuccessful bubble excitation.

ACKNOWLEDGMENTS

We are grateful for an initial grant from Hospital for Sick Children Foundation (Toronto), subsequent funding from Johnson & Johnson Company and Brain Child, and substantial financial support from a grant to RSCC by the Natural Sciences and Engineering Council of Canada. We would also like to thank the continuing support and advice of Dr. Jahan Tavakkoli and Dr. Howard Ginsberg. The loan of a calibrated hydrophone through the generosity of Dr. E. C. Everbach at Swarthmore College is also much appreciated.

- ¹K. W. Lam, P. Bascom, A. Karpelson, J. M. Drake, and R. S. C. Cobbold, "Ultrasound Bubble Excitation for Doppler Measurement of Cerebrospinal Fluid Shunt Flow," Amsterdam, IEEE Engineering in Medicine and Biology Society (EMBS), Paper 110, November 1996.
- ²K. W. Lam, J. M. Drake, and R. S. C. Cobbold, "Non-invasive cerebrospinal fluid shunt flow measurement by Doppler ultrasound using ultrasonically excited bubbles: A feasibility study," *Ultrasound Med. Biol.* **25**, 371–389 (1999).
- ³T. G. Leighton, *The Acoustic Bubble* (Academic, London, 1994).
- ⁴L. A. Crum and G. M. Hansen, "Generalized equations for rectified diffusion," *J. Acoust. Soc. Am.* **72**, 1586–1592 (1982).
- ⁵E. A. Neppiras and E. E. Fill, "A cyclic cavitation process," *J. Acoust. Soc. Am.* **46**, 1264–1271 (1969).
- ⁶R. E. Apfel, "Acoustic cavitation prediction," *J. Acoust. Soc. Am.* **69**, 1624–1633 (1981).
- ⁷R. Y. Nishi, "Ultrasonic detection of bubbles with Doppler flow transducers," *Ultrasonics* **10**, 173–179 (1972).
- ⁸B. A. Hills and D. C. Grulke, "Evaluation of ultrasonic bubble detectors in vitro using calibrated microbubbles at selected velocities," *Ultrasonics* **13**, 181–184 (1975).
- ⁹M. F. Gillis, M. T. Karagianes, and P. L. Peterson, "Bends: Detection of circulating gas emboli with external sensor," *Science* **161**, 579–580 (1968).
- ¹⁰D. L. Miller, A. R. Williams, and D. R. Gross, "Characterization of cavitation in a flow-through exposure chamber by means of a resonant bubble detector," *Ultrasonics* **22**, 224–230 (1984).
- ¹¹D. R. Gross, D. L. Miller, and A. R. Williams, "A search for ultrasonic cavitation within the canine cardiovascular system," *Ultrasound Med. Biol.* **11**, 85–97 (1985).
- ¹²A. R. Williams, M. Delius, and W. Schwarze, "Investigation of cavitation in flowing media by lithotripter shock waves both *in vitro* and *in vivo*," *Ultrasound Med. Biol.* **15**, 53–60 (1989).
- ¹³J. A. Ivey, E. A. Gardner, J. B. Fowlkes, and J. M. Rubin, "Acoustic generation of intra-arterial contrast boluses," *Ultrasound Med. Biol.* **21**, 757–767 (1995).
- ¹⁴F. R. Young, *Cavitation* (McGraw-Hill, Oxford, 1989).
- ¹⁵E. A. Neppiras, "Acoustic Cavitation," *Physics Reports* **3**, 159–251 (1980).
- ¹⁶W. Lauterborn, "Numerical investigation of nonlinear oscillations of gas bubbles in liquids," *J. Acoust. Soc. Am.* **59**, 283–293 (1976).
- ¹⁷A. I. Eller and L. A. Crum, "Instability of the motion of a pulsating bubble in a sound field," *J. Acoust. Soc. Am.* **47**, 762–767 (1970).
- ¹⁸W. J. Galloway, "An experimental study of acoustically induced cavitation in liquids," *J. Acoust. Soc. Am.* **26**, 849–857 (1954).
- ¹⁹F. R. Young, *Cavitation* (McGraw-Hill, Oxford, 1989), pp. 50, 121–123.
- ²⁰P. A. Dayton, K. E. Morgan, A. L. Klibanov, G. Brandenburger, K. R. Nightingale, and K. W. Ferrara, "A preliminary evaluation of the effects of primary and secondary radiation forces on acoustic contrast agents," *IEEE Trans. Ultrason. Ferroelectr. Freq.* **44**, 1264–1277 (1997).
- ²¹R. K. Gould, "Rectified diffusion in an 11-kHz sound field," *J. Acoust. Soc. Am.* **56**, 1740–1746 (1974).
- ²²G. K. Batchelor, *An Introduction to Fluid Mechanics* (Cambridge U.P., Cambridge, 1967), pp. 237–367.
- ²³P. S. Epstein and M. S. Plesset, "On the stability of gas bubbles in liquid-gas solutions," *J. Chem. Phys.* **18**, 1505–1509 (1950).
- ²⁴J. B. Fowlkes, P. L. Carson, E. H. Chiang, and J. M. Rubin, "Acoustic generation of bubbles in excised canine urinary bladders," *J. Acoust. Soc. Am.* **89**, 2740–2744 (1991).

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.

The force between two parallel rigid plates due to the radiation pressure of phonons

O. Bschorr

Aeroacoustics, Munich D-81679, Germany

(Received 5 October 1998; revised 24 May 1999; accepted 16 September 1999)

Larraza *et al.* measured the force between two plates in an external high-intensity sound field. In analogy to the conventional “Casimir effect” it was called the “acoustical Casimir effect.” The purpose of this paper is to derive the force between two plates without external sound fields, that is, induced only by thermal-pressure fluctuations, i.e., by phonons in air. For this reason, the approximated Debye method will be used. Thereby, air will be described as a phonon field and the number of modes in free air and between the plates with distance d will be estimated. A deficiency of modes in the d -space results in a deficiency of radiation pressure and therefore leads to an attractive plate pressure $f = \pi kT/18 d^3$ Pa. In ambient air and for a plate distance $d = 10^{-5}$ m the pressure $f = 0.7 \cdot 10^{-6}$ Pa follows. In comparison to the Casimir forces due to the zero-point field of the vacuum, the pressure is of the same size and can also be measured. © 1999 Acoustical Society of America. [S0001-4966(99)07512-8]

PACS numbers: 43.25.Qp [MAB]

INTRODUCTION

Electromagnetical vacuum fluctuations induce electrical dipoles on a conducting plate. Between two interacting plates an attractive force results. In 1948 this mechanism was theoretically predicted by Casimir.¹ The Casimir forces are very small, but in the following decades they were confirmed experimentally. Larraza *et al.*² transferred the Casimir experiment into acoustics. They measured the force between two plates in an external sound field. They used plate distances between 1.2 to 60 mm and a sound intensity of 133.5 dB with a bandwidth from 5 to 20 kHz. In the space between the two plates, lower-frequency modes are suppressed and therefore an attractive force results. The force is also repulsive when the distance between the plates is comparable to the half-wavelength associated with the lower edge of the frequency band. Because of the similarity of the equipment, Larraza *et al.* called this the “acoustic Casimir effect.”³ In this paper, the acoustic Casimir effect without an external sound field will be investigated. Also, in a gas at rest there are thermodynamical pressure fluctuations and Brownian motions. For very small plate distances an attractive force between the two plates can be expected.

I. MODEL

For the estimation of the acoustical Casimir effect in gaseous media, a semiclassical method is used, with which

Debye determined the energy density of phonons in solids. Additionally, the radiation pressure of the phonons is taken into account. Therefore, up to the Debye frequency ν_B the gas is considered as an elastic medium transmitting only longitudinal waves with the sound velocity c (in contrast to Debye’s solids there are no translational degrees of freedom). The number of eigenmodes dn per unit volume in the gaseous continuum within the frequency interval $d\nu$ is

$$dn = 4\pi \frac{v^2 d\nu}{c^3}. \quad (1)$$

According to Debye, the total number of modes is given by the Loschmidt number L —the number of molecules per unit volume. A three-dimensional elastic system with L masses has $3L$ eigenmodes. With $\int dn = 3L$ the border frequency ν_B results as

$$\nu_B = c \sqrt[3]{\frac{9L}{4\pi}}. \quad (2)$$

At a gas temperature T the energy content of an eigenmode (phonon) is $a = h\nu / (\exp h\nu/kT - 1)$ and so the total phonon energy is given by $u = \int a dn$ (k =Boltzmann constant, h =Planck quantum constant, $\Theta = h\nu_B/k$ =Debye temperature)

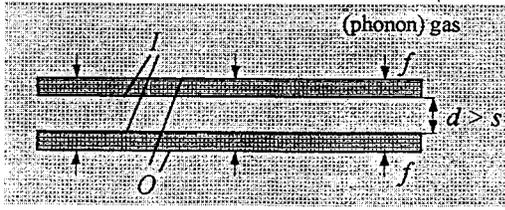


FIG. 1. Two rigid plates in a distance d within a gas. The gas considered as a phonon field gives an attractive pressure f between the plates (acoustic Casimir effect).

$$u = \int_0^{\nu_B} \frac{h\nu}{\exp(h\nu/kT) - 1} \frac{4\pi\nu^2 d\nu}{c^3}$$

$$= \begin{cases} 3LkT \left(1 - \frac{3}{8} \frac{\Theta}{T} + \dots \right) & \text{for } T \gg \Theta, \\ \frac{3}{5} \pi^4 LkT \left(\frac{T}{\Theta} \right)^3 & \text{for } T \ll \Theta. \end{cases} \quad (3)$$

Compared to the room temperature T , the Debye temperature Θ of gases (atmospheric air: $\Theta=4.5$ K) is much lower, so that with $\Theta/T \rightarrow 0$ the energy of the phonons per unit volume follows:

$$u = 3LkT. \quad (5)$$

An isotropic phonon field with energy density u has a sound intensity $i = uc/4\pi$. A wave with intensity i vertically reflected from a surface causes a radiation pressure $p = 2i/c$. The summarized half-sphere radiation pressure is:

$$p = \frac{1}{3}u. \quad (6)$$

This pressure corresponds to the macroscopic gas pressure. With the introduction of the molecular weight m , the gas density $\rho = Lm$ and the gas constant $R = k/m$ the phonon radiation pressure p can be expressed as

$$p = LkT = Lm \frac{k}{m} T = \rho RT. \quad (7)$$

For $\Theta/T \rightarrow 0$ this result is identical with the thermodynamical gas pressure. To show existence and size of the acoustical Casimir effect, the simplified Debye model is sufficient and strict quantum calculation of the Debye frequency ν_B —and also of the cutoff frequency ν_C [Eq. (8)]—will be neglected.

II. ACOUSTIC CASIMIR EFFECT

In a homogeneous gas, two rigid plane plates are spaced with parallel distance d (Fig. 1). The dimensions of the plates are great compared to d . The free-flight length s of the gas molecules ought to be small compared to the plate distance, i.e., $s < d$. In contrast to Larazza the gas is at rest and no external sound fields exist. The plates are assumed as perfect

acoustic reflectors and absolutely rigid, so that no surface vibrations by the gas phonons occur (in the electric case, this corresponds to nonconductive plates). Because of the thermodynamic balance and because of the great difference of impedances and wave velocities in the gas and in a metal plate, these assumptions are not very stringent. With the plate distance d a cutoff-wavelength λ_C and a cutoff frequency ν_C will be introduced

$$\lambda_C = 2d, \quad \nu_C = c/\lambda_C = c/2d. \quad (8)$$

For wavelengths $\lambda < \lambda_C$ in the space between the plates, free-field condition can be supposed with an undisturbed three-dimensional mode field. The spectral energy and therefore also the radiation pressure on the inner (I) and outer (O) side of the plates are the same. In the force balance the radiation pressure is compensated in this range and it is not necessary to calculate it. In the range $\lambda > \lambda_C$ on the outer side (O) of the plate, a full-mode field exists, but on the inner side (I) modes up to λ_B do not appear. With the corresponding cutoff frequency ν_C , the difference of the phonon energy $u(d)$ at the inner and outer plate side for $\Theta/T \rightarrow 0$ and $\nu_C < \nu_B$ is

$$u(d) = \int_0^{\nu_C} \frac{h\nu}{\exp(h\nu/kT) - 1} \frac{4\pi\nu^2 d\nu}{c^3} = \frac{\pi kT}{6d^3}. \quad (9)$$

With Eq. (6) the resulting radiation pressure $f = u(d)/3$ causes attraction of the two plates. According to the notation of Larazza *et al.*, this effect is called ‘‘acoustic Casimir effect.’’ The pressure f between the two attracting plates is

$$f = \frac{\pi kT}{18d^3}. \quad (10)$$

Surprisingly, the pressure f only depends on the gas temperature T and not on density, pressure, and the kind of gas. The pressure f is extremely small. With a plate distance $d = 10^{-5}$ m and the ambient temperature $T = 288$ K, the acoustic Casimir pressure results $f = 0.7 \cdot 10^{-6}$ Pa. The equipment for measuring the acoustic as well as the electromagnetic Casimir effect is similar. With the electromagnetic Casimir pressure $g = \pi h c_L / 480 d^4$, the pressure ratio f/g is ($c_L = \text{light velocity}$)

$$\frac{f}{g} = \frac{80}{3} \frac{kT}{hc_L} d. \quad (11)$$

At the plate distance $d = 1.8 \cdot 10^{-6}$ m both effects show the same resulting pressure. Therefore, it will be possible to also measure the acoustic Casimir effect induced by the thermic pressure fluctuations in a gas.

¹H. B. G. Casimir, ‘‘On the attraction between two perfectly conducting plates,’’ *Proc. K. Ned. Akad. Wet.* **51**, 793–796 (1948).

²A. Larazza, ‘‘The force between two parallel plates due to the radiation pressure of broad band noise: An acoustic Casimir effect,’’ *J. Acoust. Soc. Am.* **103**, 2267–2272 (1998).

³A. Larazza, ‘‘An acoustic Casimir effect,’’ *Phys. Lett. A* **248**, 151–155 (1998).

Structural radiation mode sensing for active control of sound radiation into enclosed spaces

Ben S. Cazzolato and Colin H. Hansen

Department of Mechanical Engineering, University of Adelaide, SA, 5005, Australia

(Received 3 March 1999; accepted for publication 27 August 1999)

In the recent article by Cazzolato and Hansen [J. Acoust. Soc. Am. **104**, 2878–2889 (1998)] it was shown that it is possible to derive for a structure some set of surface velocity distributions, referred to as radiation modes, which are orthogonal in terms of their contributions to the acoustic potential energy of a coupled cavity. The technique used an orthonormal decomposition to derive an expression for the radiation modes which was based on prior work for free-field sound radiation. It will be shown in the following letter that for the special case involving the calculation of global internal potential energy it is possible to use a simple approach which requires no orthonormal decomposition since the expression for the global potential energy is already in a form that can be easily diagonalized. © 1999 Acoustical Society of America. [S0001-4966(99)04112-0]

PACS numbers: 43.50.Ki [MRS]

INTRODUCTION

In the recently published article by Cazzolato and Hansen¹ an expression for the structural radiation modes orthogonal to the acoustic potential energy in an enclosure was derived. The approach used an orthonormal decomposition of the acoustic error weighting matrix, $\mathbf{\Pi}$, to obtain the eigenvectors, \mathbf{U} , and eigenvalues, \mathbf{S} , of the matrix (i.e., $\mathbf{\Pi} = \mathbf{Z}_a^H \mathbf{\Lambda} \mathbf{Z}_a = \mathbf{U} \mathbf{S} \mathbf{U}^T$, where \mathbf{Z}_a is the structural-acoustic transfer function matrix and $\mathbf{\Lambda}$ is a diagonal weighting matrix). The approach taken was primarily because of precedence, since the technique had been used in the past for radiation into the free space. However, on closer inspection of the governing equations it can be shown that if global control of the acoustic space is the objective of an active noise control system, then it is not necessary to decompose the radiation matrix. By collecting the appropriate terms that comprise the radiation matrix, a pseudo-eigenvector–eigenvalue decomposition is obtained, and while not truly orthonormal, it does result in an adequate approximation of the exact orthogonal data set.

As already stated, the following approach is only applicable to radiation modes which are orthogonal in terms of their contributions to the total acoustic potential energy in the acoustic space, to cases for which the modal density is low and the flexible structure forms a large part of the bounding surface of the structure. If the latter two conditions do not hold, then at high frequencies the internal radiation mode shapes degenerate to approximately the free field radiation mode shapes.² This is because the assumption (used in the current simplification) that the acoustic mode shapes integrated over the radiating surface are orthogonal, no longer holds. As mentioned in passing by Cazzolato and Hansen,¹ it is possible to derive a set of radiation modes which are orthogonal in terms of their contributions to the potential energy in some subspace of the interior cavity, such as the space around a passenger's head. The alternative approach which is presented here is not suitable for such subspaces and the previous formulation¹ must be used.

I. PREVIOUS FORMULATION

In the paper by Cazzolato and Hansen,¹ the theory of sound transmission through a structure into a contiguous cavity was developed with the transmitted sound field derived in terms of radiation modes. Using a modal-interaction approach to the solution of coupled problems, the response of the structure was modeled in terms of its *in vacuo* mode shape functions and the response of the enclosed acoustic space was described in terms of the rigid-wall mode shape functions.³ The response of the coupled system was then determined by solving the modal formulation of the Kirchhoff–Helmholtz integral equation. The following two sections have been taken directly from Cazzolato and Hansen¹ for the sake of completeness.

A. Global error criteria

An appropriate global error criterion for controlling the sound transmission into a coupled enclosure is the total time-averaged frequency dependent acoustic potential energy, $E_p(\omega)$, in the enclosure⁴

$$E_p(\omega) = \frac{1}{4\rho_0 c_0^2} \int_V |p(\vec{\mathbf{r}}, \omega)|^2 d\vec{\mathbf{r}}, \quad (1)$$

where $p(\vec{\mathbf{r}}, \omega)$ is the acoustic pressure amplitude at some location $\vec{\mathbf{r}}$ in the enclosure, ρ_0 is the density of the acoustic fluid (air), c_0 is the speed of sound in the fluid and V is the volume over which the integral is evaluated. The frequency dependence, ω , is assumed in the following analysis but this parameter will be omitted in the following equations for the sake of brevity. Using the modal interaction approach to the problem,³ the acoustic pressure at any location within the cavity is expressed as an infinite summation of the product of rigid-wall acoustic mode shape functions, ϕ_i , and the modal pressure amplitudes, p_i , of the cavity

$$p(\vec{\mathbf{r}}) = \sum_{i=1}^{\infty} p_i \phi_i(\vec{\mathbf{r}}). \quad (2)$$

The modal expansion for the acoustic potential energy evaluated over n_a acoustic modes is then given by

$$E_p = \mathbf{p}^H \mathbf{\Lambda} \mathbf{p}, \quad (3)$$

where \mathbf{p} is the $(n_a \times 1)$ vector of acoustic modal amplitudes and $\mathbf{\Lambda}$ is a $(n_a \times n_a)$ diagonal weighting matrix, the diagonal terms of which are

$$\Lambda_{ii} = \frac{\Lambda_i}{4\rho_0 c_0^2}, \quad (4)$$

where Λ_i is the modal volume of the i th cavity mode, defined as the volume integration of the square of the mode shape function,

$$\Lambda_i = \int_V \phi_i^2(\vec{\mathbf{r}}) dV(\vec{\mathbf{r}}). \quad (5)$$

The pressure modal amplitudes, \mathbf{p} , within the cavity, arising from the vibration of the structure are given by the product of the $(n_s \times 1)$ structural modal velocity vector, \mathbf{v} , and the $(n_a \times n_s)$ modal structural-acoustic radiation transfer function matrix, \mathbf{Z}_a ,⁵

$$\mathbf{p} = \mathbf{Z}_a \mathbf{v}. \quad (6)$$

The (l, i) th element of the radiation transfer function matrix \mathbf{Z}_a is the pressure amplitude of the acoustic mode l generated as a result of structural mode i vibrating with unit velocity amplitude. Substituting Eq. (6) into Eq. (3) gives an expression for the acoustic potential energy with respect to the normal structural vibration,

$$E_p = \mathbf{v}^H \mathbf{\Pi} \mathbf{v}, \quad (7)$$

where the error weighting matrix $\mathbf{\Pi}$ is given by

$$\mathbf{\Pi} = \mathbf{Z}_a^H \mathbf{\Lambda} \mathbf{Z}_a. \quad (8)$$

It should be noted that the error weighting matrix $\mathbf{\Pi}$ is not necessarily diagonal which implies that the normal structural modes are not orthogonal contributors to the interior acoustic pressure field. It is for this reason that minimization of the modal amplitudes of the individual structural modes (or kinetic energy) will not necessarily reduce the total sound power transmission.

B. Diagonalization of the error criteria

Since $\mathbf{\Pi}$ is real symmetric it may be diagonalized (using a singular value decomposition) to yield the orthonormal transformation:

$$\mathbf{\Pi} = \mathbf{U} \mathbf{S} \mathbf{U}^T, \quad (9)$$

where the unitary matrix \mathbf{U} is the (real) orthonormal transformation matrix representing the eigenvector matrix of $\mathbf{\Pi}$ and the (real) diagonal matrix \mathbf{S} contains the eigenvalues (singular values) of $\mathbf{\Pi}$. The physical significance of the eigenvectors and eigenvalues is interesting. The eigenvalue can be considered a radiation efficiency (or coupling strength⁶) and the associated eigenvector gives the level of participation of each normal structural mode to the radiation mode; thus it indicates the modal transmission path.⁶

Substituting the orthonormal expansion of Eq. (9) into Eq. (7) results in an expression for the potential energy of the cavity as a function of an orthogonal radiation mode set,

$$E_p = \mathbf{v}^H \mathbf{U} \mathbf{S} \mathbf{U}^T \mathbf{v} = \mathbf{w}^H \mathbf{S} \mathbf{w}, \quad (10)$$

where the elements of \mathbf{w} are the velocity amplitudes of the radiation modes defined by

$$\mathbf{w} = \mathbf{U}^T \mathbf{v}. \quad (11)$$

Equation (11) demonstrates that each radiation mode is made up of a linear combination of the normal structural modes, the ratio of which is defined by the eigenvector matrix \mathbf{U} . As the eigenvalue matrix, \mathbf{S} , is diagonal, Eq. (10) may be written as follows,

$$E_p = \sum_{i=1}^n s_i |w_i|^2, \quad (12)$$

where s_i are the diagonal elements of the eigenvalue matrix \mathbf{S} and w_i are the modal amplitudes of the individual radiation modes given by Eq. (11).

The potential energy contribution from any radiation mode is equal to the square of its amplitude multiplied by the corresponding eigenvalue. The radiation modes are therefore independent (orthogonal) contributors to the potential energy and the potential energy is directly reduced by reducing the amplitude of any of the radiation modes. As mentioned previously, the normal structural modes are not orthogonal radiators since the potential energy arising from one structural mode depends on the amplitudes of the other structural modes. The orthogonality of the radiation modes is important for active control purposes as it guarantees that the potential energy will be reduced if the amplitude of any radiation mode is reduced.⁷

II. ALTERNATIVE FORMULATION

It will be shown here that the previous approach to diagonalize the error weighting matrix $\mathbf{\Pi}$ used in Sec. II A via the orthonormal transformation was unnecessary for the cost function being global potential energy. This is because in this case the expression used to define $\mathbf{\Pi}$ was already written in terms of a diagonal matrix $\mathbf{\Lambda}$ and a fully populated matrix \mathbf{Z}_a and its Hermitian transpose. In situations where the control objective is not global but rather a subspace, the following approach cannot be used because the error weighting matrix $\mathbf{\Pi}$ does not have an inner matrix which is diagonal. For example, the error weighting matrix $\mathbf{\Pi}$ for minimizing the sum of the squared pressures over some subspace is given by⁸

$$\mathbf{\Pi} = \mathbf{Z}_a^H \mathbf{Z}_w \mathbf{Z}_a, \quad (13)$$

where $\mathbf{Z}_w = \mathbf{\Phi}_e^* \mathbf{\Phi}_e^T$ and $\mathbf{\Phi}_e$ is the mode shape matrix at the error sensor locations within the subspace. Clearly \mathbf{Z}_w is not diagonal (unlike $\mathbf{\Lambda}$) but fully populated and therefore it is necessary to use the approach in Sec. II A.

The reformulation of the radiation modes orthogonal to the internal potential energy will now be presented. The interior acoustic potential energy is given by

$$E_p = \mathbf{v}^H \mathbf{Z}_a^H \mathbf{\Lambda} \mathbf{Z}_a \mathbf{v}, \quad (14)$$

where⁵

$$Z_a(l,i) = \frac{j\rho_0 S \omega}{\Lambda_l(\kappa_l^2 + j\eta_{a_l}\kappa_l k - k^2)} B_{l,i}, \quad (15)$$

where $B_{l,i}$ is the (l,i) th element of the $(n_a \times n_s)$ nondimensional coupling coefficient matrix, \mathbf{B} ,⁵ κ_l and η_{a_l} are the wave number and modal loss factor of the l th acoustic mode, respectively, and S is the total surface area of the bounding structure. Now \mathbf{Z}_a can be written in matrix form,

$$\mathbf{Z}_a = \mathbf{Y}\mathbf{B}, \quad (16)$$

where \mathbf{Y} is the $(n_a \times n_a)$ diagonal acoustic resonance matrix whose elements are given by

$$Y_{l,l} = \frac{j\rho_0 S \omega}{\Lambda_l(\kappa_l^2 + j\eta_{a_l}\kappa_l k - k^2)}. \quad (17)$$

Therefore, the potential energy may be expressed as

$$E_p = \mathbf{v}^H \mathbf{B}^H \mathbf{Y}^* \Lambda \mathbf{Y} \mathbf{B} \mathbf{v} \quad (18)$$

or

$$E_p = \mathbf{y}^H \mathbf{\Omega} \mathbf{y}, \quad (19)$$

where \mathbf{y} is the $(n_a \times 1)$ modal amplitude column vector of the radiation modes given by

$$\mathbf{y} = \mathbf{B} \mathbf{v} \quad (20)$$

and the $(n_a \times n_a)$ diagonal frequency-dependent weighting matrix, $\mathbf{\Omega}$, is given by

$$\mathbf{\Omega} = \mathbf{Y}^* \Lambda \mathbf{Y}. \quad (21)$$

Evaluating the diagonal weighting matrix, the elements are given by

$$\Omega_{ll} = \frac{\rho_0 c (S k)^2}{4 \Lambda_l ((\kappa_l^2 - k^2)^2 + (\eta_{a_l} \kappa_l k)^2)}. \quad (22)$$

It is clear that Eq. (19) is the same format as Eq. (10) with a fully-populated participation matrix and a diagonal weighting matrix. The radiation efficiency filters used by Cazzolato and Hansen (Ref. 1, Fig. 1) to weight the modal amplitudes to provide the inputs to the active noise control system are therefore given by the square root of the diagonal weighting matrix, $\mathbf{\Omega}$, which is equal to the magnitude of the product of the $(n_a \times n_a)$ diagonal frequency-dependent acoustic resonance matrix, \mathbf{Y} , and the square root of the modal volume matrix, $\mathbf{\Lambda}$. By induction, it is possible to define a corresponding mode shape matrix

$$\mathbf{\Xi} = \mathbf{\Psi} \mathbf{B}^T, \quad (23)$$

where $\mathbf{\Psi}$ is the structural mode shape matrix.

Premultiplying Eq. (23) by $\mathbf{\Psi}^T$ and integrating over the surface of the structure gives

$$\frac{1}{S} \int_s \mathbf{\Psi}^T(\vec{\mathbf{x}}) \mathbf{\Xi}(\vec{\mathbf{x}}) dS(\vec{\mathbf{x}}) = \frac{1}{S} \int_s \mathbf{\Psi}^T(\vec{\mathbf{x}}) \mathbf{\Psi}(\vec{\mathbf{x}}) \mathbf{B}^T dS(\vec{\mathbf{x}}), \quad (24)$$

and, using the principle of modal orthogonality, the following expression is obtained:

$$\frac{1}{S} \int_s \mathbf{\Psi}^T(\vec{\mathbf{x}}) \mathbf{\Xi}(\vec{\mathbf{x}}) dS(\vec{\mathbf{x}}) = \mathbf{M} \mathbf{B}^T, \quad (25)$$

where \mathbf{M} is the $(n_s \times n_s)$ diagonal matrix with diagonal elements given by

$$M_i = \frac{1}{S} \int_s \Psi_i^2(\vec{\mathbf{x}}) dS(\vec{\mathbf{x}}). \quad (26)$$

The left-hand term of Eq. (25) is the same as the expression for the nondimensional coupling coefficient matrix ($\mathbf{B}^T = (1/S) \int_s \mathbf{\Psi}^T(\vec{\mathbf{x}}) \mathbf{\Phi}(\vec{\mathbf{x}}) dS(\vec{\mathbf{x}})$) with the exception that the mode shape matrix of the radiation mode has been used in place of the acoustic mode shape matrix corresponding to the acoustic mode shape at the enclosure boundary. Therefore it follows that the radiation mode shape matrix is identical to the acoustic mode shape matrix in which column i scaled by some scalar term M_i , i.e.,

$$\mathbf{\Xi}(\vec{\mathbf{x}}) = \mathbf{\Phi}(\vec{\mathbf{x}}) \mathbf{M}, \quad (27)$$

cf. the same expression in terms of the structural mode shapes $\mathbf{\Xi}(\vec{\mathbf{x}}) = \mathbf{\Psi}(\vec{\mathbf{x}}) \mathbf{U}$.¹ It should be noted that since the mode shapes for the current formulation obviously do not vary with frequency, it is only appropriate to compare this current formulation with that of the ‘‘fixed-shape’’ radiation modes presented in Sec. II of the previous paper.

The approach just described is only suited to low frequencies where the modal density of the acoustic system is low since this ensures that the rows of the \mathbf{B} matrix are unique (column-orthogonal). As the number of the acoustic modes is increased, the likelihood of the acoustic mode shapes across the vibrating surface being orthogonal decreases. When nonorthogonality occurs, the advantage of this current approach begins to break down. To ensure uniqueness, it is possible to collect all the acoustic modes which have the same surface pressure pattern into a ‘‘single’’ radiation mode. This results in removal of the redundant line in the \mathbf{B} matrix and adds the corresponding terms in the diagonal weighting matrix $\mathbf{\Omega}$. The SVD approach has the advantage that this occurs automatically. It has been shown numerically and experimentally⁸ that, for low frequencies, the two approaches for calculating the radiation mode shapes lead to identical levels of control. This has been shown not to be the case at high frequencies,² especially when the radiating structure is small compared to the bounding surface of the cavity, which is when the internal radiation modes shapes degenerate to approximately the free field radiation mode shapes. This is because the acoustic response in the cavity becomes diffuse and can no longer be considered modal. In this situation \mathbf{B} is no longer column-orthogonal and therefore a SVD is necessary to orthogonalize the expression for the radiation matrix.

The current formulation is not only applicable to active noise control but has important implications for passive control of sound transmission into cavities. This shows that when attempting to minimize the sound transmission into cavities it is just as important to have an understanding of the dynamics of the receiving space as an understanding of the dynamics of the exciting structure. Dynamic absorbers and

co-located sensor/actuator pairs act to increase the impedance the structure “sees” at the mount point. Therefore, using the acoustic mode shapes to guide placement of such devices would likely achieve good results very quickly without having to analyze the dynamics of the structure. Obviously further refinement and optimization would have to take into consideration the dynamics of both the structure and the cavity.

III. CONCLUSION

The results presented by Cazzolato and Hansen¹ still hold since no new assumptions have been presented. The advantage of the current approach is that there is no need for the SVD to derive the mode shape matrices of the radiation modes contributing orthogonally to the global potential energy of the enclosure, which not only simplifies the analysis, but also decreases computation times. Note, however, that the approach outlined here is not suitable for cases where the cost function is the potential energy in a subspace of the enclosure. In this case the analysis presented previously in Ref. 1 must be used.

The approach presented here has important implications for the design of active control systems using radiation modal control. Only the dynamics of the cavity are required to design the control system. The radiation mode shapes are

identical in shape to the acoustic mode shapes of the cavity, and the radiation efficiencies of the radiation modes can be easily derived from the cavity resonance terms. Therefore, the modal sensor shapes need to be identical to the acoustic shapes at the enclosure boundary and the frequency weighting (radiation efficiency) filters need to emulate the modal interface coupling that occurs between the structure and the cavity to enable a successful active noise control system to be implemented.

¹B. Cazzolato and C. Hansen, “Active control of sound transmission using structural error sensing,” *J. Acoust. Soc. Am.* **104**, 2878–2889 (1998).

²M. Johnson, “Active control of sound transmission,” Ph.D. dissertation, The University of Southampton, 1996.

³F. Fahy, *Sound and Structural Vibration: Radiation, Transmission, and Response* (Academic, London, 1985).

⁴P. Nelson, A. Curtis, S. Elliott, and A. Bullmore, “The active minimization of harmonic enclosed sound fields, part I: Theory,” *J. Sound Vib.* **117**(1), 13 (1987).

⁵S. Snyder and C. Hansen, “The design of systems to actively control periodic sound transmission into enclosed spaces, Part I: Analytical models,” *J. Sound Vib.* **170**, 433–449 (1994).

⁶F. Bessac, L. Gagliardini, and J.-L. Guyader, “Coupling eigenvalues and eigenvectors: A tool for investigating the vibroacoustic behaviour of coupled vibrating systems,” *J. Sound Vib.* **191**, 881–899 (1996).

⁷M. Johnson and S. Elliott, “Active control of sound radiation using volume velocity cancellation,” *J. Acoust. Soc. Am.* **98**, 2174–2186 (1995).

⁸B. Cazzolato, “Sensing systems for active control of sound transmission into cavities,” Ph.D. dissertation, The University of Adelaide, 1999.

Voice onset times and burst frequencies of four velar stop consonants in Gujarati

Manish K. Rami, Joseph Kalinowski, Andrew Stuart, and Michael P. Rastatter^{a)}

*Department of Communication Sciences and Disorders, East Carolina University, Greenville,
North Carolina 27858-4353*

(Received 12 July 1999; revised 29 July 1999; accepted 20 August 1999)

Voice onset times (VOT) and burst frequencies of two aspirated (i.e., /k^h/, /g^h/) and two unaspirated (i.e., /k/, /g/) Gujarati velar stop consonants were investigated in an effort to provide characteristic acoustic information. Stop consonants in a monosyllabic vowel-consonant-vowel (VCV) production were obtained from eight native speakers of Gujarati. Differences in mean VOT and burst frequency as a function of voicing and aspiration were examined. A significant voicing by aspiration effect was found for VOT ($p=0.026$). The two voiced stops, while not significantly different from each other ($p=0.278$), had significantly shorter VOTs than voiceless stops. The aspirated /k^h/ had a significantly longer VOT than the unaspirated /k/ ($p=0.0013$). With respect to burst frequency, voiced stops had significantly higher burst frequencies than voiceless stops ($p=0.002$). There was no significant difference between mean burst frequencies of aspirated and unaspirated stops ($p=0.058$). © 1999 Acoustical Society of America. [S0001-4966(99)00312-4]

PACS numbers: 43.70.Fq, 43.70.Kv [AL]

INTRODUCTION

It has been shown that phoneme categories for stop consonants can be differentiated by differences in the temporal aspects of the onset of the periodic laryngeal vibration or glottal pulsing and the articulatory events associated with the release of the consonant burst (Fant, 1969; Ladefoged, 1996; Lieberman, 1977). This period, between the release burst of the stop sound and the onset of vibration of the vocal folds, is generally referred to as voice onset time (VOT). While VOT is quite effective in separating phonemic categories in most two- and three-category languages, it is inadequate in demarcating stop-consonant contrasts in four-category languages such as Hindi and Marathi (Ladefoged and Maddieson, 1996; Lisker and Abramson, 1964). A four-category language is one in which there are four stops at one place of articulation (e.g., four velars, four dentals, and so on). In these languages, aspiration is used in conjunction with the voicing element to produce four categories of stop consonants.

In a seminal cross-language study of initial stops in 11 languages, Lisker and Abramson (1964) reported VOT values for the voiced unaspirated and voiced aspirated stops in four-category Hindi and Marathi languages. Although there were orderly differences in mean values of VOT, there was an overlap in the ranges of variation in distributions of VOTs of the two voiced stops, and the two proved to be nearly similar. Lisker and Abramson indicated that the voiced aspirates differed from the other voiced category by the presence of a low-amplitude “buzz” assimilated with noise in the interval following stop release. It has also been suggested that aside from formant transitions, the burst frequency could also provide phonetically salient information (Halle, Hughes,

and Radley, 1957; Stevens, 1972, 1980) in these cases. The frequency of the release burst is determined by turbulent air-flow at the release of the stop.

The Gujarati language, similar to Hindi and Marathi, has four velar stops. Specifically, Gujarati contains two aspirated (i.e., /k^h/, /g^h/) and two unaspirated (i.e., /k/, /g/) velar stop consonants. Gujarati is spoken by 44 million people worldwide and is ranked twenty-third in the top 100 languages by population (Grimes, 1996). While the linguistic percept of these sounds to the native Gujarati speaker is apparent, information pertaining to the acoustic properties is unknown. While perceptual discrimination of the four velar Gujarati stops may be based on VOT distinctions, it remains possible that the information provided by the burst spectrum of each consonant might provide salient perceptual information as well. Toward that end, VOTs and the consonant-stop burst frequencies of the four velar Gujarati stop consonants: /k/, /g/, /k^h/, and /g^h/ were examined.

I. METHOD

A. Participants

Eight normal native speakers of Gujarati ($M=33.6$ years, $s.d.=18.5$; six males and two females) served as participants. While participants had learned Gujarati as their first language, all were bilinguals with English as their second language.

B. Apparatus

Speech samples were recorded in a quiet room with a microphone (Apple model Plaintalk) placed approximately 40 cm from lips with an orientation of 0° azimuth and 20° altitude. The microphone output was line-fed into a personal computer (Apple model Power Macintosh 7100/80) interfaced with an analog-to-digital input/output board (Digidesign model Audiomedia NuBus). Amplitude waveforms

^{a)} Author to whom correspondence should be addressed. Electronic mail: rastatterm@mail.ecu.edu

were generated by using a commercially available speech and sound signal analysis program (InfoSignal, Inc. model Signalyze 3.12). A fast Fourier transfer spectral analysis of the speech samples was also undertaken to obtain spectrograms and spectra that were used for acoustic analysis. The analog speech signals were digitally sampled at 44.1 kHz and quantized at 16 bit.

C. Procedure

Each participant produced 25 utterances of each of the four velar stop consonants of Gujarati (i.e., /k/, /g/, /k^h/, and /g^h/ in serial order) in a single test session. These consonants were recorded in a carrier phrase using a neutral schwa before and after the target consonant (i.e., have a VCV tʃ^he). Each utterance was produced after a normal breath with a pause before the target sound to decrease the effects of varying lung volume on VOT (Hoit, Soloman, and Hixon, 1993). Participants successively produced the phrases for each stop.

VOTs were measured from the amplitude waveforms of each stop. VOT was defined as the duration between the consonantal burst and the first glottal cycle (Lisker and Abramson, 1964). Burst frequency was measured from the spectra of each consonant. Spectra were obtained by placing a 10-ms window centered at the point of highest amplitude in the burst. The burst frequency was defined as the peak with the highest amplitude (Stevens, 1980).

Ten percent of the data were reanalyzed in an effort to provide an index of intrajudge and interjudge agreement. Intrajudge agreement for VOT and burst frequency, as measured by Pearson Product correlation, was 0.98 ($p < 0.001$) and 0.99 ($p < 0.001$), respectively. The mean standard error was also used to index intrajudge reliability (Ferguson, 1976). The intrajudge mean standard errors for VOT and burst frequency measurements were 0.2 ms and 2.19 Hz, respectively. Interjudge agreement was evaluated with a second independent judge unaware of the purpose of the study. Pearson Product correlations indexing interjudge agreement were 0.97 ($p < 0.001$) and 0.98 ($p < 0.001$) for VOT and burst frequency, respectively. Interjudge mean standard errors of measurement proved to be 0.39 ms for VOT and 2.97 Hz for the burst frequency, respectively.

II. RESULTS

Mean individual values of VOT and burst frequency were obtained from each participants' 25 of the four velar stop-consonant tokens. Grand means and standard deviations of VOT and burst frequency for the four velar stop consonants of all the participants are presented in Table I.

A repeated two-factor analysis of variance (ANOVA) was used to assess differences in mean VOT as a function of voicing and aspiration. Significant main effects were found for voicing [$F(1,7) = 153.48$, $p < 0.0001$, $\omega^2 = 0.94$] and aspiration [$F(1,7) = 13.86$, $p < 0.007$, $\omega^2 = 0.58$]. As well, a significant voicing by aspiration interaction was observed [$F(1,7) = 7.99$, $p = 0.026$, $\omega^2 = 0.43$]. This interaction is depicted in Fig. 1. Two orthogonal signal-*df* comparisons were undertaken to investigate the voicing by aspiration interaction. It was found that the mean VOT for the aspirated voice-

TABLE I. Grand means and standard deviations of voice onset time (ms) and burst frequency (Hz) for the four velar Gujarati stop consonants ($n = 8$). Standard deviations are presented in parentheses.

	Velar stop consonant			
	Voiced		Voiceless	
	/g/	/g ^h /	/k/	/k ^h /
Voice onset time	-37.3 (9.8)	-29.2 (7.3)	40.6 (21.1)	75.0 (33.2)
Burst frequency	1555 (198)	1531 (176)	1451 (146)	1427 (140)

less stop, /k^h/, was significantly longer than the unaspirated voiceless stop /k/ [$F(1,7) = 26.8$, $p = 0.0013$]. VOTs for voiced stops /g^h/ and /g/, although significantly shorter than the voiceless stops, were not statistically different from one another [$F(1,7) = 1.379$, $p = 0.278$].

Differences in mean burst frequency as a function of voicing and aspiration were also examined with a repeated two-factor ANOVA. Results showed a significant main effect of voicing [$F(1,7) = 25.2$, $p = 0.002$, $\omega^2 = 0.729$]. The main effect of aspiration [$F(1,7) = 5.14$, $p = 0.058$, $\omega^2 = 0.315$] and the interaction between voicing and aspiration [$F(1,7) = 0.005$, $p = 0.947$, $\omega^2 = 0$] were not significant. In other words, voiced stops had significantly higher burst frequencies than voiceless stops. There was no significant difference between mean burst frequencies of the aspirated and unaspirated stops.

III. DISCUSSION

The findings of this study suggest that the voicing feature of velar stop production in Gujarati parallels that of the two four-category languages of Hindi and Marathi (Lisker and Abramson, 1964). That is, the two voiced aspirated /g^h/ and unaspirated /g/ Gujarati stops reside on the negative side of the VOT continuum while the voiceless aspirated /k^h/ and unaspirated /k/ are located in the positive half of the VOT

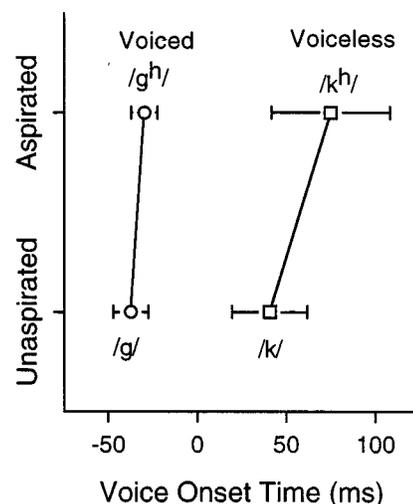


FIG. 1. Mean voice onset time as a function of voicing and aspiration for the four velar stop consonants in Gujarati. Error bars represent plus/minus one standard deviation of the mean.

continuum (see Table I and Fig. 1). Burst frequency seems to be an additional conspicuous acoustic property delineating the voiced and voiceless Gujarati stops. Statistically significant higher burst frequencies were associated with the voiced stops (see Table I). Burst frequencies above 1500 Hz were associated with voiced velar stop productions, while those below defined the voiceless velar stops. Given that distinctions exist in VOT and burst frequency between the voiced and voiceless Gujarati stop consonants, it is conjectured that information provided by these acoustic properties could contribute to the perceptual distinctiveness of voicing.

While VOT and burst frequency appear to be a conspicuous acoustic property delineating voicing, differences between the aspirated and unaspirated stops within each voicing category remain less clear. Although significant differences existed in VOT between the voiceless stops, such was not the case with the voiced stops. Compare mean differences in VOTs of approximately 35 and 8 ms for the voiceless and voiced stops, respectively. It may be the case that burst frequency provides an additional cue for differentiating the voiced aspirated /g^h/ and unaspirated /g/ Gujarati stops. Although differences in burst frequency were not significant, the effect size was large (Cohen, 1988). Higher burst frequencies were evident with the unaspirated voiced stops. Such a finding suggests that there may be a slight shift in point of constriction between the aspirated and unaspirated velar stops. Acoustical information contained in the burst frequency may provide additional salience for the discrimination of the aspirated versus unaspirated voiced velar stops in the Gujarati. It may also be the case that yet an undetermined acoustic property not explored in this study may be the conspicuous property that listeners cue into for perceptual distinctiveness. This speculation can only be explored in future acoustical and perceptual experiments.

In the case of Hindi, an acoustic and fiberoptic study by Benguerel and Bhatia (1980) has shown that there are differences in VOT for the voiceless aspirated and unaspirated and the voiced unaspirated stops. However, for the voiced aspirated the authors suggested that a third type of voicing known as murmur existed (i.e., the simultaneous presence of voicing and turbulence). Though similar data for Gujarati are currently unavailable, it is conjectured that a murmur may be found in the Gujarati voiced aspirated stop and could be an additional salient acoustic property that aids listeners in distinguishing the unaspirated from the aspirated stops.

In conclusion, the current findings show systematic acoustic differences among VOTs and burst frequencies of the four Gujarati velar stop consonants. It appears that these differences may be the conspicuous acoustic properties that contribute to the perceptual discrimination of voicing. In ad-

dition, VOT may distinguish aspiration for the voiceless velar stops /k^h/ and /k/. Additional information provided by burst frequency may be used to make aspiration distinctions between the voiced stops aspirated /g^h/ and unaspirated /g/ as the VOT continuum overlaps for these consonant pairs. Unlike the general phonetic types of two- and three-category languages, the four-category languages seem to be employing attributes of alternate acoustic events for determining phonetic boundaries, such as information provided by the burst frequency of the consonant as found in the current study. It should be pointed out that it is difficult to comment about the exact voice source characteristics and the laryngeal control as has been done in the case of Hindi aspirated voiced stops (Benguerel and Bhatia, 1980) until similar data from other places of articulation (e.g., palatal, alveolar, dental, etc.) in Gujarati are available. Research employing techniques designed to observe perceptual and physiological differences among the velar Gujarati aspirated and unaspirated consonant pairs will also help answer these questions more completely.

ACKNOWLEDGMENTS

The authors thank and acknowledge the thoughtful comments of Dr. Arthur S. Abramson, Dr. Peter Ladefoged, and Dr. Anders Löfqvist on an earlier version of this paper. Presented in part at the North Carolina Speech, Language, and Hearing Association Annual Convention, Charlotte, NC, 24 April 1998.

- Benguerel, A.-P., and Bhatia, T. K. (1980). "Hindi stop consonants: an acoustic and fiberoptic study," *Phonetica* **37**, 134–148.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Erlbaum, Hillsdale, NJ, 1988).
- Fant, G. (1969). "Stops in CV syllables," in *Speech Transmission Labs Quarterly Progress Reports, 41* (Royal Institute of Technology, Stockholm).
- Ferguson, G. A. (1976). *Statistical Analysis in Psychology and Education*, 4th ed. (McGraw-Hill, New York).
- Grimes, B. F. (Ed.). (1996). *Ethnologue: Languages of the World*, 13th ed. (Summer Institute of Linguistics, Dallas).
- Halle, M., Hughes, G. W., and Radley, J. P. A. (1957). "Acoustic properties of stop consonants," *J. Acoust. Soc. Am.* **29**, 107–116.
- Hoit, J. D., Soloman, N. P., and Hixon, T. J. (1993). "Effect of lung volume in voice onset time (VOT)," *J. Speech Hear. Res.* **36**(3), 516–520.
- Ladefoged, P. (1996). *Elements of Acoustic Phonetics*, 2nd ed. (University of Chicago Press, Chicago).
- Ladefoged, P., and Maddieson, I. (1996). *The Sounds of the World's Languages* (Blackwell, Cambridge, MA).
- Lieberman, P. (1977). *Speech Physiology and Acoustic Phonetics: An Introduction* (Macmillan, New York).
- Lisker, L., and Abramson, A. S. (1964). "A cross language study of voicing in initial stops: Acoustical measurements," *Word* **20**, 384–422.
- Stevens, K. N. (1972). "Quantal nature of speech," in *Human Communication: A Unified View*, edited by E. E. David, Jr. and P. B. Denes (McGraw-Hill, New York).
- Stevens, K. N. (1980). "Acoustic correlates of some phonetic categories," *J. Acoust. Soc. Am.* **68**, 836–842.